

UNIVERSITE PARIS IX- DAUPHINE
UFR SCIENCES DES ORGANISATIONS

**CONCURRENCE EN ASSURANCE SANTE :
ENTRE EFFICACITE ET SELECTION**

THESE
pour l'obtention du titre de

Docteur en Sciences Economiques
(arrêté du 30 mars 1992)

Présentée et soutenue par

Agnès COUFFINHAL

JURY

Directrice de Thèse

Lise ROCHAIX-RANSON
Professeur à l'Université de la Méditerranée

Rapporteurs

Louis EECKHOUDT
Professeur aux Facultés Universitaires Catholiques de Mons

Carol PROPPER
Professeur à l'Université de Bristol

Jean-Charles ROCHET
Professeur à l'Université de Toulouse I

Suffragants

Louis-André GERARD-VARET
Directeur d'Etudes à l'EHESS, G.R.E.Q.A.M.

Pierre-Marie LARNAC
Professeur à l'Université Paris IX - Dauphine

Emile LEVY
Professeur à l'Université Paris IX - Dauphine

Christine MEYER
Responsable du Département Assurance Maladie
Fédération Nationale de la Mutualité Française

Septembre 1999

"L'Université n'entend donner aucune approbation ni improbation aux opinions émises dans les thèses : ces opinions doivent être considérées comme propres à leurs auteurs".

REMERCIEMENTS

En premier lieu, je tiens ici à exprimer la gratitude que j'éprouve envers Lise Rochaix-Ranson. Sa rigueur intellectuelle et sa conception infiniment généreuse du rôle d'un directeur de thèse m'ont donné l'envie d'entamer ce parcours, l'énergie nécessaire pour le mener à terme et le désir de continuer à travailler avec elle aussi longtemps qu'elle le voudra bien.

Je tiens aussi à remercier du fond du cœur Christine Meyer et Dominique Polton qui ont, en premier lieu, joué un rôle central dans le financement de cette recherche. Ensuite, grâce à elles, j'ai pu découvrir l'univers de la santé qui m'était largement étranger. Mais surtout, leur exigence et leur ouverture d'esprit, leur approche rigoureuse des problématiques réputées techniques ou politiques, m'ont encouragée à faire le lien entre théorie et pratique et ont fait naître chez moi le désir de continuer dans cette voie.

Sans le Professeur Emile Levy et Jean de Kervasdoué, rien n'aurait été possible. Leur bienveillance à mon égard qui s'est manifestée concrètement à maintes reprises et leurs conseils éclairés m'ont été infiniment précieux.

Pour ses encouragements, son aide, mais aussi son accueil chaleureux à Toulouse, je remercie vivement le Professeur Jean-Charles Rochet.

Je tiens aussi à remercier les Professeurs Louis-André Gérard-Varet, Louis Eeckhoudt, Pierre-Marie Larnac et Carol Propper, qui ont accepté de faire partie du Jury.

Mes collègues de SANESCO et du CREDES, dont certains sont devenu des amis, m'ont soutenu avec constance et m'ont aidé à mener de front la recherche et le travail. Ils ont en outre largement contribué à faire avancer l'une et l'autre. Je tiens particulièrement à remercier ici, par ordre d'apparition, Frédéric Bousquet, Raphaël Spira et Michel Grignon.

Je suis aussi redevable de l'aide de nombreuses personnes, auxquelles je tiens à exprimer ici ma reconnaissance : Didier Blanchet, Pascale Breuil, les dauphinois, notamment Philippe Bernard et Jérôme Wittwer, Farid Gasmi, Geneviève Gégou, Emmanuelle Guihéneuf, Agnès Gramain, Laurence Hartmann, le Professeur Claude le Pen, Ronan Mahieu, Edward Norton, Frédéric Rupprecht, Véronique Seltz, Maître Bernard Servet et toutes les personnes qui ont accepté le travail, ô combien ingrat, de relire les différentes versions de cette thèse. Je suis malheureusement certaine de ne pas citer toutes les personnes qui m'ont apporté leur soutien à un moment ou un autre durant ces cinq années, mais j'espère que qu'elles accepteront mes remerciements chaleureux et sincères.

Enfin, et peut être surtout, merci tous mes proches, dont certains ont déjà été cités :

- les amis de toujours ou de longtemps qui ont su pardonner mes disparitions et qu'il me tarde de revoir autrement qu'entre deux portes,
- ceux que j'ai rencontré pendant cette période mouvementée et qui n'ont pas été effrayés,
- mes parents et ma famille,

Ils m'ont, tour à tour, soutenue, encouragée, supportée et distraite... Bruce a fait tout cela et encore plus. Tous les jours. J'espère que je pourrai le leur rendre au centuple.

Je tiens à avoir une pensée à la mémoire de mon parrain et de Mimi et une autre pour tous ceux de mes proches que je n'ai pas pu soutenir autant que je l'aurais souhaité dans les épreuves qu'ils traversent.

A Bruce,

A Mum et Tatakoko.

SOMMAIRE

INTRODUCTION GENERALE	9
------------------------------------	----------

PARTIE I : LE SERVICE D'ASSURANCE EN SANTE

INTRODUCTION	15
---------------------------	-----------

CHAPITRE 1 : PRINCIPES FONDAMENTAUX DE L'ASSURANCE	17
---	-----------

SECTION 1 : FORMALISATION DU RISQUE	18
--	-----------

SECTION 2 : PARTAGE DU RISQUE INDIVIDUEL	22
---	-----------

SECTION 3 : PARTAGE DU RISQUE SOCIAL.....	31
--	-----------

SECTION 4 : LE SERVICE D'ASSURANCE.....	40
--	-----------

CONCLUSION	47
-------------------------	-----------

CHAPITRE 2 : LE RISQUE ET LA REPARATION DU DOMMAGE EN SANTE	48
--	-----------

SECTION 1 : FACTEURS DE RISQUE ET NATURE DE L'ALÉA EN SANTÉ	49
--	-----------

SECTION 2 : DÉTERMINANTS DE L'UTILISATION DES SOINS	64
--	-----------

CONCLUSION	74
-------------------------	-----------

CHAPITRE 3 : DECLINAISON A LA SANTE DES PRINCIPES ASSURANTIELS	75
---	-----------

SECTION 1 : OBJET DE L'ASSURANCE SANTE.....	76
--	-----------

SECTION 2 : PRISE EN COMPTE DES COMPORTEMENTS STRATÉGIQUES PAR L'ASSUREUR	81
--	-----------

SECTION 3 : TARIFICATION DE L'ASSURANCE.....	92
---	-----------

CONCLUSION	101
-------------------------	------------

CONCLUSION DE LA PARTIE I.....	102
---------------------------------------	------------

PARTIE II : DE L'ANTISELECTION A LA SELECTION :
LES LIMITES A L'EFFICACITE

INTRODUCTION	104
CHAPITRE 4 : ASSURANCE PRIVEE ET ANTISELECTION : APPROCHES THEORIQUES	106
SECTION 1 : OPTIMUM ET ÉQUILIBRE AVEC ANTISÉLECTION.....	108
SECTION 2 : ANTISÉLECTION ET INTERVENTION DE L'ÉTAT	129
CONCLUSION	150
CHAPITRE 5 : ASSURANCE SANTE ET ANTISELECTION : APPROCHES EMPIRIQUES	151
SECTION 1 : LA SEGMENTATION DES RISQUES	152
SECTION 2 : LES PREDICTIONS DU MODÈLE DE ROTHSCHILD ET STIGLITZ	160
SECTION 3 : L'INTERVENTION DE L'ÉTAT.....	176
CONCLUSION	180
CHAPITRE 6 : LA SELECTION EN ASSURANCE SANTE	182
SECTION 1 : ACCES A L'ASSURANCE ET REVENU	184
SECTION 2 : ACCES A L'ASSURANCE ET SELECTION DIFFÉRENTIELLE	200
CONCLUSION	223
CONCLUSION DE LA PARTIE II.....	224

PARTIE III : CONCURRENCE ENTRE ASSUREURS
DANS LES SYSTEMES DE SANTE

**CHAPITRE 7 : FONCTIONS DE L'ASSURANCE PRIVEE
DANS LES SYSTEMES MIXTES 230**

SECTION 1 : UNE APPROCHE METHODIQUE DES ARTICULATIONS PUBLIC-PRIVÉ.....232

SECTION 2 : DÉCLINAISON DE LA TYPOLOGIE FONCTIONNELLE DE L'ASSURANCE PRIVÉE.....243

CONCLUSION269

**CHAPITRE 8 : UNE CONCURRENCE ENTRE ASSUREURS
POUR L'EFFICACITE PRODUCTIVE..... 271**

SECTION 1 : RÉGLEMENTATION DE LA CONCURRENCE : ENJEUX ET PRINCIPES273

SECTION 2 : LA PRATIQUE DE L'AJUSTEMENT AU RISQUE297

CONCLUSION DE LA PARTIE III..... 340

CONCLUSION GENERALE346

INTRODUCTION GENERALE

INTRODUCTION GENERALE

Les années récentes ont vu s'intensifier le recours à des mécanismes de marché, y compris en Europe, par des pays qui cherchent à améliorer l'efficacité de leurs systèmes de santé. Certaines réformes ont explicitement visé à introduire ou à encourager la concurrence dans l'assurance santé. Or, dans le même temps, la quasi-totalité des pays développés a cherché à maintenir ou à étendre l'accès à une couverture maladie.

On analyse traditionnellement la tendance à l'universalisation de la couverture santé comme la traduction d'une volonté politique d'améliorer la justice sociale. Dans cette perspective, en dehors de toute référence à des préférences individuelles, un certain nombre de principes intangibles déterminent les droits fondamentaux de tout membre de la société. L'Etat peut et doit intervenir pour que ces droits soient respectés. La recherche d'équité peut donc conduire à amender le principe de consentement à payer (*willingness to pay*), qui repose sur la capacité de payer (*ability to pay*) et indirectement sur la distribution des richesses. Si l'accès à certains biens permet d'établir ces droits fondamentaux, et si la capacité de payer ces biens est jugée inadéquate au regard des critères énoncés, l'Etat peut agir pour que leur distribution ne soit pas seulement fondée sur le consentement à payer.

Les soins de santé sont, parmi d'autres, des biens considérés comme relevant de ce que Tobin (1970) qualifie de principe "*d'égalitarisme spécifique*". L'assurance, qui permet à des personnes exposées à des risques d'opérer ex ante une redistribution de leur propre richesse entre différents états du monde, ne permet pas de corriger des inégalités qui peuvent être jugées socialement inacceptables dans l'accès aux soins. Ces inégalités sont de deux ordre :

- les personnes dont l'état de santé est dégradé doivent, pour se soigner ou pour s'assurer, payer plus que celles qui sont en bonne santé. Comme le soulignent van de Ven et van Vliet (1992) "*a chronically ill person would pay, say, one hundred times*

the premium of a chronically well person of the same age. In most societies that is considered undesirable¹.

- les personnes moins aisées, pour un niveau de risque donné, doivent consacrer une part plus grande de leurs ressources pour obtenir une couverture ou pour avoir accès à un traitement donné.

Il est, en réalité, difficile de distinguer ce qui est injuste au regard de l'un ou l'autre critère car la capacité d'une personne à payer des soins ou une assurance dépend de son niveau de risque relativement à son revenu. En outre, ces inégalités se conjuguent car il existe une corrélation inverse entre niveau de risque, et richesse ou classe sociale. Si l'Etat n'opère pas de redistribution, les plus pauvres se voient, en moyenne, obligés de consacrer plus de ressources à se soigner ou à se couvrir que les personnes les plus aisées.

La volonté de maintenir ou d'instaurer une assurance santé universelle obéit indéniablement à une logique de justice sociale. Sans chercher à méconnaître cette évidence, notre travail, résolument centré sur une analyse en terme d'efficacité, conforte l'idée que l'utilisation de mécanismes concurrentiels visant à améliorer l'efficacité des systèmes d'assurance maladie n'est pas incompatible avec le respect de principes d'équité.

Pour autant, il apparaît paradoxal de faire peser des contraintes fortes sur les choix des individus en généralisant la couverture maladie et, dans le même temps, de recourir à des mécanismes concurrentiels dont le succès même repose sur les arbitrages réalisés librement par des agents économiques rationnels. Lorsque la couverture est obligatoire, uniforme et socialisée, on peut s'attendre à ce que l'efficacité soit moindre que dans un système où les individus sont libres de s'assurer et de faire jouer pleinement la concurrence entre assureurs pour obtenir des services coût-efficaces.

¹ *Un malade chronique paierait, environ, cent fois la prime d'une personne en bonne santé chronique du même âge. Dans la plupart des sociétés, cet état de fait est jugé indésirable.*

L'objectif de cette thèse est de montrer que ce paradoxe n'est qu'apparent. Pour ce faire, il importe de considérer séparément efficacité allocative et productive. En particulier, on montrera que, sur le marché de l'assurance santé, l'organisation d'une concurrence porteuse d'efficacité productive nécessite de limiter les choix des assurés et des patients. Autrement dit, on relativise l'importance de l'arbitrage efficacité-équité dans le débat sur la concurrence pour mettre en lumière les termes celui qui prévaut entre efficacité allocative et efficacité productive. Nous soulignerons alors que la tendance à l'universalisation, qui relève en priorité d'une logique d'équité, s'explique, aussi, par la recherche de l'efficacité productive.

Sur le plan méthodologique, notre cadre de référence et le point de départ de nos raisonnements sont la théorie de l'assurance, dont nous explorons en détail certains aspects pertinents au regard de notre sujet. Cette démarche nous amène à proposer des clarifications et des extensions de certains modèles.

Parallèlement, nous cherchons à déterminer en quoi ces modèles permettent de rendre compte du fonctionnement de l'assurance santé. Pour ce faire, nous nous appuyons sur la littérature empirique disponible. Nonobstant la difficulté d'obtenir des informations de qualité, l'intérêt de procéder à un travail économétrique spécifique sur données françaises, dans notre perspective, aurait été limité par le fait que l'assurance privée joue, en France, un rôle marginal dans la couverture des risques liés à la santé.

En outre, la littérature empirique ne permettant jamais que d'éclairer un aspect particulier, dans un contexte institutionnel précis et à une période donnée, il nous semble intéressant d'analyser différents systèmes de santé dans lesquels l'assurance privée ou la concurrence entre assureurs est présente. Comme le soulignent Björkman et Altenstetter (1997), ces comparaisons internationales ont un objectif principal : *"to extract lessons, or more appropriately cautions and caveats about easily voiced proposals for change versus the harsh realities of politics and the inertial constraints of bureaucratic organizations"*² (pp.1). Au delà d'une perspective d'analyse en terme de politique économique, l'observation du fonctionnement des marchés permet d'alimenter

² *D'extraire des leçons, ou de façon plus appropriée, des précautions et des mises en gardes sur des propositions de réformes qu'il est facile de formuler mais qui se heurtent aux dures réalités de la politique et de l'inertie des organisations bureaucratiques.*

la réflexion théorique en conduisant à reformuler certaines interrogations et à ouvrir de nouvelles pistes de recherche.

Notre démonstration s'articule autour de trois parties.

La première vise à définir le service d'assurance en santé. Il apparaît en effet nécessaire, pour savoir quel rôle peut jouer la concurrence, de déterminer précisément les caractéristiques du service qu'un assureur peut rendre, en dehors de tout contexte institutionnel, mais en prenant en compte la nature des risques qu'il couvre et les caractéristiques du marché des soins sur lequel il intervient en tant que tiers-payeur. Nous montrons notamment comment les assureurs peuvent avoir intérêt à intervenir de façon active dans la gestion des soins et à améliorer l'efficacité de la prise en charge des personnes qu'ils couvrent.

Dans la seconde partie de ce travail, nous nous tournons vers l'étude du fonctionnement du marché du service ainsi défini et nous mettons en lumière les conditions sous lesquelles la concurrence n'est pas source d'efficacité. Nous abordons cette analyse en partant des modèles standards de la théorie de l'assurance, qui montrent qu'une asymétrie d'information sur le risque en faveur des assurés peut conduire à la disparition du marché, à cause de l'antisélection. Tout en reconnaissant que ce phénomène semble expliquer pour partie la dynamique du marché de l'assurance santé, nous sommes progressivement amenés à développer une analyse théorique qui montre comment une concurrence non réglementée, sur ce marché, génère des formes de sélection, par le risque et par le revenu. Celles-ci limitent fortement les bénéfices attendus de la concurrence en terme d'efficacité productive. Etant donnée la priorité accordée à la maîtrise des dépenses de santé, il apparaît logique que la plupart des réformes des systèmes d'assurance santé s'articulent, explicitement ou non, autour de la régulation des pratiques de sélection.

Dans la troisième partie, nous procédons à une analyse comparative des systèmes de santé et de leurs réformes. Elle nous conduit tout d'abord à mettre en évidence que l'assurance privée joue un rôle plus marginal dans les systèmes de santé que ce que les conclusions de la première partie n'auraient laissé entrevoir. L'assurance volontaire reste fondamentalement complémentaire de l'assurance publique en remplissant des fonctions différentes selon les caractéristiques de ce service public. En fait, les réformes ou les

évolutions spontanées des systèmes d'assurance traduisent l'émergence d'un consensus autour d'un mode de gestion particulier de la concurrence : celle-ci ne porte pas tant sur la fonction d'assurance que sur la gestion du risque. Une analyse détaillée de ces réformes confirme bien que, pour promouvoir l'efficacité productive, l'organisation de la concurrence nécessite de réaliser un arbitrage au détriment de l'efficacité allocative, ce qui conforte le caractère marginal de l'assurance volontaire.

PARTIE I

LE SERVICE D'ASSURANCE EN SANTE

PARTIE I

LE SERVICE D'ASSURANCE EN SANTE

"Sécurité sociale : le mythe de l'assurance" (Dufourq, 1994), *"Sécurité sociale : pour l'assurance"* (Caussat, 1994), *"la référence assurantielle en matière de protection sociale : apport et limites"* (Blanchet, 1996)... sont quelques exemples d'articles aux titres évocateurs qui soulignent le regain d'intérêt suscité par la réflexion sur la nature des systèmes de protection sociale mis en place ou généralisés au vingtième siècle. Au cours des vingt dernières années, l'augmentation rapide des dépenses de santé et les difficultés croissantes des systèmes de protection sociale à faire face à leurs engagements, dans un contexte économique difficile, ont conduit à ce que les vertus de la concurrence soient de plus en plus systématiquement mises en avant. Dans un tel contexte, l'étude de la nature exacte du service qui est, ou qui peut être rendu par l'assurance privée, est donc centrale. En effet, si l'on parvient à déterminer ce qui relève de l'assurance et de la solidarité, ce qui peut être mis en œuvre de façon plus efficace par le privé ou le public, on peut espérer, sur la base d'arbitrages démocratiques, recentrer l'intervention de l'Etat sur ce qui est jugé nécessaire et, tout en maintenant un degré de solidarité nationale, améliorer la gestion globale des systèmes de santé.

L'objectif de cette première partie est de proposer une réflexion économique sur la nature du service qui peut être rendu par l'assurance dans le domaine de la santé. Pour ce faire, nous adoptons une démarche en trois temps. Dans un premier chapitre, nous rappelons les principes fondamentaux de la théorie de l'assurance. Dans le deuxième chapitre, nous étudions la nature des risques auquel l'individu est confronté pour ce qui concerne sa santé et la façon dont il détermine les moyens à mettre en œuvre lorsqu'il subit un dommage. Enfin, dans le chapitre 3, nous proposons une synthèse qui montre en quoi les caractéristiques des biens étudiées, soins et santé, ont un impact sur la façon dont l'assurance est mise en œuvre.

Nous cherchons dans cette partie à décrire le service d'assurance. Nous mettons donc l'accent sur les relations bilatérales qui s'établissent entre des personnes qui souhaitent se couvrir contre la réalisation de dommages et les producteurs de ce service. Implicitement, nous étudions plutôt l'assurance dans la mesure où elle résulte de transactions volontairement consenties entre des individus et des producteurs privés, en concurrence ou non entre eux. Il faut cependant remarquer que certains aspects que nous évoquons, comme le risque moral, sont susceptibles de concerner aussi un assureur public.

En revanche, nous écartons de cette première partie l'antisélection qui émerge lorsque les assurés détiennent des informations sur leur niveau de risque auxquelles les assureurs n'ont pas accès. En effet, l'antisélection, contrairement au risque moral, soulève surtout des difficultés dans le cadre d'un marché sur lequel les assureurs sont en concurrence : elle peut entraîner la disparition du marché. L'existence d'antisélection a évidemment des conséquences sur la façon dont les transactions entre assureurs et assurés se déroulent, mais l'étude de ce phénomène nous semble plus naturellement relever d'une analyse au niveau du marché. Nous le traiterons donc dans la seconde partie de ce travail qui est consacrée à l'étude de la concurrence entre assureurs en santé.

CHAPITRE 1

PRINCIPES FONDAMENTAUX DE L'ASSURANCE

L'univers dans lequel les agents économiques évoluent est marqué par une grande incertitude. En fonction de leurs goûts, les individus peuvent s'exposer ou non à certains risques s'ils estiment pouvoir en tirer une utilité directe (en pratiquant un sport réputé dangereux) ou un revenu plus élevé (en investissant sur les marchés financiers). Toutefois, une partie de l'incertitude à laquelle les individus sont exposés est largement indépendante de leurs choix individuels et se traduit par la réalisation de dommages. Le fait de subir un vol ou d'avoir un accident sont des événements largement imprévisibles qui diminuent l'utilité de l'individu. Dès lors qu'il a conscience de l'existence de ce type d'aléa et s'il éprouve de l'aversion pour le risque, l'individu va chercher à se prémunir contre la réalisation de ces sinistres ou, du moins, à en compenser les effets.

La théorie économique décrit un ensemble de mécanismes qui permettent de se couvrir contre ces aléas ou de répartir de façon optimale leurs conséquences en fonction de la nature des risques. L'objectif de ce chapitre est de présenter les principaux enseignements de la théorie économique concernant les mécanismes de couverture contre des risques, en particulier l'assurance. Leur présentation nous conduira :

- partant de la reconnaissance du fait que les dotations d'un individu peuvent être aléatoires, à distinguer deux types de risques : le risque social et le risque individuel (Section 1) ;
- à présenter les mécanismes théoriques optimaux de partage de ces deux types de risques (Section 2 et Section 3). Nous soulignerons que l'assurance n'est qu'un des mécanismes possibles de partage du risque individuel et nous expliquerons en quoi il est légitime de le placer au centre de notre analyse ;
- enfin, nous étudierons le fonctionnement des marchés d'assurance, en particulier quand l'individu peut non seulement s'assurer, mais aussi diminuer le risque auquel il est exposé (Section 4).

1. FORMALISATION DU RISQUE

L'extension du modèle d'équilibre général aux économies avec incertain a été initiée par Arrow (1953, 1964) et Debreu (1959) auxquels on doit le développement des concepts de bien et de prix contingents (1.1). Leur analyse conduit à distinguer deux types de risques : le risque social et le risque individuel (1.2).

1.1 Bien et prix contingents

Le bien contingent est défini par ses caractéristiques physiques, la date à laquelle il est disponible, mais aussi par les conditions dans lesquelles il est disponible. Plus précisément, la quantité disponible de ce bien dépend de la réalisation d'une variable aléatoire exogène. Le prix de ce bien contingent est le montant payé par (à) l'agent qui s'engage à accepter (livrer) une unité du bien considéré (Debreu, 1959).

Les individus, plus ou moins bien dotés en biens selon l'état du monde qui se réalise, procèdent à des échanges sur des marchés contingents. Les prix se fixent de manière à ce que la demande excédentaire sur chaque marché contingent soit nulle. Dans un univers incertain, comme dans le cas général, l'application des principes de la théorie économique conduit à montrer que si un équilibre de marché concurrentiel existe et si les prix des biens contingents sont déterminés par les marchés, alors, pour une distribution initiale donnée des ressources, l'équilibre est optimal au sens de Pareto. Autrement dit, étant donnée l'allocation initiale des ressources, il n'est pas possible d'améliorer la situation de certains individus sans nuire à d'autres. Ce lien entre concurrence et optimalité repose sur la logique selon laquelle :

- 1 le bien être social est maximisé quand les utilités individuelles sont maximisées ;
- 2 les utilités individuelles sont maximisées quand les personnes peuvent choisir librement leurs allocations de ressources ;
- 3 et donc, le bien être social est maximisé quand les personnes sont libres de choisir.

Plus généralement, la notion de bien contingent permet d'étendre aux économies avec incertitude les principes de l'équilibre général et les théorèmes fondamentaux de la théorie du bien-être¹: il est possible d'associer un équilibre à tout optimum et un équilibre est un optimum.

Sans perte de généralité (cf. Arrow, 1953, 1964, et Drèze, 1971), on analyse dans ce chapitre l'allocation d'un bien unique (richesse ou consommation des individus) et on se limite au cadre d'une économie d'échange. En effet, la production, dans l'acception traditionnelle du terme, n'a pas de lien direct avec la problématique de partage du risque en santé.

1.2 Risque individuel et risque social

La définition de l'incertitude à laquelle les agents économiques font face a été progressivement affinée pour laisser apparaître une distinction fondamentale entre le risque individuel et le risque social. Arrow et Lind (1970) étudiant les décisions d'investissements publics introduisent pour la première fois l'idée qu'il faut différencier deux cas : "*investments where the benefits and costs varied greatly, depending on the state of nature, but where the difference between total benefits and total costs was constant for every state*" (pp. 418) et les autres. Ils établissent donc une distinction entre les investissements selon que la réalisation d'événements aléatoires conduit ou non à la variation de l'espérance de richesse totale. Lorsque la richesse globale ne varie pas, le risque est individuel (cela serait par exemple le cas si x entreprises, de même taille, et désignées de façon aléatoire, faisaient faillite tous les ans). Lorsque la richesse globale varie, le risque est dit social (un événement aléatoire se réalise qui a pour conséquence que toutes les entreprises font faillite une année).

Le modèle de Malinvaud (1972, 1973) repris et complété par Cass, Chichilnisky et Wu (1991, 1996) formalise la coexistence, dans une économie, de risques individuels et sociaux. Le cadre d'analyse présenté dans les paragraphes suivants nous servira de référence.

¹ En faisant l'hypothèse traditionnelle de convexité des ensembles de consommation.

² *Les investissements pour lesquels les gains et les pertes varient fortement d'un état du monde à l'autre mais pour lesquels la différence entre les coûts et les gains totaux est constante.*

1.2.1 Définition de l'aléa au niveau individuel

L'économie est composée de N individus.

Chaque individu peut se trouver dans S états du monde, mutuellement exclusifs, numérotés $(1, \dots, s, \dots, S)$. La réalisation de l'état s pour un individu i , lui donne droit à une allocation w_s , et la probabilité π_s de réalisation de cet état du monde est la même pour tous les individus.

Autrement dit, dans ce modèle :

- l'aléa est défini au niveau individuel et les risques sont indépendants,
- l'espérance de richesse est la même pour tous les individus (condition dite d'identique exposition au risque).

1.2.2 Distinction des états collectifs et des états sociaux

Un état collectif est défini par le vecteur des dotations des N individus. Il y a donc S^N états collectifs E . On peut établir une partition de ces états collectifs en ce que l'on appelle des états sociaux. Un état social Ω est l'ensemble des états collectifs qui aboutissent à la même allocation agrégée. Chaque état social est donc l'ensemble des états collectifs qui sont caractérisés par un même nombre d'individus dans chacun des états S $(n_{1/\Omega}, n_{2/\Omega}, \dots, n_{S/\Omega})$ avec, par construction :

$$\sum_i n_{i/\Omega} = N \quad \forall \Omega$$

Notons que :

- Pour N personnes et S états du monde individuels, il y a $\binom{N+S-1}{S-1}$ états sociaux du monde possibles.
- Chaque état social est un ensemble de $\frac{N!}{n_{1/\Omega}! n_{2/\Omega}! \dots n_{S/\Omega}!}$ états collectifs.

1.2.3 Distinction du risque social et du risque individuel

Pour illustrer la différence entre les deux types de risques sociaux et individuels, on peut considérer que deux tirages successifs sont effectués. Le premier tirage représente la réalisation du risque social, le second correspond celle du risque individuel.

1. Le premier détermine l'état social dans lequel l'ensemble des individus se trouve et donne la proportion des individus qui seront dans un état s . Il est associé à un niveau de richesse agrégée $W_\Omega = \sum_s n_{s,\Omega} w_s$.

2. Le second correspond au tirage des $n_{1,\Omega}$ personnes dans la population qui sont dans l'état 1 et qui obtiennent w_1 , puis des $n_{2,\Omega}, \dots, n_{S,\Omega}$ qui sont dans les états 2 à S parmi les N individus. A l'issue de ce second tirage chaque individu connaît sa dotation individuelle.

La probabilité pour chaque individu d'être dans l'état s , sachant que l'état social est Ω , peut être calculée par la formule de Bayes :

$$\pi_{s,\Omega} = \frac{\pi_{s,\Omega}}{\sum_s \pi_{s,\Omega}}$$

Etudier les mécanismes de partage du risque social conduit à analyser les arbitrages qu'effectuent les individus entre différents états sociaux (et donc différents niveaux de richesse agrégée). Etudier le partage du risque individuel revient, pour un état social donné, à étudier le comportement économique des individus dans l'ensemble des états collectifs qui correspondent à cet état social et à déterminer les arbitrages qu'ils effectuent face à un risque que l'on qualifiera désormais de risque individuel.

Nous étudierons séparément les deux cas : dans un premier temps nous nous penchons sur les mécanismes de partage du risque individuel, dont l'assurance fait partie, puis sur ceux de partage du risque social.

2. PARTAGE DU RISQUE INDIVIDUEL

La configuration de l'économie est telle que les individus sont soumis à un aléa et la dotation globale est constante. Par rapport au cadre présenté précédemment, on se situe dans un état social donné Ω . Le nombre d'individus dans un même état individuel s est donc connu :

$$n_{s,\Omega} = N \pi_{s,\Omega}$$

Les résultats suivants s'appliquent aux états sociaux dans lesquels tous les individus n'obtiennent pas la même allocation (dans le cas contraire, il n'y a pas de risque individuel). Afin d'alléger la notation, on considère donc que le premier tirage au sort a abouti à déterminer l'état Ω , et on ignore l'indice correspondant. En particulier, la probabilité notée maintenant π_s correspond en réalité à $\pi_{s,\Omega}$. La dotation globale est $W = \sum_s n_s w_s$, pour tout état collectif E .

Les préférences des individus vérifient l'axiomatique de von Neumann et Morgenstern. De plus, chaque individu a des préférences strictement monotones et strictement convexes. Autrement dit chaque individu éprouve de l'aversion pour le risque.

On s'interroge dans un premier temps sur la répartition optimale des ressources dans une telle économie (2.1) avant de décliner les mécanismes qui permettent de l'atteindre : les échanges marchands de biens contingents (2.2), la couverture réciproque (2.3) et l'assurance (2.4).

2.1 L'optimum

Deux conditions doivent être remplies pour atteindre l'optimum : l'allocation est faisable et il ne doit pas exister d'autre allocation qui permettrait d'augmenter l'utilité d'au moins un individu sans diminuer celle d'un autre.

On montre que l'optimum peut être caractérisé de la façon suivante :

- les consommations de chaque individu sont identiques, quel que soit l'état collectif. Autrement dit, en adoptant le point de vue d'un individu, la consommation ne dépend pas de la réalisation de la variable s ;

- la consommation de l'individu est égale à sa dotation moyenne sur l'ensemble des états du monde : dans le cas particulier d'identique exposition au risque, la consommation est la même pour tous les individus. Dans un cas plus général, la consommation de chacun est égale à sa dotation espérée.

Ces résultats sont indépendants du degré d'aversion au risque des individus. A partir du moment où leur utilité est concave, tous se couvrent intégralement contre le risque individuel.

2.2 L'équilibre des marchés Arrow-Debreu

L'allocation Pareto optimale décrite ci-dessus peut être obtenue par des échanges de biens contingents sur des marchés de type "Arrow-Debreu". On définit p_E comme étant le prix contingent à l'état collectif E . Chaque individu maximise son espérance d'utilité sous contrainte budgétaire :

$$\begin{aligned} \max. \quad & \sum_E \pi_E u_i(c_E^i) \\ \text{sc.} \quad & \sum_E p_E c_E^i = \sum_E p_E w_E^i \end{aligned}$$

Chaque marché contingent est à l'équilibre si $\sum_i c_E^i = \sum_i n_i w_i = W$, pour tout E .

On montre qu'à l'équilibre, la consommation est constante pour tous les individus, quel que soit E , et qu'elle est égale à la dotation moyenne de l'individu sur l'ensemble des états du monde (Henriet et Rochet, 1991, pp. 100-102). Dans le cas particulier d'identique exposition au risque, cette consommation est la même pour tous les individus.

A l'équilibre, les prix relatifs d'équilibre entre deux états du monde sont définis de façon unique et ils sont égaux au rapport des probabilités des états collectifs.

Pour deux états E et E' :

$$\frac{p_E}{p_{E'}} = \frac{\pi_E}{\pi_{E'}}$$

Pour atteindre l'équilibre "Arrow - Debreu", il faut donc ouvrir $n_{1/\Omega}! n_{2/\Omega}! \dots n_{s/\Omega}!$ $N!$ marchés contingents (un par état collectif). Les coûts de la mise en œuvre de ce mécanisme de partage de risque apparaissent donc considérables (Gérard-Varet, 1997).

Deux autres possibilités s'offrent aux individus pour atteindre l'optimum par un mécanisme de marché : la première est la mise en œuvre de contrats de couverture réciproque, la seconde est l'assurance. Le premier modèle présente avant tout un intérêt théorique car il est très général et assez méconnu, le second correspondant au mécanisme le plus couramment étudié de couverture de risques individuels. Nous verrons cependant que sa mise en œuvre nécessite que des conditions plus strictes soient vérifiées.

2.3 La couverture réciproque

Cass, Chichilnisky et Wu (1991, 1996) montrent qu'avec $S-I$ contrats de prix actuariel (dits contrats de couverture réciproque), l'ensemble des individus d'une économie avec risque individuel, peut atteindre l'allocation optimale sans qu'il soit nécessaire d'ouvrir l'ensemble de ces marchés contingents. Le modèle postule une identique exposition au risque, mais nous verrons que cette hypothèse peut être relâchée.

2.3.1 Définition des contrats et programme du consommateur

Chacun des $S-I$ contrats concerne la réalisation, pour un individu, de l'état s . La prime unitaire est de π_s . L'individu choisit pour chaque état s une quantité x_s^i .

Les termes du contrat sont tels que l'individu :

- "obtient" $(1 - \pi_s) x_s^i$ si l'état du monde s se réalise ;
- "paie" $\pi_s x_s^i$ dans tous les autres états du monde.

π_s est un prix qui peut être assimilé à une prime, car, pour obtenir 1 dans l'état s , il faut payer π_s dans tous les états du monde. Le terme actuariel signifie que ce prix est égal à la probabilité de réalisation de l'état qui donne lieu au versement de 1.

x_s^i peut être positif ou négatif, autrement dit, l'individu peut être en situation de payer une somme ou d'en recevoir une si l'état du monde s se réalise. C'est en cela que l'on peut qualifier ce mécanisme de "couverture réciproque".

Il n'existe pas de contrat pour se couvrir contre l'état Σ (choisi au hasard).

Dans les conditions posées, le programme individuel est donc :

$$\begin{aligned} \text{max.} \quad & \sum_s \pi_s u_i(c_s^i) \\ \text{sc.} \quad & c_s^i = w_s + x_s^i - \sum_{s' \neq \Sigma} \pi_{s'} x_{s'}^i \quad \text{pour tout } s \neq \Sigma \\ & c_\Sigma^i = w_\Sigma - \sum_{s \neq \Sigma} \pi_s x_s^i \end{aligned}$$

ou, après substitution :

$$\text{max.} \quad \sum_{s \neq \Sigma} \pi_s u_i \left(w_s + x_s^i - \sum_{s' \neq \Sigma} \pi_{s'} x_{s'}^i \right) + \pi_\Sigma u_i \left(w_\Sigma - \sum_{s \neq \Sigma} \pi_s x_s^i \right)$$

2.3.2 Solution

On montre alors que la consommation de chacun est constante dans tous les états du monde, et égale à l'espérance de dotation sur l'ensemble des états du monde. Cette valeur c est la même pour tous, par hypothèse d'identique exposition au risque. Tous les individus demandent ou offrent la même quantité x dans un état s .

On retrouve donc le résultat concernant le partage optimal du risque individuel, lorsque l'allocation totale dans l'économie est fixée : l'allocation de la richesse est telle que chacun a la même consommation dans tous les états du monde. Autrement dit le risque individuel est totalement assuré.

Le modèle de Cass, Chichilnisky et Wu postule l'existence de H catégories d'individus. Tous les individus d'une même catégorie sont identiquement exposés au risque (ils ont tous la même richesse dans les s états du monde) et ont la même fonction d'utilité. Dans ce cas, le partage du risque individuel impose l'existence de $H(S-1)$ contrats d'assurance.

On peut cependant remarquer que :

1. l'allocation ne dépend pas du niveau d'aversion au risque des individus. L'hypothèse faite par les auteurs que les individus d'une même catégorie ont la même fonction d'utilité n'est pas nécessaire à l'obtention du résultat. Il suffit que les fonctions d'utilité soient convexes ;
2. de plus, l'hypothèse que les individus d'une même catégorie sont identiquement exposés au risque peut être partiellement relâchée. Si, entre deux états du monde quelconques s et s' , l'allocation de tous les individus varie de la même quantité alors :
 - l'allocation globale dans l'économie est constante ;
 - les individus consomment, dans tous les états du monde, leur dotation espérée ;
 - la dotation globale est égale à la consommation de tous les individus ;
 - les contrats de couverture réciproque sont équilibrés.

On pourrait décrire cette configuration de la manière suivante : l'état Σ correspond pour chaque individu à un état de santé n'entraînant pas de dépense de soins. Les dotations initiales des individus w^i_{Σ} reflètent leurs niveaux de richesse. Les $S-1$ autres états correspondent à $S-1$ maladies aléatoires qui entraînent des soins d'un montant a_1, \dots, a_{S-1} . La proportion dans la population de personnes qui souffrent d'une maladie s est connue. La richesse globale dans l'économie est donc constante. Dans cette économie, les individus peuvent en théorie se couvrir complètement contre le risque individuel de tomber malade en utilisant la couverture réciproque.

Au terme de cette analyse, il semble donc que le résultat de Cass, Chichilnisky et Wu peut être généralisé à l'ensemble des situations qui remplissent les deux conditions énoncées et s'applique notamment à la couverture des risques liés à la santé.

Pour conclure sur ce mécanisme de partage réciproque du risque individuel, on peut souligner qu'un type d'institution proche de ce modèle existe : la "société de secours mutuel" ou la mutuelle de fermes (*farm mutual*), qui apparaît aux USA autour des années 1875 (Heflebower, 1980). Face à un risque d'incendie et devant la difficulté de trouver une assurance commerciale, les fermiers s'organisent localement et sont appelés à contribuer à la réparation d'un dommage lorsque celui-ci survient. Ce modèle, auquel nous reviendrons, s'applique aussi théoriquement dans le cas où la richesse globale varie

et où le risque n'est pas seulement individuel. Il implique une forme de solidarité ex post (une fois le dommage survenu), ou du moins un engagement ex ante qui se traduit par des versements ex post. Cette solidarité n'est pas nécessairement transposable à grande échelle. De plus, une telle configuration implique des paiements décentralisés (puisque chacun doit verser une somme aux personnes qui subissent les dommages) dont la gestion est très lourde.

2.4 Le mécanisme d'assurance

Les échanges sur les marchés contingents, et les contrats de couverture réciproque peuvent être remplacés par des contrats offerts par une compagnie d'assurance qui proposerait une couverture actuarielle. Elle prélève les mêmes primes que celles du modèle avec contrat réciproque et assure, de façon centralisée, le versement des "indemnités" aux personnes qui se trouvent ex-post dans l'état *s*. Dans cette configuration, le risque est transféré à une entité externe au groupe des personnes se couvrent.

Ce mode d'organisation est l'une des méthodes de mise en œuvre de ce que Marshall (1974) appelle le principe de réserve :

"Individuals who experience risk must lay up reserves. While they may do this individually, under some circumstances it is cheaper for them to combine reserves. In practice they do so by buying insurance contracts. (...) thus the insurance industry exists to exploit economies in pooling reserves" ³ (pp.477).

L'étude du comportement des agents économiques en environnement incertain dans une situation où la richesse globale est constante permet de conclure qu'une personne risquophobe confrontée à une série d'événements aléatoires cherche à stabiliser sa richesse entre les différents états du monde.

La théorie de l'assurance oppose plus nettement les cas où la personne est dans un état favorable et où sa richesse est élevée, et les cas où elle subit un dommage et où sa richesse diminue d'un montant donné.

³ Les personnes confrontées à des risques doivent constituer des réserves. Elles peuvent le faire à titre individuel, mais il peut s'avérer moins onéreux, dans certaines circonstances, de les regrouper. En pratique elles réalisent cette opération en souscrivant des contrats d'assurance...cette industrie exploite donc les économies (ndlt d'échelle) dans la mutualisation des réserves.

Afin d'achever la présentation des mécanismes de partage du risque individuel, on décrit ici la version la plus simple du modèle où il existe une compagnie d'assurance et où l'individu est soumis à un risque unique.

2.4.1 Hypothèses et définition du contrat d'assurance

L'assurance est un mécanisme qui permet aux individus de transférer de la richesse conditionnelle entre les états du monde " malade " (noté ici état 2) et " en bonne santé " (état 1), de façon à maximiser leur bien-être.

Les individus ont la même dotation w dans le cas où ils ne sont pas malades (l'état Σ dans le paragraphe précédent). Ils sont identiquement exposés au risque et confrontés à un dommage unique, l'événement aléatoire "tomber malade" associé à une probabilité π . La maladie est assimilée à une perte monétaire a , qui représente par exemple le coût du traitement que chacun doit suivre pour être guéri.

Les individus maximisent l'espérance d'utilité suivante :

$$(1 - \pi)u(c_1) + \pi u(c_2)$$

Sans assurance,

$$c_1 = w$$

$$c_2 = w - a$$

La dotation espérée de l'individu est :

$$E(c) = (1 - \pi)w + \pi(w - a) = w - \pi a$$

On définit le contrat d'assurance selon deux dimensions : le prix et la quantité. Le "prix" unitaire p correspond au montant de prime qu'il faut déboursier pour obtenir un franc de remboursement en cas de dommage et la quantité x reflète le taux de couverture obtenu par l'assuré. L'individu paie px dans les deux états du monde et reçoit x dans le cas où il tombe malade. Dans cette configuration, l'individu choisit la quantité d'assurance x qui maximise son espérance d'utilité:

$$(1 - \pi)u(w - px) + \pi u(w - a - px + x)$$

2.4.2 Résolution et interprétation

L'annulation de la dérivée partielle de l'équation précédente par rapport à x permet d'obtenir son maximum (la fonction u est concave) :

$$-(1-\pi)pu'(w-px) + \pi(1-p)u'(w-a-px+x) = 0 \quad (1)$$

L'entreprise qui fournit le service :

- d'un côté, encaisse (globalement) N fois le montant de la prime px , où N est le nombre de personnes dans l'économie ;
- de l'autre, dans la mesure où la loi des grands nombres s'applique, les $N\pi$ personnes qui subissent le dommage reçoivent x .

Si l'entreprise évolue dans un environnement concurrentiel, elle réalise un profit nul lorsque :

$$Npx - N\pi x = 0$$

Soit,

$$p = \pi$$

La prime unitaire est égale à la probabilité de réalisation du dommage.

La condition (1) est alors remplie si :

$$w - px = w - a - px + x$$

Autrement dit, si le dommage est intégralement couvert :

$$x = a$$

Le résultat semble donc identique à celui présenté précédemment (on pourrait le généraliser à une série de dommages d'un montant $a_1 \dots a_{S-1}$). Ici, le prix unitaire de la couverture π est dit actuariel parce qu'il est égal à la probabilité de réalisation du dommage, ce qui est une conséquence du fait que le marché est concurrentiel et que les profits sont nuls. La quantité d'assurance demandée et obtenue par l'assuré est a . L'assurance est complète.

Dans cette transaction, le rôle de l'assureur consiste essentiellement à mutualiser les risques, autrement dit, à les regrouper. Plus précisément, il doit constituer un groupe

suffisamment grand de personnes exposées à des risques *identiques* pour faire jouer la loi des grands nombres.

Il nous semble utile de clarifier, à ce stade, la définition du concept de mutualisation qui fait souvent, dans le vocabulaire des économistes et des praticiens, référence à des notions différentes. Le terme "mutualisation des risques", qui qualifie toute opération de mise en commun de risques, ne fait pas seulement référence à des opérations d'assurance actuarielles : si toute assurance pratique une mutualisation, la mutualisation peut emprunter d'autres voies que l'assurance actuarielle. Une entité publique qui prélève des cotisations et rembourse les frais médicaux des personnes couvertes réalise aussi une mutualisation des risques, sans qu'il y ait tarification au risque. Comme le souligne Blanchet (1996), on tend à parler d'assurance "*sitôt que des mécanismes de mutualisation sont utilisés pour couvrir un risque*", ce qui conduit à une définition trop large du champ de l'assurance. A l'inverse, le terme de "mutualisation" est parfois employé pour caractériser une situation dans laquelle il y a *péréquation* entre niveaux de risques différents, autrement dit pour qualifier une situation dans laquelle les paiements ne sont pas actuariels. Ainsi une mutuelle, qui choisit de ne pas appliquer une tarification au risque réalise, au même titre qu'une compagnie d'assurance, une mutualisation des risques. Elle s'en différencie en ce qu'elle combine cette mutualisation avec une *péréquation explicite* entre niveaux de risques.

Quoi qu'il en soit, dans le cas où l'assurance est actuarielle, si l'individu a une probabilité π de subir la maladie, et que celle-ci occasionne une perte de revenu de a , il paie πa à un assureur pour qu'il lui verse le montant a s'il subit un dommage pendant la durée de la période contractuelle. Ainsi, il égalise son revenu dans les deux cas de figure : $w - \pi a$ s'il ne tombe pas malade, $w - a + a - \pi a$ dans le cas contraire. D'une situation où l'individu risquait une perte a , il passe à une situation où sa richesse est constante dans toutes les circonstances. Cette situation est tout à la fois un équilibre et un optimum de Pareto. L'intervention d'un nouvel intermédiaire sur le marché, l'assureur, permet donc théoriquement d'atteindre les résultats optimaux de partage du risque individuel. Ce mécanisme présente en outre l'avantage de diminuer le nombre de transactions à effectuer, y compris par rapport à une situation de couverture réciproque⁴.

⁴ Où les transactions sont déjà moins nombreuses que dans le cas des échanges sur les marchés contingents.

On peut trouver, dans cette simplicité, la raison fondamentale de la prédominance dans la réalité de ce mécanisme : l'entreprise d'assurance est le prestataire privilégié du service de couverture du risque individuel. Henriet et Rochet (1997b) proposent, dans un autre contexte, une justification de l'intermédiation par l'assurance en terme d'engagement : *"une fois le risque réalisé, un individu bien portant n'a aucune incitation à financer les coûts de ceux qui sont tombés malades. Une simple promesse ex ante n'est pas crédible... l'intermédiation d'une compagnie d'assurance rend cette promesse crédible en lui donnant un caractère d'engagement"*. De fait, la couverture réciproque que nous avons évoquée précédemment suppose une contrainte morale forte et force est de constater que les contrats procédant de ce mécanisme se sont développés dans les communautés restreintes, fortement soudées et dans lesquelles le contrôle réciproque et la menace de sanction d'un manquement à ses obligations sont crédibles (communauté rurale de proximité). Enfin, le mécanisme d'assurance permet, lorsque les risques sont lourds, une nécessaire diversification du risque sur un groupe assez large, incompatible avec cette notion de contrôle.

Au terme de cette présentation des mécanismes de partage du risque individuel, il nous semble naturel de mettre l'accent sur l'assurance. Nous nous tournons maintenant vers l'étude des moyens qui pourraient permettre aux individus de se prémunir contre le risque que nous avons qualifié de social.

3. PARTAGE DU RISQUE SOCIAL

Nous présentons ici les résultats de la théorie économique concernant le partage du risque que nous avons qualifié de social. Nous considérons que les individus reçoivent, dans chaque état social, la dotation moyenne associée à la réalisation de l'ensemble des états collectifs qui le composent. Autrement dit, il n'y a pas de risque individuel, mais seulement un risque social. Signalons que cette hypothèse, qui permet de simplifier la présentation, n'influence pas le résultat final : lorsque risque social et risque individuel coexistent, le risque individuel est effectivement assuré complètement pour chaque état social (comme dans les modèles que nous venons de présenter).

Chaque état Ω est associé à un niveau de richesse agrégée différent. On distingue chaque état Ω par un indice t ($t = 1$ à T) qui augmente avec le niveau de richesse agrégée W_{Ω} . Les probabilités associées aux états du monde sociaux sont $\pi_{\Omega_1}, \dots, \pi_{\Omega_T}$.

Après avoir décrit les propriétés de l'optimum (3.1) et de l'équilibre (3.2), nous nous interrogerons sur l'existence d'institutions qui, sur le marché du risque collectif, pourraient constituer le pendant des assurances sur le marché du risque individuel (3.3).

3.1 L'optimum

3.1.1 Position du problème

Les conditions auxquelles l'allocation est Pareto optimale s'expriment de la façon suivante :

1. $\sum_i c_{\Omega t}^i = \sum_i w_{\Omega t}^i = W_{\Omega t} \quad \forall t$ (condition de faisabilité)
2. Il n'existe pas d'autre allocation faisable qui augmente l'espérance d'utilité d'un individu sans diminuer celle d'un autre.

Pour déterminer les allocations optimales, on considère une fonction d'utilité collective :

$$U = \sum_i \lambda_i \left[\sum_t \pi_{\Omega t} u_i(c_{\Omega t}^i) \right]$$

Le choix des λ_i (> 0) est arbitraire. A chaque choix de λ_i correspond une allocation optimale (Gollier, 1992, pp. 6 par exemple).

On maximise U sous les T contraintes reflétant les conditions de faisabilité, par la méthode de Lagrange.

3.1.2 Résolution et analyse

Les conditions de premier ordre peuvent être exprimées de la façon suivante :

$$\lambda_i u'_i(c_{\Omega t}^i) = \lambda_j u'_j(c_{\Omega t}^j) \quad \forall i, j \text{ et pour tout } t$$

C'est la *condition de Borch* (Borch, 1962, 1990, par exemple), dont on peut déduire une série de résultats.

a) Egalité des taux marginaux de substitution

Tout d'abord, pour deux états sociaux quelconques, le taux marginal de substitution entre ces deux états est le même pour tous les individus. Ceci signifie qu'à l'optimum il n'y a plus d'échanges mutuellement avantageux possibles entre deux individus.

b) Monotonicité des consommations individuelles quand la richesse agrégée augmente

Considérons deux états t et u tels que $W_{\Omega_t} < W_{\Omega_u}$

On montre que nécessairement $c^i_{\Omega_t} < c^i_{\Omega_u} \quad \forall i$. Autrement dit, lorsqu'on considère les états Ω_1 à Ω_T (classés par niveau croissant de dotation globale), on voit que la consommation de tous les individus augmente d'un état à l'autre et ce, indépendamment de leur dotation individuelle.

Par exemple :

- si la dotation d'un individu est minimale en Ω_1 et maximale en Ω_T , sa consommation sera la plus faible en Ω_1 et la plus élevée en Ω_T ;
- si un individu a une dotation constante dans tous les états sociaux, il pourrait apparaître entièrement assuré. Il n'est pourtant pas optimal qu'il conserve ce profil de consommation lissé, mais préférable que le risque social soit partagé entre tous les individus.

Ce résultat correspond à l'intuition selon laquelle si tous les individus ont de l'aversion pour le risque (ce qui est le cas), il faut diversifier ce risque autant que faire se peut.

c) Mise en commun des ressources agrégées

Le fait que la consommation ne dépende qu'indirectement de la dotation individuelle peut être démontré plus spécifiquement.

En remarquant que $c^i_{\Omega_t} = f(w^1_{\Omega_t}, \dots, w^i_{\Omega_t}, w^k_{\Omega_t}, \dots, w^N_{\Omega_t})$ et en différenciant la condition de Borch, on peut montrer (1962, pp. 427-428) que :

$$\frac{\partial c^i_{\Omega_t}}{\partial w^k_{\Omega_t}} = \frac{\partial c^i_{\Omega_t}}{\partial W_{\Omega_t}}$$

La consommation d'un individu dans l'état t dépend de la dotation globale de l'état W_{Ω_t} , et ne dépend de la dotation individuelle que dans la mesure où celle-ci a un impact sur la dotation globale.

Dans le contexte de la réassurance, Borch interprète ce résultat de la façon suivante :

*"Any Pareto optimal set of treaties is equivalent to a pool arrangement, i.e., all companies hand their portfolios over to a pool, and agree on some rule as to how payments of claims against the pool shall be divided among the companies"*⁵ (1962, pp. 428).

Tout se passe en fait comme si les individus mettaient toutes leurs allocations en commun et s'entendaient sur une règle de partage de la richesse agrégée.

d) Caractérisation de la règle de partage

On peut tout d'abord rappeler qu'une infinité de règles de partage Pareto optimales sont possibles puisque les N constantes λ_i sont choisies arbitrairement. Wilson (1968) va plus loin dans la caractérisation de la règle de partage du risque social.

Si on note l'indice d'Arrow-Pratt de tolérance absolue au risque pour l'individu i et pour le niveau de consommation⁶:

$$t_i(c_{\Omega_t}^i) = - \frac{u'_i(c_{\Omega_t}^i)}{u''_i(c_{\Omega_t}^i)}$$

On démontre à partir de la condition de Borch que :

$$\frac{\partial c_{\Omega_t}^i}{\partial W_{\Omega_t}} = \frac{t_i(c_{\Omega_t}^i)}{\sum_j t_j(c_{\Omega_t}^j)} \quad (2)$$

On appelle "tolérance au risque absolue du groupe", la somme des tolérances au risque des individus (dénominateur). La relation (2) signifie que l'augmentation marginale de la consommation d'un individu, lorsque la richesse agrégée croît, est égale à sa part dans la tolérance au risque absolue du groupe. Gollier (1992, pp. 8) démontre que cette règle

⁵ *Tout ensemble de contrats Pareto optimal est équivalent à une mutualisation : toutes les entreprises (d'assurance) mettent en commun leurs portefeuilles et s'accordent sur une règle de répartition entre les compagnies.*

⁶ *"Half the tolerable variance per unit of compensating risk premium for infinitesimal risks"* quand la consommation est $c_{\Omega_t}^i$ (Wilson, 1968, pp. 121)

de partage est celle qui minimise la somme des primes de risque individuelles (au sens de Pratt⁷) dues au partage d'une augmentation infinitésimale de la richesse agrégée.

On peut noter, à la lumière de la relation (2), que :

- plus une personne a une tolérance au risque élevée, plus sa part optimale dans la prise en charge du risque social est grande ;
- à la limite, si une personne (ou un groupe de personnes) est neutre au risque, on montre qu'elle assume le risque social dans sa totalité (sous réserve que son espérance d'utilité soit supérieure ou égale à celle qui correspond à son allocation initiale et que ses consommations soient positives) ;
- la règle de partage peut être linéaire (et ne pas dépendre du niveau de richesse agrégée) si certaines conditions sont remplies (constance de l'aversion absolue au risque entre autres, voir par exemple Wilson, 1968 ou Eliashberg et Winkler, 1981).

Au terme de cette analyse, on peut résumer simplement les caractéristiques d'une allocation optimale du risque agrégé : il faut que ce risque soit diversifié sur tous les membres du groupe et que la répartition tienne compte de la tolérance au risque de chacun.

3.2 L'équilibre

Les résultats précédents caractérisent l'optimum et donc tout équilibre concurrentiel. Nous étudions plus précisément ici les mécanismes de marchés qui permettent le transfert du risque social.

Pour chaque état social Ω_r , un marché contingent est ouvert, sur lequel les transactions se font à un prix p_r .

⁷ $P_i = -\frac{\sigma^2 u_i''}{2u_i'}$

3.2.1 Position du problème

Chaque individu maximise son espérance d'utilité sous contrainte budgétaire :

$$\begin{aligned} \max. \quad & \sum_t \pi_{\Omega_t} u_i(c_{\Omega_t}^i) \\ \text{sc.} \quad & \sum_t p_t c_{\Omega_t}^i = \sum_t p_t w_{\Omega_t}^i, \text{ contrainte à laquelle on associe } \lambda_i \end{aligned}$$

Chaque marché contingent est à l'équilibre si $\sum_i c_{\Omega_t}^i = \sum_i w_{\Omega_t}^i = W_{\Omega_t}$ pour tout Ω_t .

Les conditions de premier ordre du programme de chaque consommateur sont que :

$$\frac{\pi_{\Omega_t}}{\pi_{\Omega_{t'}}} \frac{u'_i(c_{\Omega_t}^i)}{u'_i(c_{\Omega_{t'}}^i)} = \frac{p_t}{p_{t'}} \quad \text{pour tout } i \text{ et quels que soient } t \text{ et } t'.$$

Ce résultat peut aussi être exprimé de la façon suivante :

$$\frac{u'_i(c_{\Omega_t}^i)}{u'_i(c_{\Omega_{t'}}^i)} = \frac{u'_j(c_{\Omega_t}^j)}{u'_j(c_{\Omega_{t'}}^j)} \quad \text{pour deux individus et deux états quelconques.}$$

3.2.2 Interprétation des résultats

Les fonctions d'utilité marginale étant décroissantes, si la consommation de l'individu i est plus élevée dans l'état t que dans l'état t' , ceci est aussi vrai pour tout autre individu. La dotation agrégée de l'état t est donc plus élevée que celle de l'état t' .

Deux remarques :

1. on retrouve donc le résultat de monotonie des consommations individuelles quand l'allocation globale augmente. Plus généralement, le premier théorème du bien-être s'applique : l'allocation compétitive est une allocation Pareto efficiente. La condition de premier ordre du programme d'équilibre est semblable à celle du programme d'allocation optimale présenté précédemment. Pour faire le lien entre les deux, il suffit que les λ_i (du programme d'allocation paretien) soient tels que les contraintes budgétaires du programme d'équilibre individuel soient saturées ;

2. en terme de prix, ceci signifie que :

$$\frac{P_t}{\pi_{\Omega_t}} < \frac{P_{t'}}{\pi_{\Omega_{t'}}$$

si la richesse agrégée est plus élevée dans l'état t que dans l'état t' .

Comme le conclut Marshall (1974) : *(probability) "adjusted prices rise as aggregate (and hence individual) consumption falls. Increasing adjusted prices are the incentives which assure the parallel decline of consumption for all individuals"*⁸ (pp. 479).

3.3 Le principe de mutualité

Après avoir traité la question du risque individuel et vu qu'il pouvait (et devait) être complètement diversifié, il est apparu qu'il n'en était pas de même pour le risque collectif. Une allocation efficiente du risque collectif est telle qu'une variation de la richesse globale affecte tous les individus "dans le même sens" mais pas forcément dans les mêmes proportions. C'est ce que Marshall appelle le principe de mutualité, qui permet le partage du risque, même si la loi des grands nombres ne s'applique pas : *"the mutuality principle is that individuals take on an obligation to share losses suffered by others, the quid pro quo being the reciprocity of the arrangement"*⁹ (pp 477).

Arrow (1996) souligne que les contrats de transfert de risque observés dans la réalité couvrent le plus souvent des risques individuels et qu'en tout état de cause, tous les agents économiques n'assument pas la fonction de porteur de risque. Cette fonction échoit plutôt à des entreprises d'assurance, dont les contrats ne sont pas conditionnés par la réalisation de l'état social.

La portée de ces remarques est néanmoins atténuée car :

- les compagnies d'assurances sont des intermédiaires financiers dont les actionnaires participent à la diversification du risque sur un plus grand nombre de personnes ;

⁸ Les prix ajustés aux probabilités augmentent lorsque la consommation agrégée (et donc individuelle) diminue. Cette augmentation des prix ajustés est le mécanisme incitatif qui garantit la décroissance de la consommation de tous les individus.

⁹ Le principe de mutualité veut que les individus s'engagent à partager les pertes subies par d'autres, en contrepartie d'un engagement réciproque.

- les traités de réassurance aboutissent aussi à une plus grande diversification du risque.

Marshall remarque cependant qu'un autre mécanisme de partage du risque social existe, dans lequel les individus sont partie prenante de façon directe et qui permet une allocation du risque collectif : le contrat mutuel participatif. "*Mutual and participating stock insurance companies issue contracts which include, besides the obligation to indemnify loss, a dividend to the consumer which depends on the overall performance of the company*"¹⁰ (pp. 483).

Même si les individus couverts par ce type de contrat obtiennent tous le même "dividende" et non un dividende proportionnel à leur tolérance au risque, ce type de mécanisme semble pouvoir assurer une répartition plus efficace du risque qu'un simple contrat d'assurance qui permet seulement de diversifier le risque individuel. L'hypothèse est explorée et validée par Dionne et Doherty (1993). Ils étudient un modèle dans lequel coexistent risque individuel et risque social. Les assurances peuvent proposer des contrats déterminés en fonction du seul risque individuel, ou qui comprennent un dividende conditionné à la réalisation d'un état social particulier. L'allocation des ressources, dans ce second cas, est Pareto supérieure.

En partant de l'idée que chacun peut être exposé à une série de risques connus, nous avons, dans un premier temps, souligné que la conjonction de ces risques peut se traduire par deux types de situations :

- dans certains cas, la réalisation de ces risques conduit à des situations où le niveau de richesse agrégée est constant. On dit alors que le risque n'est qu'individuel ;
- dans d'autres, elle se traduit par des situations où la richesse globale de l'économie varie, le risque étant alors social.

Les conditions optimales de partage de ces risques et les mécanismes de marché qui permettent de les atteindre diffèrent fondamentalement, selon que l'on étudie le partage

¹⁰ *Les mutuelles et certaines compagnies d'assurance émettent des contrats qui incluent, à côté de l'obligation de rembourser une perte, un dividende qui dépend de la performance globale de l'entreprise.*

des risques quand la richesse est constante ou quand elle varie. Risque individuel et risque social coexistent, mais on peut montrer que ceci n'a aucune incidence sur les conditions optimales de partage des risques. Pour un niveau de richesse agrégée, le risque individuel est entièrement diversifié et les individus ont une allocation constante dans tous les états collectifs. Si la richesse agrégée augmente (resp. diminue), la richesse de chacun augmente (resp. diminue) dans des proportions qui dépendent de son aversion au risque.

Sans négliger l'intérêt théorique de la question du partage du risque social, on peut présenter deux types de raisons qui justifient notre choix de ne pas approfondir cet aspect dans la suite de ce travail.

1. La coexistence de risques individuels et collectifs permet de caractériser l'économie dans son ensemble. Or, notre objectif est d'étudier un ensemble de risques particuliers liés à la santé. Il semble évident qu'au niveau de la population, la réalisation de certains risques santé, comme par exemple des épidémies, peuvent conduire à une variation de la richesse globale, par des mécanismes économiques directs ou indirects faisant intervenir un grand nombre de marchés et d'agents économiques. Pour autant, notre approche étant résolument centrée sur l'étude d'un risque particulier et sur sa couverture, il serait trop réducteur de l'étudier dans sa dimension sociale (au sens où nous l'avons entendu dans cette section) sans le replacer dans l'ensemble des risques qui contribuent à faire varier la richesse globale. La question du risque "social" ne nous intéresse donc qu'à la marge, dans la mesure où le risque social conditionne le fonctionnement du partage du risque individuel.
2. En outre, les mécanismes de partage du risque global ne concernent probablement pas aussi directement les individus que les entités qui les couvrent pour le risque individuel. Exception faite des contrats participatifs que nous avons signalés, ce sont des agents économiques intermédiaires qui vont réaliser le partage des risques agrégés par des mécanismes de réassurance ou de diversification de portefeuille. Il est d'ailleurs symptomatique que les principaux résultats en la matière aient été initiés dans un traité de réassurance (Borch, 1962).

4. LE SERVICE D'ASSURANCE

Parmi les différents mécanismes de couverture que nous venons de présenter, nous focaliserons désormais notre analyse sur l'assurance qui permet aux individus de transférer le risque sur une entité extérieure sans avoir à entrer dans de multiples transactions inter-individuelles. C'est le mécanisme le plus économe en coûts de transactions et, de fait, celui qui correspond à la méthode la plus répandue de couverture des dommages. L'assurance est un service rendu par une entreprise particulière et le fonctionnement de ce marché a été l'objet de nombreux développements théoriques. Nous présentons ici les principaux résultats qui permettent de cadrer l'analyse ou qui trouveront un écho particulier dans le cas de la maladie. Ils portent sur l'offre et la demande d'assurance (4.1) et le fait que la détermination du risque n'est pas totalement exogène (4.2).

4.1 L'offre et la demande d'assurance

On peut à ce stade se demander à quelles conditions l'assurance est en mesure d'offrir le service concerné (4.1.2) et caractériser plus précisément la demande du service (4.1.2).

4.1.1 Les conditions d'assurabilité d'un risque

Pour que l'entreprise d'assurance puisse rendre ce service, un certain nombre de conditions qui portent sur les risques considérés doivent être réunies. Dans la réalité, le risque individuel peut être pris en charge par une compagnie d'assurance si (Eeckhoudt, 1992) :

- les risques sont purs. Ils ne se manifestent que par des pertes (ce ne sont pas des risques spéculatifs) ;
- la perte, si elle advient, doit être mesurable ;
- un nombre suffisant d'individus sont exposés au même risque ;
- les risques doivent être largement indépendants afin que la loi des grands nombres s'applique. Cette condition peut s'interpréter comme le fait de dire que le risque n'est qu'individuel et non social au sens où nous l'avons entendu précédemment ;

- l'individu n'a pas d'influence sur le risque. Les contrats d'assurance peuvent cependant être aménagés pour prendre en compte le problème du risque moral ;
- l'assureur doit être en mesure ex ante d'évaluer le risque (nous étudierons longuement les difficultés que cette hypothèse soulève dans le domaine de la santé) ;

Par rapport aux modalités théoriques de partage de risques sur des marchés contingents ou de couverture réciproque, les conditions dans lesquelles l'assurance peut fonctionner sont plus strictes. Schmit (1986) passe en revue ces conditions théoriques d'assurabilité et leur substitue une condition plus large mais plus réaliste : l'assureur doit simplement pouvoir prédire avec une marge d'erreur minimale les pertes associées à son portefeuille de risque. Si, par exemple, les risques sont corrélés, un assureur capable de mesurer cette corrélation et de se réassurer peut les assumer. Pour qu'un risque assurable au sens où Schmit l'entend soit effectivement assuré, une condition de faisabilité doit en outre être remplie : la prime associée au service d'assurance ne doit pas être trop élevée, sans quoi il n'y aura pas de demande pour ce marché.

L'article de Schmit invite donc à reconsidérer les conditions théoriques d'assurabilité d'un risque et à les confronter au fonctionnement effectif des marchés. Nous adopterons par la suite une approche semblable pour présenter le rôle des assureurs dans la couverture du risque maladie.

4.1.2 La demande d'assurance

En admettant que l'assureur soit en mesure de rendre le service considéré, on peut s'interroger sur les caractéristiques de la demande qui lui est adressée.

Dans le cas où le dommage est unique, si le prix unitaire p est actuariel ($p = \pi$), l'individu se couvre complètement ($x = a$), quels que soient son niveau d'aversion au risque, la probabilité de réalisation (π) ou la taille du dommage (a) et son revenu (w) (si tant est qu'il lui permet de s'acquitter de la prime d'assurance). En revanche, si le prix n'est pas actuariel, ce qui peut s'expliquer (y compris en concurrence parfaite) par l'existence de coûts de gestion, la demande d'assurance varie en fonction de l'ensemble de ces paramètres.

Considérons que $x=F(p, w, \pi, a)$.

On montre que (Henriet et Rochet, 1991) la quantité d'assurance demandée :

- augmente lorsque π et a augmentent, autrement dit lorsque, pour un prix et un niveau de richesse donnés, la taille ou la fréquence du dommage augmentent. Ce dernier résultat doit cependant être nuancé si le coût réel de l'assurance augmente avec la probabilité de dommage¹¹. Dans ce cas particulier, la demande d'assurance sera moins élevée pour les dommages fréquents (Ehrlich et Becker, 1972) ;
- diminue quand la richesse augmente si et seulement si l'aversion absolue pour le risque est décroissante, ce qui est une hypothèse communément admise (Arrow, 1971). On peut expliquer ce résultat apparemment paradoxal par l'idée que la valeur relative du bien assuré diminue lorsque la richesse augmente ;
- diminue lorsque le prix unitaire de l'assurance augmente, si l'aversion relative pour le risque de l'individu est inférieure à 1.

Dans le cas, plus général, où le dommage est représenté par une variable aléatoire (risques multiples) dont on connaît la distribution, le contrat se décline en deux dimensions : la prime et l'indemnité reçue en cas de dommage. On montre alors que si le taux de chargement est positif, tout contrat optimal comporte une indemnisation complète au delà d'une franchise. L'existence de ces coûts rend l'indemnisation des petits sinistres inefficace. Schlesinger (1981) montre, sous certaines conditions, que le niveau de la franchise diminue (et donc la demande d'assurance augmente) si l'individu a plus d'aversion au risque, si la probabilité de dommage augmente et si sa richesse initiale est plus faible. Ces résultats sont similaires à ceux que l'on trouve dans le cas du risque unique.

¹¹ Lorsque le prix est égal à la probabilité augmentée d'un taux de chargement et que ce taux de chargement augmente avec la probabilité de dommage, ce qui correspond à l'hypothèse que plus les dommages sont fréquents, plus les coûts de gestion par dommage seront élevés.

4.2 L'endogénéisation des risques

Ce panorama introductif reste incomplet car nous n'avons évoqué que les possibilités qu'a l'individu de partager un risque parfaitement exogène avec d'autres ou de le transférer sur une entité extérieure : la compagnie d'assurance. Or, on peut faire l'hypothèse qu'il est possible de diminuer le risque auquel on est exposé moyennant un effort. Deux approches de cette question sont privilégiées par l'analyse économique :

- soit l'individu arbitre entre les différentes possibilités qui lui sont offertes pour se protéger contre le risque ;
- soit on considère que le niveau d'effort est choisi par l'individu après qu'il s'est assuré et que l'assureur ne peut observer cet effort : il y a risque moral.

4.2.1 Auto-protection, auto-assurance et assurance

Si le risque n'est pas exogène, deux nouveaux moyens s'offrent à l'individu pour se prémunir contre la réalisation un dommage. Le premier consiste pour l'individu à diminuer la probabilité d'occurrence d'un dommage et le second à limiter le montant de la perte si celui-ci survient. Ces deux mécanismes et leur impact sur la demande d'assurance ont été étudiés par Ehrlich et Becker (1972) qui les appellent respectivement auto-protection et auto-assurance. Ils formalisent l'idée que l'individu arbitre entre les possibilités qui lui sont offertes de se protéger contre le risque ou de le transférer à une entité extérieure par un contrat d'assurance. Nous résumons ici les principaux résultats de leur analyse.

a) L'auto-assurance

L'auto-assurance consiste pour l'individu à réaliser un investissement (coûteux) qui lui permet de diminuer le montant du dommage. C'est pour lui, au même titre que l'assurance, un moyen de redistribuer de la richesse entre les états du monde puisque son investissement a un impact sur le montant du dommage. S'il n'y a pas d'assurance, il va choisir un montant d'auto-assurance positif, sous réserve qu'il soit risquophobe et que sa propre productivité dans cette activité dépasse un certain seuil. On montre que

l'incitation à s'auto-assurer est plus faible lorsque les dommages sont rares¹². L'individu ayant, dans ce cas, une faible probabilité de tirer un bénéfice de son investissement, son coût relatif augmente et son intérêt diminue.

Lorsqu'assurance et auto-assurance sont simultanément possibles, l'auto-assurance est moins élevée que dans le cas précédent. Si le prix de l'assurance croît, l'individu lui substitue de l'auto-assurance (et vice versa).

b) L'auto-protection

L'auto-protection consiste pour l'individu à investir pour diminuer la probabilité de réalisation de dommages. Cette activité ne lui permet pas de redistribuer directement de la richesse entre les états du monde, ce qui est son objectif ultime lorsqu'il éprouve de l'aversion pour le risque. On montre, de fait, que l'auto-protection peut être entreprise par les individus quelle que soit leur attitude face au risque. Le niveau d'auto-protection dépend des goûts, de la productivité des individus, de leur revenu et des caractéristiques du dommage. L'étude des propriétés de la demande d'auto-protection dépend de la conjonction de l'ensemble de ces facteurs et le seul résultat stable est que l'auto-protection diminue quand sa productivité baisse.

Si la possibilité de s'assurer existe, l'auto-protection peut être découragée puisqu'il existe un mécanisme qui permet de transférer de la richesse entre les deux états du monde. A l'inverse, l'auto-protection, en diminuant la probabilité de dommage, fait baisser le coût de l'assurance. Elle peut donc théoriquement être stimulée par la possibilité de s'assurer. Les deux mécanismes peuvent ainsi se contrecarrer.

Si l'assurance est actuarielle, on montre que l'auto protection peut être plus élevée que dans le cas où l'assurance n'est pas disponible et que l'augmentation de la productivité dans l'activité d'auto-protection augmente à la fois la demande d'assurance et d'auto-protection. Les deux activités sont donc complémentaires. Si le prix de l'assurance est indépendant de la probabilité de dommage (toute variation de la probabilité de dommage est compensée exactement par celle du taux de chargement), la présence de l'assurance décourage l'auto-protection.

¹² alors que sous certaines conditions que nous avons présentées, celle de s'assurer est plus élevée.

4.2.2 Le risque moral

Dans l'analyse d'Ehrlich et Becker, déployer des efforts pour diminuer son niveau de risque ou transférer le risque sur un assureur représente de fait une alternative. L'individu choisit simultanément un niveau d'effort et d'assurance en fonction de ses seules préférences et du niveau de différentes variables exogènes (prix de l'assurance, dotations initiales...).

L'interaction entre la capacité de l'individu à moduler son niveau de risque et l'assurance est plus souvent étudiée dans le cadre d'une asymétrie d'information. Dans ce cas, le niveau "d'effort" :

- n'est pas observable par l'assureur ;
- il peut être modifié par l'individu après qu'il a signé le contrat d'assurance.

L'individu couvert par une assurance n'est pas incité à réaliser les investissements qui permettent de diminuer son risque dans la mesure où il n'en récolte pas les bénéfices. L'assureur doit alors proposer des contrats qui prennent en compte ce qu'on appelle le risque moral.

Ces contrats d'assurance ne pouvant être explicitement fondés sur le niveau d'effort qui n'est pas observable, ils portent sur les conditions d'assurance (prime et couverture en cas de dommage) et leur élaboration procède de la logique suivante :

1. le niveau d'effort sur lequel se fonde implicitement le contrat est celui qui maximise l'utilité de l'individu pour un contrat donné (contrainte d'incitation) ;
2. le profit de l'assurance est positif ou nul, compte tenu du niveau d'effort effectivement choisi ;
3. parmi l'ensemble des contrats possibles, on choisit celui qui maximise l'utilité de l'individu.

Winter (1992) propose une revue de littérature sur la forme des contrats en présence de risque moral.

Si l'effort de l'assuré conduit à une réduction de la probabilité de dommage, et si le coût de l'effort n'est pas trop élevé, on montre que le contrat doit comporter une franchise, l'assuré étant remboursé si la perte dépasse un certain seuil et pour la différence entre le

montant de la perte et ce seuil. Si le coût de l'effort dépasse un certain seuil, la franchise optimale est nulle.

Si l'effort de l'assuré permet de diminuer le montant du dommage aléatoire sans en modifier la probabilité de réalisation, et si l'individu a une aversion absolue au risque non croissante, le contrat couvre l'individu lorsque la perte est inférieure à un montant donné. Le contrat comporte en outre un copaiement au delà, copaiement qui ne diminue pas quand la perte augmente. Zeckhauser (1995) a récemment souligné l'importance de ce résultat dans le cas où les pertes peuvent être très élevées et en particulier quand l'absence d'effort des individus peut distordre la distribution de probabilité du dommage en faveur d'événements très rares mais très coûteux : la part optimale de l'assuré dans la couverture peut être élevée pour des niveaux de perte faible (franchise) mais aussi pour des dommages très coûteux (l'assureur peut éventuellement plafonner sa participation au remboursement du dommage).

CONCLUSION

Au terme de cette présentation des principaux résultats de la théorie économique en incertain, nous pouvons noter que :

1. les résultats de la théorie économique peuvent être étendus aux situations incertaines. Si les individus sont libres de déterminer l'allocation de leurs ressources et si les marchés sont parfaits, les transactions fondées sur l'intérêt individuel aboutissent à une situation Pareto optimale ;
2. les personnes qui font face à un risque individuel vont chercher à se prémunir contre la variation de leur richesse entre les états du monde. Parmi les mécanismes qui leur permettent d'améliorer leur bien-être, l'assurance est le plus répandu ;
3. l'existence d'aléa au niveau individuel peut se traduire par des situations dans lesquelles la richesse globale disponible dans l'économie varie. Toutefois, on peut considérer que la gestion de ce phénomène relève de la responsabilité des assureurs, dans le cadre de contrats de réassurance. Dans la mesure où ce sont les comportements des individus qui nous intéressent, il ne semble donc pas nécessaire d'approfondir l'analyse de la gestion du risque social. Cette démarche est légitimée par le fait que l'existence d'un tel risque ne modifie pas les mécanismes de partage du risque individuel ;
4. si l'on s'en tient à l'étude de l'assurance, il faut cependant prendre en compte la possibilité que peuvent avoir les individus d'influencer leur propre niveau de risque. Dans le cas où cet effort n'est pas observable, la théorie semble montrer que l'existence du risque moral peut être gérée par les assureurs, sans remettre fondamentalement en cause leur capacité à offrir de l'assurance.

CHAPITRE 2

LE RISQUE ET LA REPARATION DU DOMMAGE EN SANTE

Pour l'Organisation Mondiale de la Santé, la santé est la conjonction d'un état de bien-être physique, mental et social. Pour l'économiste, la notion de bien-être renvoie au concept de fonction d'utilité. L'analyse économique conduit donc à considérer que l'état de santé d'un individu est un des arguments de sa fonction d'utilité.

Or, cet état de santé est en partie aléatoire. Comme le souligne implicitement la définition de l'OMS, il est déterminé par les caractéristiques de l'individu et de l'environnement dans lequel il évolue. Ces facteurs de risque peuvent donner lieu à l'apparition de maladies. Si les circonstances le permettent, les individus malades vont avoir recours au système de soins pour rétablir, dans la mesure du possible, leur état de santé initial.

Le fait de tomber malade ou d'avoir un accident constitue pour l'individu un dommage, contre lequel il peut souhaiter se prémunir par la souscription d'une assurance. Pour réparer ce dommage, il va consommer des soins. Or, *"les soins de santé, sans être uniques, cumulent un ensemble de caractéristiques qui nécessite d'enrichir considérablement la boîte à outils standard... Si les économistes de la santé n'ont pu à ce jour apporter toutes les réponses espérées (attentes au demeurant souvent excessives), ils ont contribué à clarifier la nature des services sur ce marché"* (Rochaix, 1996, pp. 508).

Ce chapitre a pour objectif d'explicitier la place de l'incertitude dans la détermination de l'état de santé d'un individu (Section 1) et de rappeler comment l'analyse économique explique le recours aux soins (Section 2). Nous ne cherchons pas à présenter l'ensemble des apports de l'économie de la santé à *"la boîte à outils standard"*. Nous présentons donc les seuls aspects qui nous semblent pertinents pour comprendre en quoi la mise en œuvre du mécanisme d'assurance est conditionnée par la nature des biens que l'on étudie : santé et soins.

1. FACTEURS DE RISQUE ET NATURE DE L'ALEA EN SANTE

A une vision assez déterministe de l'état de santé, inspirée notamment des travaux précurseurs de Grossman (1.1), on peut en première analyse opposer une approche stochastique qui semble plus en phase avec notre problématique (1.2). Les conjuguer dans un cadre unique permet de cerner la nature du risque en santé (1.3).

1.1 L'approche déterministe

Le terme "déterminisme" recouvre deux notions :

- d'une part, l'idée que les individus interviennent activement dans la production de leur propre santé et que leurs comportements, y compris de recours aux soins, contribuent à ce processus. Ils déterminent leur état de santé ;
- d'autre part, l'idée qu'il existe des facteurs explicatifs de l'état de santé des populations : les *déterminants* de l'état de santé.

Grossman (1972) a explicité la relation entre santé et consommation de soins dans le cadre d'une analyse du cycle de vie (1.1.1). La relation de causalité entre utilisation des soins et détermination de l'état de santé peut néanmoins être nuancée par l'étude plus globale des déterminants environnementaux de l'état de santé (1.1.2).

1.1.1 Le déterminisme comportemental

L'association d'une consommation de soins à une variation de l'état de santé repose sur l'hypothèse de l'existence d'une fonction de production :

"health care is one argument in a very complex production function, not a utility function, whose output is (expected to be) improved health status. The connection between health care and health is an external, technical relation, about which specialized providers develop specialized knowledge" (Evans, 1983, pp. 78).

¹ La santé est un des arguments d'une fonction de production, et non d'une fonction d'utilité, très complexe, dont l'output est (ou devrait être) une amélioration de l'état de santé. Le lien entre soins et santé est une relation techniques exogène, sur laquelle des prestataires spécialisés acquièrent une connaissance spécialisée.

La clarification de la relation entre soins et santé peut être attribuée à Grossman qui, en 1972, a développé le premier modèle économique de comportement centré sur la santé : elle est introduite explicitement comme argument de la fonction d'utilité individuelle. Cette analyse décline, de manière plus fine, à la santé la théorie du "capital humain" développée par Becker (1965) qui vise à élargir le champ du calcul économique, notamment pour rendre compte du choix d'un niveau d'étude des décisions de migrer et des investissements dans la santé. Becker montre ainsi comment l'individu peut choisir de faire fructifier son capital humain (d'investir) en prolongeant ses études pour augmenter à terme son niveau de rémunération.

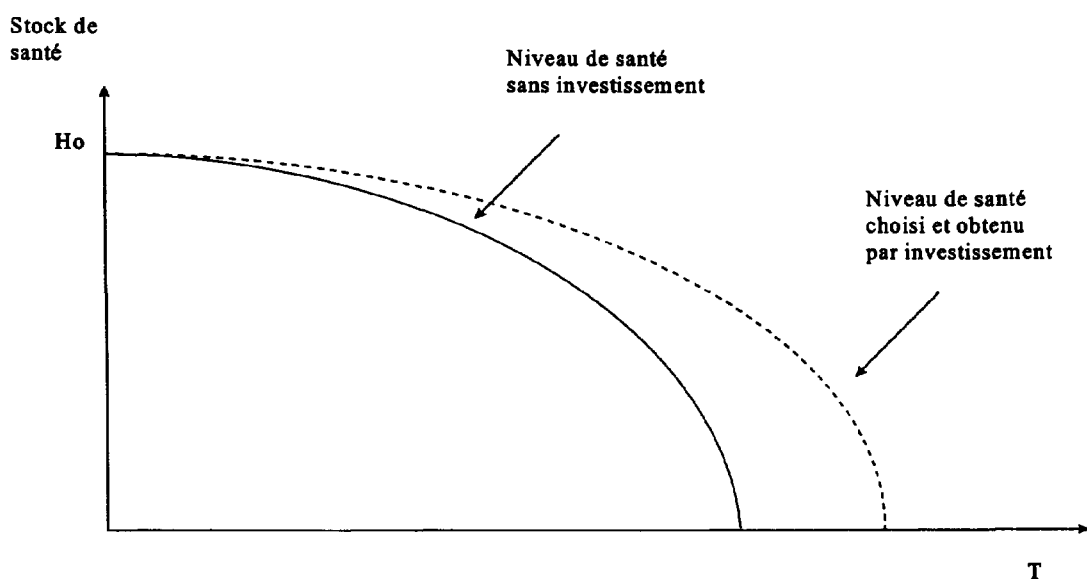
Grossman définit la santé comme un bien durable, dont un stock est alloué à l'individu au début de sa vie, et qui se déprécie au cours du temps. Sa diminution sous un seuil entraîne le décès. A chaque période, l'individu peut investir en temps passé à se soigner et en soins médicaux pour compenser la dépréciation "naturelle" de son stock de santé. Il maximise son utilité sous contrainte de ressources et de temps. Il détermine alors un niveau désiré d'état de santé, une demande de santé, et choisit le programme de consommation qui lui permet d'ajuster son stock au niveau désiré. Grossman décline son analyse en deux sous-modèles :

- dans le modèle dit de consommation, l'individu valorise l'amélioration de l'état de santé en tant que telle : être en meilleure santé augmente son bien-être ;
- dans le modèle d'investissement, l'amélioration de l'état de santé est valorisée par l'individu dans la mesure où elle lui permet d'augmenter son temps de travail et donc sa richesse.

Dans les deux cas, la demande de soins est dérivée de celle de santé. Comme le souligne Le Pen (1988), les modèles antérieurs introduisaient directement les biens et services médicaux dans la fonction d'utilité, et appliquaient la théorie classique du consommateur pour en déterminer la demande. L'apport le plus fondamental de Grossman a été de souligner que la consommation de soins n'est pas directement source d'utilité mais qu'elle permet de produire de la santé qui contribue au bien-être de l'individu. L'existence d'une telle causalité permet de justifier la consommation d'un bien qui n'est pas, a priori, source directe d'utilité. La "consommation" de soins peut, au contraire, être source de douleur et de stress, autant d'éléments qui doivent être sur-compensés par l'augmentation d'utilité liée à l'amélioration de l'état de santé.

La figure 1 illustre la définition que Grossman propose de l'état de santé d'un individu. Chacun naît avec un stock de santé H_0 , qui se déprécie à un taux croissant tout au long de sa vie. Il est possible d'investir pour compenser cette dépréciation et donc de modifier sa durée de vie. Par souci de simplification, on considère ici que le décès intervient lorsque le stock tombe à zéro. Nous utiliserons par la suite cette représentation de l'état de santé pour illustrer les différents types de risques auxquels l'individu peut être confronté.

Figure 1
Hypothèse de Grossman : Evolution du stock de santé avec le temps



L'auteur insiste, dans la définition de la santé qu'il propose, sur deux aspects qui nous intéressent particulièrement.

Il met tout d'abord en exergue l'impact du comportement individuel sur la santé puisque, dans le modèle de Grossman, l'agent économique décide en partie de sa durée de vie, compte tenu du taux de dépréciation de son stock de santé². De plus, si l'analyse de Grossman reste centrée sur les soins - "*the most important market good in the gross investment function*"³ (1972, pp.2), il reconnaît que la fonction de production est définie au niveau individuel et qu'elle comporte d'autres inputs, comme le capital humain. La

² L'impact du comportement individuel sur le taux de dépréciation de santé a été partiellement endogénéisé par la suite, cf. Claude Le Pen, (Levy, 1988).

³ *Le bien marchand le plus important dans la fonction d'investissement.*

production de santé par l'individu ne dépend en effet pas seulement de la quantité de soins qu'il consomme mais aussi :

- de son comportement individuel. Dowie (1975) et Ippolito (1981) ont, par exemple, formalisé dans un modèle de ce type, les choix de style de vie ou de comportements à risque ;
- de variables individuelles comme l'éducation, qui feront que, toutes choses égales par ailleurs, une personne sera plus ou moins incitée à investir dans son stock de santé et mieux à même de transformer des soins en une amélioration de l'état de santé.

Le second atout de cette approche en terme de stock est la prise en compte du déterminant biologique d'état de santé qu'est le vieillissement. La résolution du modèle de Grossman repose sur l'hypothèse que le taux de dépréciation du stock de santé augmente avec l'âge, à partir d'un certain point du cycle de vie. "*This correlation can be inferred because as an individual ages, his physical strength and memory capacity deteriorate*"⁴(Grossman, 1972, pp. 14).

La détérioration au cours du temps de l'état de santé et sa traduction par une augmentation des besoins médicaux, est une caractéristique inéluctable.

Grossman évoque brièvement la possibilité que le taux de dépréciation du stock de santé puisse être défini de manière aléatoire mais cette démarche n'aboutit pas à une définition précise des risques ni, a fortiori, à l'explication du besoin que peut ressentir l'individu d'avoir recours à une assurance, pour se prémunir contre eux.

Certains auteurs ont, par la suite, introduit l'incertitude dans le modèle (ainsi Dardanoni et Wagstaff (1987) introduisent une incertitude sur l'efficacité des soins), mais ils n'ont jamais été conduits à expliciter une demande d'assurance dans le cadre du modèle de Grossman.

Les tests empiriques portant directement sur ce modèle, que ce soit sous sa forme structurelle ou réduite, ainsi que ceux qui l'adoptent comme référence théorique, confirment (Le Pen, 1988) ses prédictions intuitives, mais conduisent à relativiser l'idée selon laquelle le niveau de santé résulterait d'une stratégie active d'investissement (dont

les soins seraient un input) : la relation entre état de santé et consommation de soins demeure en effet toujours significativement négative.

1.1.2 Le déterminisme environnemental

L'apparente contradiction entre l'approche de Grossman et l'idée que la santé est soumise à des chocs aléatoires assurables est renforcée par l'analyse plus détaillée des déterminants de la santé. Deux ouvrages récents "*Why are some people healthy and others not ?*" (1994, édité par Evans, Stoddart et Barer) et "*Tackling Inequalities in Health: An agenda for action*" (1995, édité par Benzeval, Judge et Whitehead) présentent les résultats d'études et de travaux d'épidémiologie qui soulignent que l'état de santé dépend non seulement du comportement des individus et de leur accès aux soins, mais aussi d'un ensemble de facteurs biologiques et environnementaux.

Parmi les facteurs biologiques, le sexe et l'âge ont une relation évidente avec l'état de santé. Au delà, que ce soit pour des raisons génétiques ou accidentelles, l'état de santé peut varier d'un individu à l'autre. L'exemple le plus simple est celui des maladies dont l'apparition est directement expliquée par la présence d'un gène (ainsi en est-il de l'hémophilie) qui fait qu'un individu éprouvera tout au long de sa vie des difficultés à "fonctionner", quel que soit son environnement. Plus généralement, le développement de la biologie génétique permet de montrer que certaines personnes sont plus susceptibles que d'autres de développer certains types de maladies, en raison de prédispositions génétiques (le diabète, les maladies cardio-vasculaires, certains cancers...). L'ensemble des facteurs biologiques qui ont une influence directe sur la santé sont des facteurs de risque strictement individuels et sur lesquels la personne n'a aucune influence.

L'environnement dans lequel les individus évoluent a aussi un impact déterminant sur leur état de santé. La qualité de l'air, de l'eau, ou plus généralement le niveau de pollution, mais aussi la salubrité de l'habitat, les conditions de travail,... sont quelques-unes des caractéristiques physiques de l'environnement qui ont une influence sur la santé. Ainsi, des enfants logés dans un environnement insalubre sont susceptibles d'être

⁴ Cette corrélation peut être déduite du fait que, lorsqu'une personne vieillit, ses forces physiques et sa mémoire se détériorent.

intoxiqués par la peinture à plomb qui "*même à faible dose est responsable d'une diminution de QI, de retards psychomoteurs, mais aussi de diminution des capacités d'apprentissage et de développement*"⁵. De même, l'environnement économique et social a une influence déterminante sur la santé.

A cet égard, deux remarques s'imposent :

- d'une part, la frontière entre ce qui relève de l'environnement "physique" et "social" n'est évidemment pas claire, ainsi que l'illustre le cas du saturnisme ;
- d'autre part, si l'environnement constitue indéniablement une variable exogène pour l'individu, en particulier au début de son existence, il influence aussi la détermination d'autres variables socio-économiques individuelles (le revenu, le patrimoine, l'emploi, le niveau d'éducation), ou comportementales (le fait d'adopter certains comportements à risque comme fumer, la qualité des relations sociales...) qui ont un impact plus ou moins direct sur l'état de santé. La prise en compte de l'environnement (considéré comme une variable partiellement exogène et déterminée collectivement) parmi les facteurs explicatifs de l'état de santé ne doit pas conduire à nier l'impact des comportements individuels sur l'état de santé, mais elle souligne à nouveau l'existence d'un certain déterminisme.

Le sens de la relation entre la qualité de l'environnement économique et social et celle de l'état de santé est connu. Moins les personnes sont favorisées, plus elles présentent des niveaux de mortalité et de morbidité élevés. L'existence de ce gradient entre différentes classes économiques ou sociales est documentée dans la quasi-totalité des pays développés. Ce gradient reste en outre observable sur une longue période en dépit des améliorations significatives de l'étendue, de l'efficacité et de l'accessibilité des soins (Hertzman *et al.*, 1994, pp. 69).

Au terme de cette analyse, il apparaît que la santé d'un individu est une combinaison de facteurs de risques :

- exogènes et déterministes. Dans ce cas, les facteurs explicatifs peuvent être déclinés au niveau individuel (facteurs biologiques) ou au niveau collectif (facteurs environnementaux au sens large) ;

⁵ Rapport sur le saturnisme publié par Médecins du Monde en Juillet 1997.

- endogènes. Le risque est alors généré par le comportement des individus ou des groupes auxquels ils appartiennent, directement ou par le truchement de leur impact sur d'autres variables.

1.2 L'approche stochastique

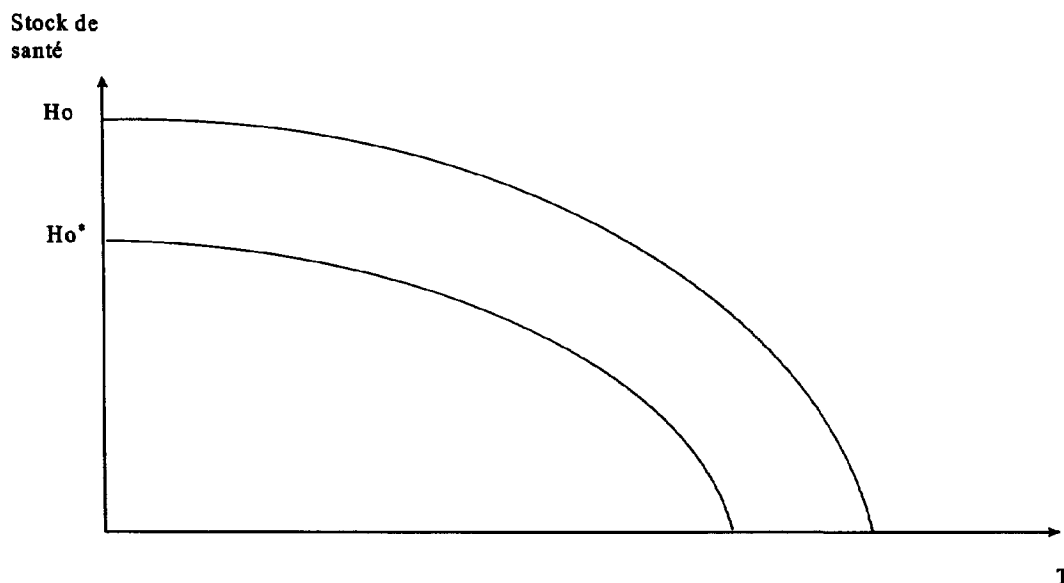
L'ensemble des déterminants que nous avons évoqués se traduisent à la fois dans l'état de santé général de l'individu mais, aussi par le fait qu'il va être atteint de certaines maladies ou pas. Les facteurs de risque se concrétisent par des "chocs" qui génèrent un besoin de prise en charge médicale. En reprenant la vision que propose Grossman de l'état de santé, on peut montrer comment l'incertitude qui pèse sur la détermination de l'état de santé caractérise l'individu tout au long de sa vie (1.2.1) puis, comment elle se traduit, à un moment donné, par la réalisation d'un dommage (1.2.2). L'objectif n'est pas ici d'amender le modèle de Grossman, qui est exclusivement centré sur les comportements individuels face à la santé, alors que notre objet d'étude est le marché du service d'assurance. Il nous semble en revanche intéressant d'illustrer les différents types de risques que nous évoquons par une approche graphique qui s'inspire des mêmes prémices (la santé est un bien durable dont le niveau décroît dans le temps).

1.2.1 Incertitude initiale

Le fait qu'un stock de santé initial ou qu'un taux de dépréciation plus ou moins élevé puisse être alloué "par la nature" à un individu représente un premier type de risque. Ces différences sont illustrées dans les figures 2 et 3 (page suivante).

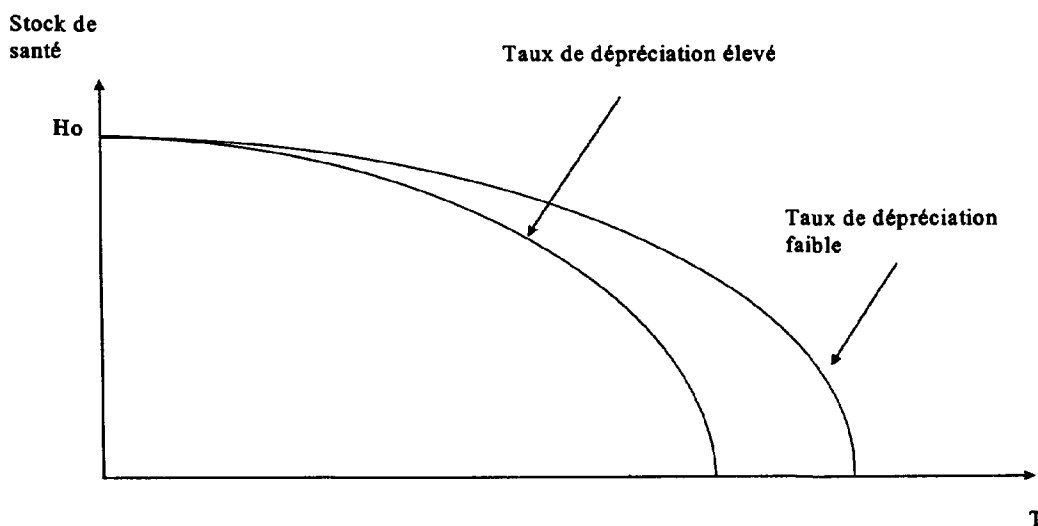
Le premier cas pourrait illustrer l'existence d'un handicap à la naissance, par exemple une surdit  cong nitale, qui n'a pas, en soi, d'impact sur le taux de d pr ciation du stock de sant .

Figure 2
Incertitude sur le niveau du stock initial de santé alloué



Dans le second cas, le taux de dépréciation diffère entre deux individus tout au long de leur vie (Figure 3).

Figure 3
Incertitude sur le taux de dépréciation du stock de santé alloué



L'exemple qui illustrerait au mieux ce cas est celui d'une maladie extrêmement rare appelée la Progeria (ou syndrome de Hutchinson-Gilford). Les enfants atteints, selon l'*Electronic Textbook of Internal Medicine* sont normaux à la naissance mais souffrent

dès leur plus jeune âge de symptômes et de maladies associées au grand âge, comme la sénilité, l'arthrose... Tout se passe comme s'ils subissaient un vieillissement très accéléré et la majorité d'entre eux décède avant l'âge de 13 ans.

Les deux illustrations que nous avons proposées sont assez caricaturales et la réalité relève plus souvent d'une combinaison des deux approches⁶. Par exemple, certains handicapés ont un capital santé plus faible que la moyenne au départ qui peut aussi se dégrader plus vite car ils sont exposés à des complications particulières (trisomie 21). De la même manière, si un individu naît avec une insuffisance rénale chronique, sans aucune autre tare médicale, son stock de santé initial est moins élevé que la normale. L'insuffisance rénale va entraîner peu à peu une anémie chronique et une déminéralisation des os ; elles entraîneront à leur tour des troubles de croissance qui eux-mêmes sont susceptibles de se traduire par des anomalies du développement psychomoteur... Le taux de décroissance du stock sera donc pour ces personnes anormalement élevé.

En tout état de cause, le fait de naître avec un stock de santé ou un taux de dépréciation différent constitue donc un premier type de risque sur la santé, qui peut se traduire par un besoin particulièrement important de prise en charge médicale tout au long de l'existence.

1.2.2 Maladie et dommage

L'apparition d'une maladie et la survenue d'un accident peuvent être considérées comme des chocs aléatoires sur l'état de santé. La référence à la notion de dommage permet ici de se rapprocher des concepts traditionnellement utilisés en économie de l'assurance. Il est alors naturel d'étudier les risques selon deux dimensions : la probabilité de subir un choc (a) et la nature de ce choc (b).

Cette analyse est complémentaire de la précédente en ce que ces dommages sont des événements aléatoires qui peuvent affecter l'état de santé, quelle que soit la situation initiale de l'individu.

⁶ Je remercie Laure Com-Ruelle, Hervé Picard et Catherine Sermet, médecins et chercheurs au CREDES pour leur aide dans la recherche d'illustrations et les discussions qu'elles nous ont values.

On peut aussi estimer que l'approche en terme de dommage constitue une vision plus détaillée de l'analyse précédente : une catégorie d'individus souffrant d'une maladie particulière (et donc subissant un dommage à titre individuel) présente, en moyenne, un taux de dépréciation relativement plus élevé ou un stock plus faible que ceux de la population générale.

a) Probabilité de réalisation

Un estimateur naturel de la probabilité de subir un dommage est la prévalence de la maladie considérée dans l'ensemble de la population. Cette approche doit cependant être nuancée au niveau individuel. La réalisation d'un événement que l'on qualifie ici de façon générique de "choc" peut être perçue comme aléatoire ou non. La contraction d'une maladie est, dans certains cas, inévitable (probabilité égale à un). A l'inverse, certains "chocs" pourraient être systématiquement évités (probabilité nulle). Le "choc" est donc évitable ou inévitable par l'individu.

Une fois exclus ces deux cas extrêmes, on peut discerner des facteurs d'hétérogénéité entre les individus. La probabilité de réalisation d'un choc particulier varie, comme nous l'avons signalé, pour des raisons exogènes et déterministes ou des raisons liées au comportement de l'individu. Dans ce dernier cas, il influence directement la probabilité de dommage, soit par des mesures de prévention spécifiques, soit en n'adoptant pas des comportements dangereux.

b) Nature du choc

Ampleur et réversibilité

L'ampleur du choc constitue une première dimension naturelle d'analyse. La mesure de l'état de santé et a fortiori d'une dégradation de cet état de santé sont des exercices difficiles. Mesurer l'état de santé nécessite de prendre en compte de multiples dimensions objectives et subjectives. Des indices synthétiques peuvent être construits et combinés pour évaluer l'état de santé d'un individu comme, par exemple, les indicateurs de risque vital et d'invalidité développés par le CREDES⁷, qui résument "*l'influence de la morbidité sur le pronostic vital ... et l'influence des diverses pathologies sur le niveau*

⁷ Dumesnil, Grandfils, le Fur et al (1997). *Santé soins et protection sociale en 1996*, Rapport CREDES N°1204, nov.

d'invalidité". Ces outils ne sont pas pour autant utilisés pour mesurer l'ampleur d'un choc particulier que subirait un individu sur son état de santé. La mesure de l'ampleur du choc renvoie plus traditionnellement à une valeur monétaire et au coût d'une maladie.

Le second élément de caractérisation des chocs porte sur leur réversibilité. Au delà des accidents ou maladies qui se traduisent par le décès de la personne, on peut établir une distinction entre des chocs réversibles et irréversibles. Cette caractéristique dépend évidemment pour partie de la nature de la maladie. Elle est aussi conditionnée par :

- d'une part, l'existence ou non d'une technique médicale qui permet de restaurer l'état de santé initial ;
- d'autre part, l'accès qu'a l'individu à ces soins (pour des raisons techniques ou financières).

Les figures 4 et 5 illustrent respectivement ces deux types de chocs sur le stock d'état de santé : la figure 4 représente une altération irréversible portant sur le stock de santé mais n'entraînant pas le décès immédiat ; la figure 5 illustre la notion de choc réversible sur le stock.

Figure 4
Choc irréversible sur le stock de santé

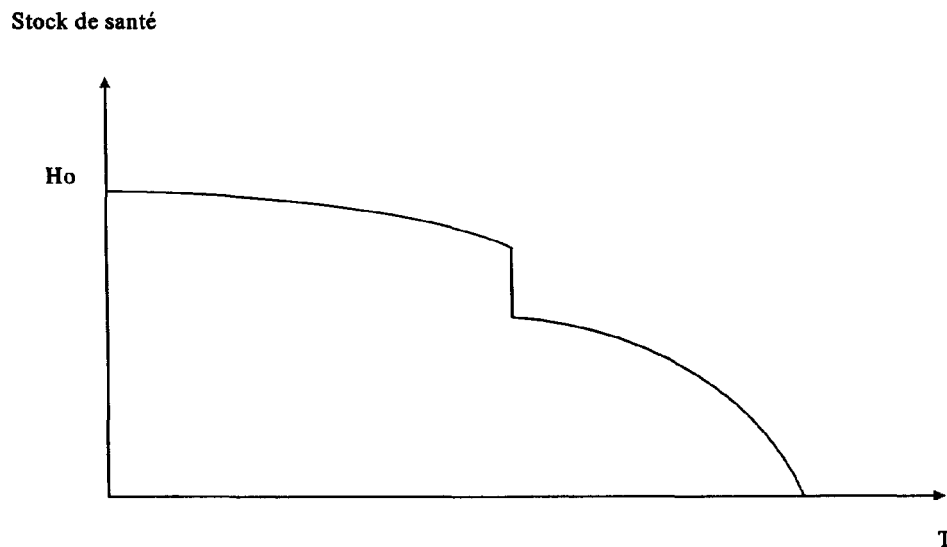
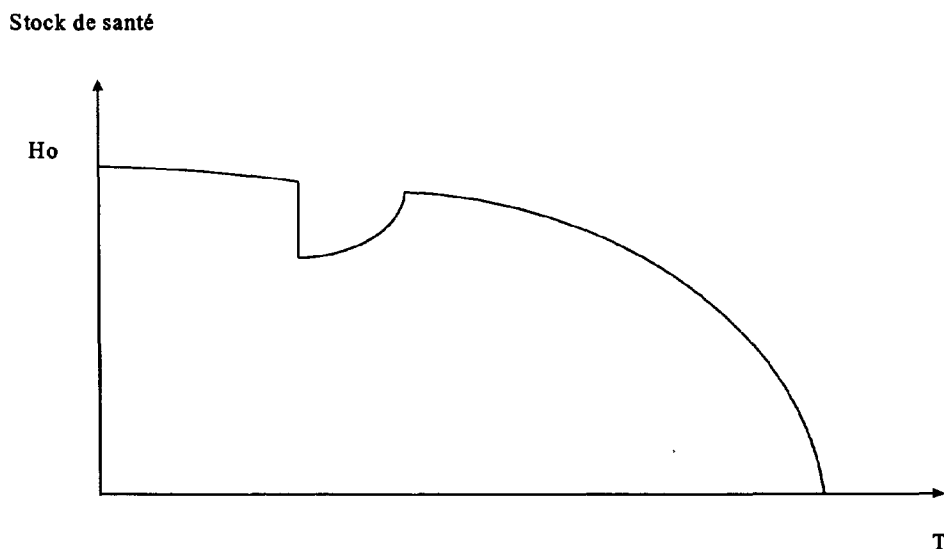


Figure 5
Choc réversible sur le stock de santé

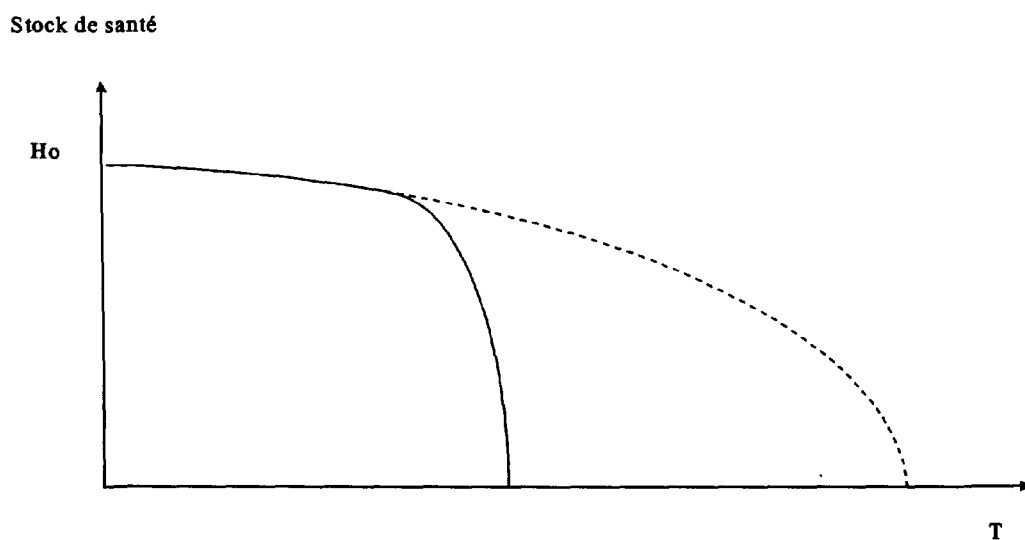


Choc sur le stock ou choc sur le taux

Un dommage peut théoriquement se décliner en deux modalités, selon qu'il porte sur le stock de santé ou le taux de dépréciation de ce stock (sachant qu'une combinaison des deux dimensions est généralement nécessaire pour décrire rigoureusement la variation de l'état de santé induite par une maladie). Ces modalités correspondent à deux types de risques fondamentalement différents. Une fracture ou plus généralement un accident correspond par exemple à des dégradations brutales de l'état de santé (figure 4 ou 5).

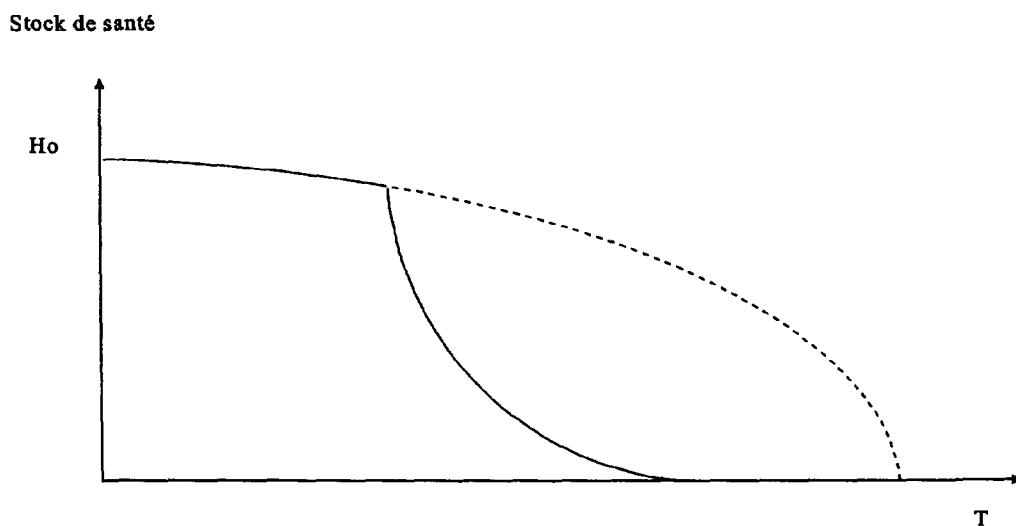
En ce qui concerne le choc sur le taux de dépréciation, on peut envisager là encore deux cas extrêmes, selon le mode d'évolution de la maladie. Un cancer correspond plutôt à une augmentation du taux de dépréciation, comparable à celle représentée dans la figure 6 (page suivante).

Figure 6
Altération du taux de dépréciation
Maladie de type cancer



Une maladie qui conduirait à une dégradation initiale brutale mais se ralentissant progressivement se traduirait par une évolution plus proche de celle qui est représentée dans la Figure 7. Les démences pourraient représenter un exemple de ce type d'évolution.

Figure 7
Altération du taux de dépréciation
Maladie de type démence



L'assimilation d'une maladie à un dommage et l'explicitation des caractéristiques de ce dommage conduisent à souligner que, exception faite du choc réversible, les dommages sur l'état de santé s'inscrivent dans le temps, ce ne qui manquera pas de soulever des difficultés pour l'assurance.

1.3 Synthèse

Le vieillissement est une donnée biologique universelle fortement corrélée avec l'état de santé. Articuler la présentation des autres facteurs de risque et de leur traduction en termes de dommage autour de l'âge, caractéristique individuelle objective et observable, permet de souligner que :

- leur importance relative varie en fonction de l'âge des individus ;
- mais aussi que l'apparition ou la révélation des autres facteurs de risque et leur traduction par des dommages ne sont pas nécessairement simultanées.

Hertzman, Frank et Evans (1994, pp. 74) proposent un découpage du cycle de vie en quatre périodes, chacune étant caractérisée par la domination d'un type de risque. En se fondant sur cette analyse, on peut proposer une brève synthèse qui souligne le lien entre les éléments déterministes et stochastiques présentés au long de cette section.

1. La période périnatale (*perinatal period*) qui est, en soi, une période de forte vulnérabilité, est aussi une période durant laquelle une grande partie des différences inter-individuelles initiales de stock ou de taux de dépréciation peuvent être décelées.
2. La période des "mésaventures" (*the period of misadventure*) s'étend de la fin de la première année de vie jusqu'au milieu de la quarantaine. Durant cette période, la santé est avant tout soumise à des risques "individuels accidentels" au sens large du terme (qui correspondent en majorité à des chocs réversibles). La notion de choc *individuel* utilisée par Hertzman, Frank et Evans renvoie au constat que la menace de grandes épidémies (à l'exception du SIDA) a largement été éliminée dans les pays développés.

Les auteurs remarquent aussi que l'état de santé des individus de cette tranche d'âge est marqué par l'apparition, chez certaines personnes relativement jeunes, de maladies résultant de dispositions génétiques ou d'exposition in utero ou périnatale. La prévalence de ces affections s'est progressivement accrue (ou est devenue plus visible), du fait de la disparition des épidémies mais aussi, paradoxalement, grâce au développement du traitement même de ces maladies, qui a permis une augmentation de la durée de vie de ceux qui en sont victimes.

3. La période suivante est celle des maladies chroniques dégénératives (*chronic degenerative diseases*) comme les problèmes cardiaques, les attaques, les cancers, l'arthrose... Ceci correspond, selon notre typologie, à l'apparition dans ces tranches d'âge d'un nombre de plus en plus élevé de chocs de plus grande ampleur parfois irréversibles. Ces chocs peuvent porter sur le stock de santé et son taux de dépréciation. Le terme de "degenerative", choisi par les auteurs, souligne le fait que l'apparition de ces maladies est pour partie explicable par l'histoire des individus pendant les périodes précédentes. Le déterminisme comportemental (style de vie, investissements dans le capital santé) et l'environnement (physique ou social) jouent un rôle prépondérant dans l'explication de l'apparition de ces pathologies.
4. Le début de la dernière période de sénescence généralisée (*generalized senescence*) se situe environ après 75 ans. Les maladies qui apparaissent sont globalement les mêmes qu'à la période précédente mais elles sont souvent combinées et leur impact sur un organisme affaibli fait qu'il devient difficile de définir une cause unique de dégradation de l'état de santé ou de décès.

La santé est donc bien au cœur d'un processus de production, qui caractérise l'individu et ne se réduit pas à la seule expression d'une relation mécanique entre soins et santé. Les individus ont une part de responsabilité dans leur état de santé, mais ils peuvent aussi subir les effets de facteurs de risques dont ils ne sont guère responsables et qui se traduiront par des différences dans la fréquence et le type de dommage auxquels ils auront à faire face.

2. DETERMINANTS DE L'UTILISATION DES SOINS

Pour un état de santé donné, la fonction de production permet théoriquement de déterminer la quantité de soins dont l'individu a objectivement besoin et, indirectement, le coût de l'assurance, en dehors de toute référence à des transactions et à des choix individuels. Or, pour bénéficier de soins, l'agent économique doit effectuer des transactions sur un marché particulier dont les caractéristiques ont été abondamment étudiées en économie de la santé. La modélisation de la demande de soins (de biens et services médicaux) développée dans les années 1970 repose sur une vision instrumentale de ce bien : l'individu consomme des soins pour atteindre l'état de santé qu'il a choisi et arbitre rationnellement entre cette consommation et les autres.

Nous exposons ici les éléments qui permettent d'enrichir cette approche, en conservant à l'esprit l'objectif de ce chapitre : mieux comprendre l'environnement dans lequel s'inscrit l'assurance maladie.

Partant d'une formalisation simple de la fonction de production de santé, nous décrivons les grands traits de la demande de soins (2.1), analyse progressivement enrichie par l'étude approfondie des outputs de la fonction de production (2.2). Nous rappelons ensuite en quoi l'hypothèse selon laquelle la demande de soins est l'expression d'une décision individuelle et parfaitement rationnelle semble manquer de réalisme (2.3).

2.1 Fonction de production et demande de soins

L'hypothèse que l'on connaît la fonction de production permet d'intégrer la demande de soins dans le cadre économique standard. Puisque les soins sont un bien marchand, ils entrent explicitement dans la contrainte budgétaire. En maximisant son utilité, l'individu détermine simultanément la quantité de soins qu'il consomme et l'état de santé qu'il atteint.

Phelps (1973) présente un exemple simple de cette démarche. Il considère une fonction d'utilité individuelle $U = U(x, H)$ où x est un panier de biens de consommation et H l'état de santé.

$$H = H_0 - \lambda + g(h)$$

H_0 est l'état de santé initial, λ est une variable aléatoire qui représente l'apparition d'une maladie ; ces deux variables sont déterminées de façon exogène. La dégradation de l'état de santé peut être compensée par la consommation d'un bien "de réparation et de remplacement" h , le soin médical. La fonction de production g a les caractéristiques classiques d'une fonction de production : la consommation de soins augmente l'état de santé de l'individu à un taux décroissant. Sur le plan analytique, ces hypothèses se traduisent de la façon suivante :

$$g'(h) > 0 \text{ et } g''(h) \leq 0$$

L'individu rationnel cherche à maximiser son utilité :

$$U = U(x, H_0 - \lambda + g(h))$$

Les soins médicaux ayant un prix unitaire p_h et le bien de consommation un prix p_x , la contrainte budgétaire est :

$$I = p_x x + p_h h$$

Ce modèle aboutit donc à la détermination d'une fonction de demande de soins dont les arguments sont les prix relatifs, le revenu et le niveau du dommage. Plus précisément, le prix des biens et le revenu deviennent les variables centrales de décision dans le choix, par l'individu, de son niveau de santé. Si un individu consomme une unité supplémentaire de soins, il réalise un arbitrage : pour lui, l'utilité marginale de cette consommation est supérieure au prix du bien.

On peut alors analyser empiriquement la demande de soins sur la base des hypothèses conventionnelles et de l'information disponible concernant les inputs de la fonction de production. Les biens et services que les personnes consomment sont ceux qui maximisent leur utilité. Autrement dit, : "*the act of demanding one set of goods implies that a person is better off than with any other bundle of goods that he or she can afford*"⁸ (Rice, 1997, pp. 402)".

⁸ Le fait de demander un panier de biens précis implique que la personne retire une utilité supérieure de la consommation de ce panier que de celle de toute autre combinaison de biens qu'elle a les moyens d'acquérir.

Les études empiriques ont, depuis longtemps, montré que l'utilisation de soins par les individus dépend du prix des biens et services médicaux ainsi que de leur revenu. La conception plus large de la fonction de production que nous avons présentée, permet de prendre en compte les paramètres de cette fonction comme autant de facteurs explicatifs de la consommation. Le niveau d'éducation est notamment une variable explicative significative de la consommation.

2.2 Caractérisation des outputs de la fonction de production

Nous savons que les soins ne sont pas le seul input de la fonction de production. Il faut toutefois compléter cette analyse car les soins ne produisent pas nécessairement de la santé (2.2.1), ni d'ailleurs seulement de la santé (2.2.2) et n'affectent pas uniquement l'individu qui en bénéficie (2.2.3).

2.2.1 Les soins ne produisent pas nécessairement de la santé

Une fois réalisé le choc aléatoire sur l'état de santé qui motive le recours aux soins, la consommation de soins est elle-même génératrice d'incertitude à plusieurs égards :

- tout d'abord, la réaction du patient à un traitement n'est pas toujours identique, ce qui modifie fondamentalement la stratégie thérapeutique qui sera adoptée. On ne sait donc pas prédire avec exactitude ex ante le montant des ressources médicales qui sera associé à l'apparition d'une maladie ;
- ensuite, les techniques utilisées comportent un risque thérapeutique : les soins administrés sont eux-mêmes susceptibles de provoquer une dégradation de l'état de santé ;
- enfin, le consommateur qui s'adresse à plusieurs prestataires de soins peut être amené à entreprendre des traitements qui s'avèrent incompatibles et dont les effets iatrogènes éventuels se traduisent par une nouvelle dégradation de l'état de santé et non par l'amélioration attendue.

L'incertitude porte donc sur l'estimation des moyens qu'il faut mettre en œuvre pour soigner un problème de santé précis et sur leur impact réel pour un patient. Sans aller jusqu'à dire que "*recovery from disease is as unpredictable as is its incidence*"⁹, (Arrow, 1963, pp 951), la définition objective d'un besoin associé à une affection est un exercice délicat. Dans une perspective de moyen ou long terme, cette difficulté est d'autant plus grande que le progrès et la diffusion des techniques médicales peuvent modifier brutalement le coût d'une procédure, comme le mode de prise en charge d'une pathologie.

2.2.2 Les soins ne produisent pas uniquement de la santé

La consommation de certains soins peut tout d'abord augmenter l'utilité des individus indépendamment de leur pouvoir curatif. A l'inverse, l'utilisation de certains biens et services médicaux, qui ne contribuent pas directement à l'amélioration de l'état de santé, peut présenter un intérêt.

Une partie des soins consommés peut être regroupée sous l'appellation de "caring" par opposition au "curing". Cette notion de caring rend compte d'une consommation motivée par un besoin d'attention spécifique, d'information, qui ne rentre pas dans la définition objective du besoin que nous avons donnée ou dans l'étude des fonctions de demande. Parmi les exceptions, on peut noter l'exemple de Clark et Olsen (1994) qui distinguent :

- une utilité qui vient du fait de recevoir des soins (*process utility*). Elle reflète le fait que les soins reçus produisent "*information, caring, warmth, interest, anxiety reduction, communication*"¹⁰ ;
- une utilité directement liée au résultat de ces soins (*outcome utility*).

Par ailleurs, la consommation de soins préventifs et de diagnostic soulève, elle aussi, des difficultés au regard de l'hypothèse de l'existence d'une fonction de production. En effet, ces soins ne contribuent, par définition, pas ou peu à la production directe de santé.

⁹ La guérison d'une maladie est aussi difficile à prévoir que sa survenue.

¹⁰ De l'information, de l'attention, de la chaleur, de l'intérêt, une réduction de l'anxiété, de la communication.

En revanche, consommer des soins préventifs ou de diagnostic peut permettre de :

1. diminuer la probabilité de réalisation d'un dommage. Ceci caractérise essentiellement les soins préventifs. La vaccination constitue un exemple concret de ce que la terminologie médicale qualifie de "prévention primaire" ;
2. diminuer le montant final du dommage. Une maladie diagnostiquée à un stade avancé est en général plus difficile et plus coûteuse à soigner. Les actes médicaux qui permettent un diagnostic précoce participent de la "prévention secondaire".

L'analyse économique de l'allocation des ressources au traitement ou à la prévention (primaire et secondaire) d'une maladie doit nécessairement être menée dans un cadre intertemporel, puisque la prévention intervient ex ante et permet de modifier la nature ultérieure du dommage. Deux modèles proposent une telle approche. Le premier ne prend en compte que la prévention primaire (Heffley, 1982) mais il introduit une double incertitude : la prévention a un impact sur la probabilité de ne pas tomber malade et les soins influencent la probabilité de redevenir en bonne santé. Les agents cherchent à maximiser le temps en bonne santé et arbitrent entre prévention et soins. Les simulations effectuées à partir du modèle montrent une très grande sensibilité du résultat aux paramètres du modèle¹¹ et l'existence probable de solutions en coin. En particulier, une personne pauvre peut avoir intérêt rationnellement à allouer l'intégralité de ses ressources à la prévention ou aux soins. Le modèle montre que le partage optimal entre prévention et soins doit être analysé pour chaque maladie (caractérisée par sa probabilité de réalisation et de guérison ainsi que la capacité de la prévention et des soins à faire évoluer ces probabilités). Le partage optimal, pour une affection donnée peut varier radicalement en fonction des ressources disponibles.

Le modèle de Eeckoudt, Godfroid, et Marchand (1998) étudie l'arbitrage entre médecine curative et prévention primaire (qui réduit la probabilité de dommage) puis secondaire, qui diminue le montant du dommage. Dans chaque cas, ils montrent qu'il existe un mélange optimal entre médecine curative et préventive et détaillent la statique comparative du modèle. Les résultats sur la prévention secondaire sont assez intuitifs et

¹¹ Les limites inférieures et supérieures des probabilités p_{11} et p_{21} , mais aussi l'impact marginal des soins sur ces probabilités.

démontrent que prévention secondaire et médecine curative sont des substituts : une sur-consommation de l'une est compensée par une sous-consommation de l'autre.

Les résultats sur la prévention primaire sont plus délicats à interpréter : si on consomme moins de soins que dans la situation optimale (rationnement), on compensera cet effet par une sur-prévention, ce qui va dans le sens de l'hypothèse de substitution. Par ailleurs, si on consomme plus de soins que nécessaire, on fera aussi plus de prévention que nécessaire, ce qui peut s'expliquer par une tentative de diminuer la probabilité d'avoir à supporter les conséquences d'un choix irrationnel de trop consommer de soins. En revanche, quel que soit le niveau de prévention choisi, le niveau optimal de soins reste le même. Ce résultat reflète le fait que l'état de santé et les soins qui permettent de l'ajuster au niveau désiré sont les variables centrales qui intéressent l'agent : on utilise la prévention primaire si on n'obtient pas l'optimum de soins, mais on ne la valorise pas en tant que telle. Comme la prévention secondaire a un impact plus direct sur l'état de santé, elle constitue un substitut plus proche aux soins. Ce modèle, qui prolonge l'analyse d'Ehrlich et Becker (1972) et l'applique à la santé, décrit les arbitrages entre ce que, dans un autre contexte, on appelle auto-protection et auto-assurance. L'individu n'a cependant pas explicitement recours à l'assurance dans ce modèle.

2.2.3 Les soins produisent des externalités

La consommation de soins peut générer des externalités : le fait qu'un individu ait accès à des soins peut avoir un impact direct sur l'utilité d'un autre individu. Ces externalités, susceptible de faire échouer l'allocation optimale des ressources par les marchés, sont regroupées en deux catégories.

a) Les externalités physiques

Dans le domaine des soins, des externalités physiques positives sont générées par le traitement ou la prévention des maladies contagieuses. En effet, plus vite on éradique une maladie contagieuse et plus les personnes vaccinées sont nombreuses, moins chaque personne court, à titre individuel, le risque de tomber malade. En présence d'une telle externalité, l'allocation optimale des ressources ne pourra être atteinte par des mécanismes concurrentiels. En effet, un individu n'a pas intérêt à financer l'achat d'un vaccin à titre individuel, dans la mesure où, si tous les autres le font, il peut bénéficier d'une réduction de sa probabilité de tomber malade.

b) Les externalités psychologiques

Les externalités psychologiques rendent compte :

1. soit de la notion de sympathie au sens premier du terme : "*individuals are affected by others' health status for the simple reason that most of them care*"¹² (Culyer 1976 pp. 89).

Cette forme d'altruisme peut être formalisée par l'intégration dans la fonction d'utilité d'un individu, non seulement de ses propres consommations de certains biens, mais aussi :

- de la consommation par d'autres individus, voire par certaines catégories d'individus, de ces mêmes biens (ici les soins), ce qui constitue une approche qualifiée parfois de paternaliste ;
- du niveau général d'offre d'un bien public (Rose-Ackerman, 1996, pp. 712-713).

On rend ainsi compte du fait que les individus attachent de l'importance au résultat de leur altruisme et qu'ils sont prêts à sacrifier des ressources pour que d'autres bénéficient de soins ;

2. soit d'une utilité spécifique que les individus tirent de l'acte de donner ou plus généralement du fait d'agir dans l'intérêt général (*group interest*). Margolis (1982) propose la définition d'un altruisme participatif qui traduit l'idée que "*individuals gain utility from participating in social acts*"¹³. L'individu partage ses ressources entre ses intérêts individuels et les intérêts du groupe, tout en agissant de façon égoïste.

La question de l'existence d'externalités psychologiques et la prise en compte de l'altruisme peuvent être utilisés pour endogénéiser la décision de certaines catégories d'individus de participer au financement de services de santé dont ils ne sont pas les bénéficiaires exclusifs (Clark et Olsen, 1994). On pourrait utiliser une démarche similaire pour expliquer la subvention par certains individus de l'acquisition d'assurance maladie par d'autres. Cependant, la spécification de fonctions d'utilité "altruistes" oblige à poser des hypothèses qui sortent du cadre standard et peuvent sembler ad-hoc.

¹² Les individus sont affectés par l'état de santé des autres pour la simple raison que la plupart d'en eux s'en soucie.

¹³ Les individus retirent une utilité du fait de participer à des actes sociaux.

2.3 Souveraineté de la décision de consommer des soins

La question de savoir si la consommation de soins peut être considérée comme l'expression d'une demande souveraine est celle qui a suscité le plus de débats en économie de la santé. Différents arguments sont avancés qui, à l'exception du premier, rendent compte du fait que l'information sur le marché des soins est imparfaite.

2.3.1 Hypothèse de rationalité individuelle

Si Arrow remarque que *"the demand for medical services is associated ... with an assault on personal integrity"¹⁴* (1963, pp. 949), Culyer (1973, pp. 52) détaille les circonstances dans lesquelles la consommation d'un individu, du fait même de sa maladie, n'est pas l'expression d'une demande au sens où l'économiste l'entend traditionnellement. Il semble évident que des personnes, dont l'intégrité mentale ou physique est altérée, ne sont pas à même de révéler leurs préférences. De même, on peut remarquer que des individus, qui doivent être considérés comme cliniquement malades, peuvent ignorer cet état de fait et n'exprimer aucune "demande" alors qu'ils auraient "besoin" d'être soignés.

2.3.2 Transparence du marché

On peut ensuite remarquer que, sur le marché des soins, *"the product and the activity of production are identical. (...) the customer cannot test the product before consuming it"¹⁵* (Arrow, 1963, pp.949). Cette caractéristique est partagée par de nombreux services (restauration, conseil, ..). En revanche, le patient n'a pas toujours l'opportunité de tester et comparer la nature du service qu'il pourrait recevoir en consultant différents offreurs, puisque son recours est conditionné par l'occurrence d'une affection, dans des conditions précises, et que, plus généralement, il n'a pas les moyens de juger de la qualité des soins dont il bénéficie dans toutes leurs dimensions (Culyer, 1973, pp.57).

¹⁴ *la demande de soins est associée à une atteinte à l'intégrité de la personne.*

¹⁵ *le produit et l'activité de production sont identiques. ... le consommateur ne peut tester le produit avant de le consommer.*

2.3.3 Choix non éclairés et rôle du producteur

La question du manque d'information du patient sur les soins dont il a besoin en cas de réalisation d'un dommage, ne peut être résumée à un problème de transparence du marché. L'hypothèse de souveraineté du patient va à l'encontre de l'intuition selon laquelle le futur patient est assez largement ignorant de son état de santé et des moyens qu'il faudrait mettre en œuvre pour l'améliorer. En initiant un épisode de soins, le patient qui présente un certain nombre de symptômes, exprime une double demande de diagnostic et de prise en charge médicalisée. Le médecin est le dépositaire d'une connaissance qui lui permet d'établir ce diagnostic, puis de mettre en œuvre ou de prescrire un traitement. Cette asymétrie informationnelle est largement irréductible : le savoir que détient le médecin est éminemment complexe et quand bien même on mettrait toute l'information à la disposition du patient, il ne serait probablement pas capable de l'utiliser. Autrement dit : "*Information in the form of skilled care is ... what is being bought from most physicians*"¹⁶ (Arrow, 1963, pp. 946).

La première conséquence est que le patient, en s'adressant au système de soins, ne connaît pas ex ante les coûts qui vont être générés. Ensuite, l'ignorance du patient peut le conduire à réaliser des arbitrages qui traduisent plus ses contraintes économiques ou ses préférences, fondées sur des critères subjectifs, que des choix thérapeutiques éclairés. La relation observée entre prix, revenu et quantité consommée peut donc sembler suspecte. Si on observe une diminution de la quantité de soins utilisée quand le prix augmente ou le revenu diminue, l'individu peut en fait renoncer à acquérir des biens ou services médicaux dont il aurait objectivement besoin mais dont il sous-évalue la valeur, par manque d'information (Peele, 1993, Rice, 1993). Enfin, le fait que le prestataire de soins détienne plus d'information que le patient sur sa maladie et sur les moyens qu'il faut mettre en œuvre pour la soigner peut, dans certaines conditions, l'inciter à utiliser le différentiel d'information en sa faveur. Il peut influencer la perception qu'a le patient de ses propres besoins ou de la capacité des soins à les satisfaire (Jones et Zanola, 1995).

¹⁶ *L'information, sous forme de soins prodigués par des experts est ce qui est acheté à la plupart des médecins.*

L'outil théorique qui semble le plus à même de rendre compte de la relation particulière entre prestataire de soins et patient, est la relation d'agence. La combinaison de l'ignorance du patient (principal) et de l'incertitude sur le résultat des soins, peut conduire le médecin (agent) à fournir une quantité ou une qualité de soins qui n'est pas optimale du strict point de vue médical.

1. Si la rémunération du médecin dépend de la quantité de services dont le patient bénéficie (paiement à l'acte), il peut être tenté de multiplier les services rendus. C'est l'hypothèse connue d'induction de la demande par l'offre. Il est très difficile de la tester, en particulier parce que les faits stylisés qu'elle pourrait permettre d'expliquer trouvent d'autres explications théoriques, aussi satisfaisantes et que le choix entre elles relève de la profession de foi (Rochaix et Jacobzone, 1997, pour un bilan récent).
2. Si la rémunération du prestataire est au contraire déconnectée de la quantité de soins qu'il fournit (salarial ou capitation), il peut être tenté de minimiser son effort, ce qui peut se traduire par une moindre qualité.

Quoi qu'il en soit, les tests des différentes hypothèses ont conduit à montrer que les prestataires de soins sont sensibles aux incitations économiques et qu'ils sont en position de tirer avantage du différentiel d'information dont ils bénéficient.

L'ensemble des éléments développés ici va dans le sens de l'idée selon laquelle la souveraineté de la décision de consommer des soins est une hypothèse qui doit être maniée avec prudence. La demande de soins doit donc être conçue comme la résultante de l'interaction des comportements de différents agents économiques plus ou moins bien informés, dans un cadre institutionnel particulier. Attribuer l'utilisation observée de ressources médicales à l'expression renouvelée d'un arbitrage rationnel de l'agent économique entre des consommations médicales dont l'impact sur la santé est connu et d'autres consommations, semble, pour le moins, réducteur.

CONCLUSION

Il nous a paru pertinent, avant de nous tourner vers l'application des principes assurantiels à la santé, d'explicitier les sources d'hétérogénéité de l'état de santé entre individus et de proposer un bilan des types de risques auxquels ils sont confrontés. Cette démarche nous a conduit à souligner que si l'état de santé est au cœur d'un processus de production, les soins ne sont qu'un des inputs de la fonction de production. De plus, il est apparu que les déterminants de l'état de santé sont multiples, exogènes ou endogènes, et que leurs interactions sont complexes. Enfin, l'analyse intertemporelle permet de rendre compte de l'interaction entre les facteurs de risque et l'apparition de dommages à un instant donné. Elle permet aussi de souligner que les dommages peuvent s'inscrire dans le temps. L'ensemble de ces éléments aura un impact sur la nature du service rendu par l'assurance.

Nous nous sommes ensuite penchés plus précisément sur les relations entre soins et santé. En dehors de toute référence à la notion d'assurance, nous avons étudié les caractéristiques du marché des soins et esquissé une présentation des déterminants de l'utilisation des soins et des imperfections de marché qui entachent le cadre dans lequel l'assureur intervient. Nous nous tournons maintenant vers l'étude du service qu'il peut rendre.

CHAPITRE 3

DECLINAISON A LA SANTE DES PRINCIPES ASSURANTIELS

L'assurance santé désigne dans le langage courant un contrat au terme duquel une partie (l'assureur) va rembourser les frais encourus par l'autre (l'assuré) auprès de prestataires de soins, pendant une période donnée, moyennant le paiement d'une prime. Ce recours aux soins est essentiellement motivé par la réalisation de "dommages", autrement dit des dégradations de l'état de santé.

La troisième étape de cette première partie consiste naturellement à décliner les principes assurantiels exposés dans le premier chapitre, compte tenu des caractéristiques des biens étudiés dans le second. Le service rendu par les assureurs dans la couverture des risques santé est en effet largement conditionné par la nature du risque considéré et par le fonctionnement du marché des soins.

Nous confronterons systématiquement la théorie et la pratique afin de souligner comment, concrètement, les assureurs mettent en œuvre le service de couverture du risque santé. Cette démarche nous conduira notamment à étudier dans quelle mesure il y a adéquation entre ce service et les attentes des assurés.

Nous nous penchons en premier lieu sur des difficultés que la théorie permet d'anticiper et qui se posent concrètement aux assureurs pour définir le champ de l'assurance santé. Nous soulignerons les raisons pour lesquelles le risque couvert n'est pas tant le fait de subir une dégradation de l'état de santé que celui de devoir financer l'achat de biens et services médicaux (Section 1). Etant donné que l'assureur finance la consommation de soins des personnes couvertes, nous verrons comment il est amené à gérer les comportements stratégiques des assurés et des acteurs du marché des soins (Section 2). Enfin, nous verrons en quoi la nature de l'activité d'assurance et le fait que le service soit tarifé au risque restreint le rôle de l'assurance dans l'ensemble des risques que l'individu peut souhaiter couvrir en santé (Section 3).

1. OBJET DE L'ASSURANCE SANTE

On identifie traditionnellement l'assurance santé (ou l'assurance maladie) à un service public ou privé, acheté par l'individu ou fourni par l'Etat, qui finance totalement ou en partie les consommations de soins. La théorie économique explicite les raisons qui font que l'assurance santé couvre les soins (1.1); elle permet aussi d'identifier les limites ou les simplifications qu'il implique (1.2).

1.1 L'assurance santé porte sur les soins

Le fait que l'assurance santé rembourse les soins peut a priori sembler surprenant : alors que l'individu cherche fondamentalement à se couvrir contre les conséquences d'une dégradation aléatoire de son état de santé, le service que lui rend l'assureur est le remboursement des soins. Alors que la santé n'est pas un bien marchand (ne dit-on pas souvent qu'elle n'a pas de prix), la valeur monétaire que l'assuré doit recevoir au titre de la réparation d'un dommage sur sa santé semble consensuelle. En outre, contrairement à l'assurance "en cas de décès"¹, qui couvre un dommage constitutif d'une atteinte à l'intégrité de la personne, comme, d'une certaine façon, la maladie, l'assurance santé n'appartient pas au champ de l'assurance vie mais à celui de l'assurance dommage. En fait, l'assuré ne choisit pas, comme en assurance vie, le montant qui lui sera versé s'il subit un dommage contractuellement décrit comme devant donner lieu à indemnisation.

Le raisonnement économique sous-jacent à l'application de la théorie de l'assurance à la santé est en réalité le suivant :

- l'état de santé des individus est soumis à des chocs aléatoires qui constituent des dommages pour les individus ;
- l'utilisation de biens et services médicaux permet de réparer ces dommages.

Dès lors, un individu qui éprouve de l'aversion pour le risque se prémunit contre les conséquences de la survenue de sinistres en santé par la souscription d'un contrat d'assurance au terme duquel l'assureur lui rembourse les dépenses engendrées pour

¹ Qui garantit à des personnes, désignées par le souscripteur, le versement d'un capital ou d'une rente s'il décède.

"réparer" ces dommages. Pour l'économiste, la fonction de production que nous avons évoquée précédemment est, une fois de plus, au cœur du raisonnement. On considère que celle-ci est connue et qu'elle a pour facteur unique les soins. Le modèle de Phelps (1973), dont les principales hypothèses ont été présentées dans le chapitre précédent, est un des premiers à intégrer explicitement une telle fonction de production de santé dans un modèle de demande d'assurance. Les individus maximisent leur fonction d'utilité $U = U(x, H_0 - \lambda + g(h))$ où x représente la consommation courante et h les soins. A chaque dégradation aléatoire de l'état de santé λ on associe, en inversant la fonction de production de santé g , une quantité de soins nécessaire pour rétablir l'état de santé initial. Cette construction théorique reflète l'idée qu'il est possible de définir objectivement le besoin :

"The concept of need in its positive sense describes an external expert's judgment about the marginal impact of a particular form of care on a particular patient's health status"² (Evans, 1983, pp. 78-79).

La définition proposée par Evans souligne que la reconnaissance du besoin, voire parfois même celle de la maladie, dépendent fortement de l'état de la science et de la technologie médicale, puisque les membres du corps médical (les experts) détiennent l'information qui permet de caractériser une altération de l'état de santé et d'évaluer les moyens qu'il faut mobiliser pour y remédier. Plus précisément, à un moment donné, il est possible d'obtenir un consensus suffisant sur les procédures qu'il faut mettre en œuvre en moyenne pour répondre à un problème de santé précis. Si l'on connaît le prix des biens et services médicaux, la maladie passe du statut de choc sur une variable non marchande à un "dommage monétaire". Si l'on connaît la distribution de probabilité d'apparition de différentes maladies, on peut donc estimer le prix de l'assurance.

Cette application de l'assurance à la santé assimile toute maladie à un choc aléatoire sur un état de santé défini de façon exogène. On ne sait pas mesurer l'équivalent monétaire d'une dégradation de l'état de santé, mais on considère qu'il est possible d'en évaluer le coût de réparation. Les résultats de la théorie de l'assurance que nous avons présentés

² Le concept de besoin, dans son acception positive décrit le jugement qu'un expert porte sur l'impact marginal d'un certain type de soin sur l'état de santé d'un patient particulier.

précédemment peuvent donc, dans la mesure où l'on fait l'hypothèse que la fonction de production est connue et déterministe, être appliqués au champ de la santé.

1.2 Le caractère limitatif de ce principe

Même si elle fait l'objet d'un large consensus, il s'agit là d'une hypothèse simplificatrice. En effet, le risque lié à la santé ne peut être résumé au seul risque financier qu'implique la consommation de soins. La désutilité liée à la maladie peut venir :

- de la douleur (physique ou psychologique) ;
- du fait qu'une perte de santé n'est pas toujours réversible ;
- mais aussi du fait que la consommation et la production des autres biens dépendent de l'état de santé (*state dependent utility*).

"In general the money equivalent loss (...) of an illness will exceed any consequent [change in] health spending by some amount which allows for pain and suffering, anxiety, lost wage and / or leisure, and a risk premium for uncertainty of outcome³" (Evans 1984, pp. 30-31).

Rares sont les analyses théoriques qui prennent en compte l'impact global de la maladie sur l'utilité de l'individu et qui déterminent sur cette base sa demande d'assurance. Zweifel (1997) étudie le cas particulier dans lequel un individu valorise plus ou moins son revenu disponible (autrement dit sa consommation non médicale) selon son état de santé. Dans son modèle, la fonction d'utilité de l'individu est différente selon qu'il est malade ou non.

Deux situations peuvent se présenter :

- l'individu profite plus de sa consommation lorsqu'il n'est pas malade (loisirs par exemple). Il choisit alors une couverture incomplète (franchise), y compris dans le cas où le prix de l'assurance est actuariel ;

³ En général, l'équivalent monétaire d'une maladie dépasse la variation des dépenses de santé qu'elle implique d'un montant qui prend en compte la douleur et la souffrance, l'anxiété, la perte de salaire et/ou de loisirs et une prime de risque qui compense l'incertitude liée au résultat du traitement.

- le cas inverse illustre une situation dans laquelle l'individu, lorsqu'il est malade, valorise particulièrement des consommations non strictement médicales (services annexes d'aide à domicile par exemple). Il cherchera, dans ce cas, à se sur-assurer, autrement dit, à avoir un revenu disponible plus élevé dans le cas où il tombe malade.

Concrètement, certains impacts induits par la dégradation de l'état de santé peuvent être couverts par des contrats d'assurance mais ceux-ci sont dissociés de l'assurance santé proprement dite :

1. des contrats d'assurance (ou des assurances sociales) permettent aux individus de se prémunir contre la perte de salaire en cas d'incapacité de travail due à une dégradation transitoire ou permanente de l'état de santé ;
2. des garanties dites "des affections redoutées" (Bague-Forst, 1995) garantissent le versement d'un capital à un assuré si une affection grave prévue au contrat est diagnostiquée. Un tel contrat pourrait être assimilé à une compensation forfaitaire de la désutilité liée à la maladie. Il s'apparente, dans les faits, davantage à un contrat d'assurance "en cas de décès" et ce type d'assurance est en général souscrit dans le cadre de l'assurance vie ;

Au delà des justifications théoriques évoquées, on peut trouver des raisons concrètes pour lesquelles l'assureur ne conditionne pas ses obligations contractuelles à l'état de santé de l'individu.

En premier lieu, celui-ci est difficile à mesurer de façon objective et incontestable. Le versement d'une indemnité compensatoire forfaitaire, liée à une variation de l'état de santé, soulèverait des difficultés pour :

- d'une part, établir un contrat en termes juridiques clairs ;
- d'autre part, déterminer si l'état de santé de l'individu s'est dégradé dans des proportions justifiant qu'il reçoive les indemnités prévues au contrat.

Ensuite, si on reste dans l'hypothèse selon laquelle l'assureur rembourse des soins, il pourrait théoriquement proposer une couverture contre certaines maladies. Mais le fait de définir le champ contractuel par rapport à des pathologies soulèverait des difficultés supplémentaires :

- de vérification de l'occurrence du dommage. L'individu peut être tenté de déclarer une maladie couverte pour bénéficier de soins dont il estime avoir besoin pour d'autres raisons. L'assureur devrait donc, au minimum, s'il basait l'assurance sur l'occurrence d'une maladie, mandater un expert pour établir un diagnostic avant de couvrir les soins. Cela reste une procédure complexe, coûteuse, et probablement injuste car le fait de souffrir ou non d'une maladie peut s'avérer difficile à établir scientifiquement ;
- de vérification de la causalité entre l'utilisation de soins et la présence d'une maladie. La consultation d'un médecin qui serait couverte par l'assurance au titre d'une maladie, peut donner lieu à une prestation de services liée à d'autres maladies sans que l'assureur en soit informé. Il serait, par ailleurs, délicat en cas de pluri-pathologie de déterminer les services utilisés au titre de chaque maladie et quand bien même on parviendrait à le faire, on multiplierait à la fois les coûts que supporterait l'individu et très probablement les risques iatrogènes.

On peut noter à titre de contre-exemple que la Sécurité sociale en France prévoit des exonérations de ticket modérateur pour certaines affections et pour les soins relatifs à ces affections. Cependant, il ne s'agit pas d'élaborer des contrats d'assurance autour de ce seul principe. L'enjeu financier étant relativement limité, les contrôles n'ont pas besoin d'être très stricts quant au motif pour lequel la personne consulte un prestataire, dans la mesure où ce dernier est susceptible de traiter la pathologie qui motive l'exonération. Ces exonérations ont cependant été vivement critiquées au motif qu'elle pouvaient faciliter une collusion tacite des médecins et de leur patients au détriment du tiers payeur.

De fait, l'assurance portant sur des biens et services médicaux, son champ est en général limité par référence à des types de soins, comme par exemple l'hospitalisation. Il est vrai qu'exclure certains types de prestations revient à ne pas couvrir certains risques (ce serait par exemple le cas si on excluait l'ensemble des soins prodigués par des dentistes) mais la convergence est beaucoup plus difficile à établir pour la majorité des soins. En

définitive, la principale raison pour laquelle "l'assurance santé" porte sur les soins est que des problèmes d'information empêchent que les dommages puissent être définis en référence à l'état de santé (Blomqvist, 1997).

2. PRISE EN COMPTE DES COMPORTEMENTS STRATEGIQUES PAR L'ASSUREUR

L'assureur, dans la mesure où il finance les soins dont il est difficile de contrôler la nécessité, la pertinence et la qualité, est amené à prendre en compte et à juguler les comportements stratégiques de ces différents acteurs.

Le risque moral, qui empêche l'allocation optimale des ressources, est l'objet de nombreuses attentions en santé, car il touche l'assurance privée comme l'assurance publique dans un contexte où la maîtrise des dépenses est un enjeu majeur. Les assureurs sont-ils en mesure de gérer le risque moral (2.1) ? Plus encore, ont-ils les moyens de contrôler les imperfections du marché des soins sur lequel ils n'interviennent qu'en qualité de tiers-payeur (2.2) ? Cette seconde question renvoie notamment à l'idée que les patients assurés sont en position d'infériorité sur le marché des soins et qu'ils ne sont pas toujours en mesure de déterminer la qualité, la nécessité et l'adéquation des traitements qui leur sont proposés.

2.1 Le risque moral

Si l'on applique directement aux soins la notion de risque moral telle qu'elle est présentée dans la théorie de l'assurance, on doit tout d'abord faire l'hypothèse que les risques liés à la santé sont partiellement endogènes. Les individus, en faisant un effort plus ou moins grand, peuvent réduire, soit la probabilité de réalisation du dommage, soit son montant. Cet effort est coûteux pour l'assuré et n'est pas observable par l'assureur. Celui-ci doit donc mettre en place des mécanismes d'incitation pour que l'assuré, qui se sait protégé en cas de dommage par son assurance, continue néanmoins à faire des efforts pour diminuer la valeur attendue de son dommage.

2.1.1 Assurance et prévention : le risque moral ex ante

L'application directe du concept de risque moral à la santé se traduit en premier lieu par l'idée que la personne couverte par une assurance va adopter des comportements plus risqués et/ou va affecter moins de ressources à la prévention. Zweifel et Breyer (1997) proposent un modèle dans lequel l'individu peut influencer la probabilité de dommage en faisant un effort, dont l'impact sur la probabilité est positif mais décroissant. Cet effort peut être assimilé à de la prévention primaire. Les deux auteurs prolongent les résultats d'Ehrlich et Becker. Ils comparent une situation où l'effort de l'assuré est observable par l'assureur et où il ne l'est pas. Le modèle prend en compte l'existence possible d'un taux de chargement proportionnel à la prime.

1. Si l'effort est observable par l'assureur, la prime est fondée sur la probabilité de dommage en fonction du niveau d'effort choisi par l'individu. On retrouve les résultats présentés dans le premier chapitre : assurance complète si le taux de chargement est nul, franchise s'il est positif jusqu'à une certaine limite au delà de laquelle l'assurance n'est plus demandée parce que son prix est prohibitif. Le "risque moral" donc est neutralisé.
2. Quand l'effort n'est pas observable par l'assureur, l'assuré ne peut bénéficier de son effort de prévention par le biais d'une diminution du prix de l'assurance. La prime d'assurance est donc liée au niveau d'indemnisation en cas de dommage et non à son effort. Lorsque le taux de chargement nul, le modèle montre que :
 - du côté de l'assuré, si l'assurance est complète, la prévention est nulle, si l'assurance décroît, la prévention tend à augmenter ;
 - comme l'assureur connaît la fonction de réaction de l'assuré, il va proposer un prix unitaire d'assurance plus élevé quand la couverture augmente (la prime augmente plus que proportionnellement à l'indemnité).

L'utilité de l'individu est finalement plus élevée dans le cas où l'effort est observable : lorsque les frais de gestion sont nuls, il obtient l'assurance complète, ce qui n'est pas le cas quand l'effort n'est pas observable.

De "l'auto-protection" que nous avons évoquée dans le premier chapitre, on passe au risque moral sur la probabilité de dommage, en ajoutant une hypothèse d'asymétrie

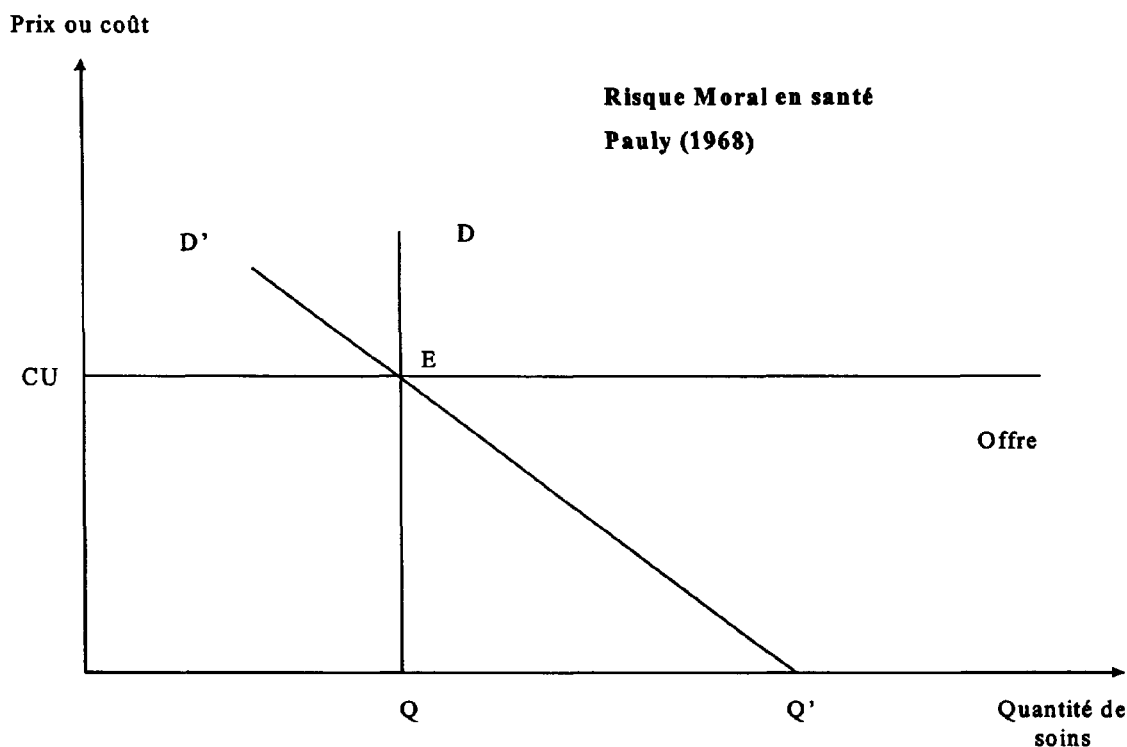
d'information. On pourrait de même passer de "l'auto-assurance" à une forme de risque moral qui conduirait l'assuré, sachant que ses dépenses seront remboursées s'il tombe malade, à avoir moins recours à la prévention secondaire⁴. A notre connaissance, le lien entre assurance et prévention secondaire n'est pas étudié dans la théorie de l'assurance appliquée à la santé.

Il existe, de plus, peu d'études empiriques sur le sujet. Faire un effort d'auto-assurance consiste, de fait, à utiliser des ressources médicales. Il est donc difficile de distinguer ce qui relève réellement de la prévention secondaire quand on observe la consommation médicale d'un individu. De plus, la prévention secondaire, qui est la conséquence (en simplifiant) d'un suivi médical régulier, dépend probablement plus des facteurs qui influencent la consommation (prix, ressources, conditions d'accès et niveau d'éducation) que du fait d'être assuré en tant que tel. L'intuition semble plutôt suggérer que la prévention secondaire, si tant est que la personne en reconnaît la nécessité, serait favorisée par l'assurance, soit parce qu'elle est encouragée par l'assureur (qui en tire un bénéfice indirect), soit parce que l'individu qui en a les moyens (peut-être justement parce qu'il est assuré) y a recours. Ces intuitions sont confirmées sur des données françaises : Genier et Jacobzone (1998) montrent en effet que la prévention tendrait plutôt à augmenter lorsque les personnes sont assurées.

2.1.2 Assurance et effet prix : le risque moral "ex post"

La notion de risque moral est plus généralement utilisée en santé pour désigner l'idée selon laquelle pour un état de santé donné, une personne assurée va consommer plus de soins qu'une personne non assurée. L'assurance rembourse à l'individu les soins qu'il utilise. Elle a donc pour effet de diminuer, voire de rendre nul le coût marginal pour l'assuré de la consommation des soins. Si l'élasticité-prix de sa demande, pour un état de santé donné, n'est pas nulle, l'assurance conduit l'individu à demander plus de soins. Le raisonnement est illustré par la figure ci-après (Pauly, 1968).

⁴ Dépistage systématique, suivi régulier destiné à une détection précoce des maladies.



CU est le coût unitaire de production des soins. Les rendements sont constants, hypothèse classique pour la production de services. Que la fonction de demande soit D ou D' , l'équilibre du marché est E et la quantité échangée est Q . L'assurance a pour conséquence d'amener le prix des soins pour l'assuré à zéro (dans le cas où elle le rembourse intégralement). Si la demande n'est pas élastique, la quantité échangée en présence d'assurance reste Q . En revanche, si la demande est élastique, l'assuré va consommer Q' lorsqu'il est assuré. Le comportement de consommation est donc modifié par la présence d'assurance. Il s'agit d'un arbitrage rationnel. Ce risque moral est constitutif d'une inefficacité à plusieurs titres :

- d'une part, l'équilibre du marché ne possède plus les propriétés souhaitables. Sans assurance, l'équilibre (prix et quantité) est déterminé par l'intersection de l'offre et la demande. Par construction, le prix est égal au coût marginal des soins. Dans le cas où la baisse de prix conduit l'individu à demander Q' , le véritable coût marginal CU est supérieur au coût marginal pour l'individu (et donc à son utilité marginale), l'utilisation des ressources médicales est donc inefficace (Feldstein, 1993) ;
- d'autre part, l'existence de risque moral relativise, pour l'individu, l'intérêt de l'assurance. Si l'assureur fonde le calcul de sa prime sur la quantité qui sera effectivement consommée, le coût de l'assurance est beaucoup plus élevé qu'en l'absence de risque moral. L'assuré peut donc préférer la perspective de devoir payer

Q unités s'il tombe malade à celle de s'acquitter dans tous les états du monde d'une prime basée sur Q' .

L'assureur, qui est de fait tiers-payeur, pour contrer cet effet "naturel" de l'assurance sur la consommation, doit connaître la fonction de réaction de l'assuré à une baisse du prix des soins. En fonction de l'élasticité, il faut alors déterminer un taux de coassurance, c'est à dire un prix P^* inférieur à CU qui sera payé par l'assuré pour chaque unité consommée en cas de dommage. A ce P^* sont associés une consommation $Q^*(P^*)$ et une prime d'assurance qui prend en compte ces paramètres. Cette prime est telle que l'agent est indifférent (ou préfère faiblement) l'assurance à la perspective aléatoire. Il y a donc arbitrage entre la réduction du risque financier auquel le patient est exposé et la perte de bien-être due aux effets du risque moral. Le taux de coassurance optimal dépend des caractéristiques des demandes de soins et d'assurance (Marquis et Phelps, 1987; Feldman et Manning, 1995; Zweifel et Breyer, 1997⁵).

On peut noter que dans cette analyse, les hypothèses qui concourraient à la définition d'une situation de risque moral en assurance et qui sont, d'une part, la capacité qu'a l'individu de modifier son niveau de risque en entreprenant un effort coûteux et, d'autre part, l'asymétrie d'information sur l'effort effectivement réalisé, ne sont plus explicitement au cœur de l'analyse. Ici, pour un dommage donné, on insiste sur le fait que les ressources et les préférences des malades vont les inciter à choisir un niveau d'utilisation différent et que l'assurance, en diminuant le prix des services, introduit dans ces choix une distorsion complexe conduisant en général les personnes à avoir un montant plus élevé de consommation que si elles n'étaient pas assurées.

Dans la réalité, le remboursement proposé par l'assureur peut porter sur l'intégralité des frais engagés, ou imposer une participation financière de l'utilisateur au financement du service.

1. Cette participation financière peut être déterminée en référence à l'unité consommée.

Les possibilités sont multiples, l'assureur peut :

- prendre en charge une proportion fixe du coût réel (coassurance) ;

⁵ Ces derniers auteurs étendent l'analyse au cas où l'utilité de l'individu dépend du fait qu'il est malade ou non.

- rembourser un montant fixe par type de soins (remboursement de l'assuré par référence à un barème) ;
 - imposer une participation fixe de l'assuré par type de service ou de bien (copaiement).
2. L'assureur peut aussi fixer des limites à ses versements en référence à la dépense totale (ou à la dépense par type de soins) :
- dans le cas d'une franchise, l'assuré n'est remboursé qu'à partir du moment où ses dépenses ont dépassé un seuil ;
 - dans le cas d'un plafond, l'assureur limite sa responsabilité totale en imposant un niveau au delà duquel l'assuré devra assumer les frais.

Parmi les études empiriques qui cherchent à mesurer le risque moral ou plus précisément l'élasticité-prix des soins, celle de la RAND (Manning et Marquis, 1989, Manning, *et al.*, 1987) reste la plus connue et fait référence. Ses résultats indiquent, comme le soulignent par exemple Levy *et al.* (1993) que la demande est effectivement sensible aux prix, à condition que le taux de coassurance soit assez élevé (de l'ordre de 25%).

La participation financière des assurés permet donc aux assureurs de limiter l'impact du risque moral ex post. Pour autant, il ne s'agit que d'un des types d'outils que l'assureur peut mettre en œuvre pour limiter l'impact des imperfections du marché des soins sur ses dépenses dans le cadre des contrats qu'il propose.

2.2 Vers une approche élargie

Certaines particularités du comportement des acteurs présents sur le marché montrent les limites de mesures strictement financières destinées à restreindre la consommation des personnes assurées. Leur analyse permet, en outre, d'expliquer pourquoi l'assureur ne se contente pas nécessairement de rembourser les soins de façon aveugle et peut choisir d'adopter un comportement plus actif dans la gestion des risques.

2.2.1 Les transactions sur le marché des soins

Tout d'abord, on peut rappeler les doutes qui pèsent sur la capacité de l'individu à déterminer les biens et services qu'il devrait consommer pour répondre à un problème de santé. Ceci conduit logiquement à s'interroger sur sa capacité à choisir avec pertinence les biens auxquels il devrait renoncer lorsque sa participation financière augmente. Sur la base des données de la RAND, Lohr *et al.* (1986) concluent, au terme d'une étude empirique, que "*cost sharing is just as likely to lower use when care is thought to be highly effective as when it is thought to be only rarely effective*"⁶. Ainsi, quand bien même l'introduction de copaiements permettrait de lutter contre le risque moral, elle conduit l'individu à renoncer à des soins selon une logique qui lui est propre et qui n'est pas nécessairement fondée sur une décision informée. L'assureur peut être amené à supporter les conséquences de ces mauvais choix et il peut avoir un intérêt direct à informer la patient sur l'efficacité des services.

Par ailleurs, dans un univers où le service n'est guère standardisé et où les prix risquent de l'être encore moins, le consommateur assuré n'est pas incité à exercer de pression concurrentielle sur les producteurs. Comme le souligne Arrow : "*insurance removes the incentives on the part of individuals, patients ... to shop around for better prices*" (1963)⁷. L'assuré peut être d'autant plus tenté de relâcher sa vigilance qu'il associe prix élevé et meilleure qualité du service rendu (à tort ou à raison puisqu'il n'a pas vraiment les moyens de juger).

En fait, l'assureur, le patient, et le prestataire de soins (en particulier le médecin) entrent dans une relation complexe.

- Nous avons vu que le médecin est théoriquement l'agent du patient, et qu'il doit déterminer les soins qui amélioreront son état de santé.
- De la même façon, le prestataire de soins est l'agent de l'assureur : "*by certifying to the necessity of given treatment or the lack thereof, the physician acts as a controlling agent on behalf of the insurance companies. Needless to say, it is far*

⁶ Les participations financières des usagers sont tout aussi susceptibles de diminuer l'utilisation des services lorsque ceux-ci sont jugés très efficaces que lorsqu'ils ne le sont que rarement.

⁷ L'assurance élimine l'incitation des individus et des patients.. à chercher le service au meilleur prix.

*from perfect check; the physicians themselves are not under any control and it may be convenient for them or pleasing to their patients to prescribe more expensive medications, private nurses, more frequent treatments, and other marginal variations of care*⁸ (Arrow, 1963, pp. 962).

Arrow souligne ici l'ambiguïté du rôle du prestataire de soins, agent de l'assureur comme de l'assuré, et dont les intérêts sont contradictoires dans la mesure où le premier finance la consommation du second. Le médecin est en outre lui-même un agent économique rationnel qui prend en compte ses propres incitations. Si son intérêt converge avec celui du patient, il peut le laisser, voire l'inciter à consommer plus que ce qui est nécessaire du strict point de vue médical (voir aussi Rochaix, 1997). Le patient n'ayant au mieux qu'une incitation financière marginale à limiter la dépense totale n'impose pas de contrôle très strict sur le prestataire. A titre d'illustration, une étude empirique menée à SANESCO⁹ sur des informations émanant de fichiers d'assurance pour la garantie dentaire avait permis de montrer que le prix moyen de la couronne dentaire augmentait significativement quand le taux de couverture dont bénéficiait l'assuré s'accroissait. Il passait de 1800 F pour une garantie remboursant le ticket modérateur, à plus de 3300F, pour les garanties dépassant 150% du tarif de convention. Or, en France, le dentiste est libre de fixer le prix de la couronne. Certes, il existe différents types de couronnes et des considérations esthétiques peuvent motiver le choix de tel ou tel matériau (métal ou céramique). Les différences de prix peuvent aussi être expliquées par la difficulté du travail. Pour autant, il y a de fortes chances pour que ces besoins spécifiques soient équi-répartis dans la population. La corrélation (que l'on observe dans différentes régions) reflète donc en partie une convergence d'intérêt entre patient et producteur aux dépens de l'assureur. En effet, malgré un prix plus élevé quand le taux de couverture augmente, le taux de remboursement global des frais réels augmente avec le niveau de garantie. Autrement dit, le patient ne pâtit généralement pas

⁸ *En certifiant de la nécessité d'un traitement ou d'une abstention thérapeutique, le médecin agit en qualité d'agent contrôleur pour le compte des compagnies d'assurance. Point n'est besoin de dire que ce contrôle est loin d'être parfait ; les médecins eux-mêmes ne sont pas contrôlés et il peut s'avérer commode pour eux et agréable pour les patients de prescrire des médicaments plus coûteux, des infirmières privées, des traitements plus fréquents, ou d'autres variations marginales de la prise en charge thérapeutique.*

⁹ Bousquet F., Couffignal A., Meyer C. (1996) *La prise en charge du risque dentaire en assurance de groupe et en assurance individuelle.* Sanesco/FFSA/Gras-Savoie.

de l'augmentation des prix, qui augmente le revenu du dentiste et le coût supporté par l'assurance.

La relation d'agence offre un cadre théorique adéquat pour décrire les relations qu'entretiennent les acteurs en présence, mais surtout pour déterminer les incitations optimales qui doivent régir leurs relations. Elle conduit aussi à reconnaître que l'assureur, plus que le patient, paie le service rendu au patient par le producteur. Il peut alors légitimement intervenir directement dans le processus de production en choisissant, par exemple, le mode de rémunération du producteur en fonction de ses propres priorités.

2.2.2 La gestion directe des transactions par l'assureur

Dans l'analyse proposée ici, l'assureur n'est plus seulement tiers-payeur mais il devient acheteur de soins pour le compte de l'assuré, voire gestionnaire du risque. Il modifie, au moins en partie, la relation entre le patient - consommateur et le professionnel de santé. Quelle est la logique sous-jacente à une telle évolution du rôle de l'assureur ?

En règle générale, le patient s'en remet à un professionnel qui détermine pour lui l'essentiel du traitement. Son choix se limite le plus souvent à celui d'un professionnel particulier. Le patient s'oriente seul dans la filière de soins, et même en faisant l'hypothèse qu'il choisit en parfaite connaissance de cause les différents prestataires, sur des critères de proximité et de qualité, tant médicale que relationnelle, il n'a accès qu'à une série d'actes médicaux et jamais à ce qu'il pourrait souhaiter acquérir : à savoir le traitement global de son problème de santé.

La complexité des prises en charge, les cloisonnements et les conflits d'intérêt qui peuvent exister entre les différents prestataires, rendent nécessaire la gestion du risque pour que le patient soit bien soigné. On entend ici la gestion du risque au sens large, à la fois gestion et coordination du processus de production de soins, et gestion du capital de santé des usagers. L'idée générale est que la production de soins est un processus qui gagne en efficacité s'il gagne en cohérence (Robinson, 1996). Une meilleure organisation des soins peut donc passer par l'émergence d'une fonction que nous

appelons "d'entrepreneur de soins"¹⁰. L'entrepreneur ne se contente en effet pas de garantir la qualité et de coordonner les soins. Il peut déployer une batterie d'outils, au premier rang desquels figurent les incitations financières en direction des producteurs de soins. L'entrepreneur de soins se comporte comme un industriel classique : il négocie les prix des facteurs de production, cherche à minimiser les rentes éventuelles de producteurs, organise le processus de soins pour l'optimiser en fonction des évolutions technologiques et surveille la qualité du service. L'information étant pour le patient assuré très difficile à obtenir sur le marché des soins, l'entrepreneur cherche à simuler une situation équivalente à celle d'un marché dans lequel un patient, parfaitement informé, négocierait les prix des actes avec les producteurs. La responsabilité de cette simulation lui incombe, car les moyens dont il dispose lui permettent de déterminer l'agencement le plus efficace des actes élémentaires entrant dans le traitement, et d'exercer une pression concurrentielle sur les prix et la qualité des soins.

Dans cette vision, l'entrepreneur de soins est un agent - parfait - du patient (van de Ven et al., 1994a). La question que l'on doit se poser alors est celle de l'institution qui va incarner cette fonction. L'assureur, qui par construction est amené à payer les soins que consomment un grand nombre d'assurés, est un candidat possible¹¹. L'émergence du *managed care* s'est d'ailleurs faite aux Etats-Unis à l'initiative de plans de santé (*health plans*). Ils l'ont utilisé pour offrir aux entreprises des contrats plus concurrentiels que ceux des assureurs traditionnels. Ces plans fournissaient un ensemble de soins aux membres affiliés en contrepartie d'une prime. Depuis la période des premiers *Health Maintenance Organizations* (HMO), dont l'activité se concentrait sur la prise en charge des salariés d'une entreprise¹², les organisations de *managed care* (MCO) ont pris une part croissante dans le marché de la santé, sous l'effet de politiques favorables au niveau fédéral¹³ et d'une évolution des comportements des employeurs et des consommateurs

¹⁰ En référence à un travail précédemment mené (Bocognano, et al., 1998). Une partie de ce travail est aussi basée sur Bocognano, et al. (1999) "Dérégulation de l'assurance maladie : quelques enseignements des expériences étrangères", soumis à publication auprès d'*Etudes et Recherches en Economie Publique*

¹¹ Mais non unique (Bocognano et al. 1998).

¹² Sur le modèle désormais largement connu du *Kaiser Foundation Health Plan*.

¹³ Dès 1971 obligation est faite aux employeurs de proposer un MCO parmi les plans d'assurance maladie s'il en existe dans la région.

désireux de limiter l'inflation des primes. Le *managed care* est désormais un produit classique aux Etats-Unis sur le marché de l'assurance maladie.

Son développement s'est accompagné d'une extraordinaire diversification des organisations qui le mettent en œuvre. Le facteur commun à toutes les MCO, et qui les distingue des systèmes d'assurance traditionnelle avec paiement à l'acte, n'est pas tant la notion de gestion des soins que la notion de sélection d'un réseau de médecins. En effet, même des plans d'assurance traditionnels se sont mis à faire de la gestion du risque, à instaurer des procédures d'entente préalable pour les soins coûteux, à exercer des contrôles sur l'utilisation de ressources onéreuses, etc. La contractualisation sélective avec les producteurs de soins est en revanche le trait distinctif des MCO ; mais à partir de cet élément commun, les organisations restent variables (Miller et Luft, 1994).

Les plus anciennes organisations sont les HMO, qui ont des relations exclusives ou non avec les prestataires et qui peuvent les rémunérer sur une base de capitation ou les intégrer dans une structure hiérarchique (salarial ou propriété des structures de soins). Dans les HMO, les assurés doivent obligatoirement s'adresser aux prestataires du réseau. Ce cadre rigide et fermé a entraîné la création de formules concurrentes intermédiaires entre les HMO et l'assurance traditionnelle. Les PPO (*Preferred Providers Organizations*) ont ainsi vu le jour dans les années quatre-vingt. Dans ces réseaux, des arrangements préférentiels sont négociés avec des groupes de médecins et des hôpitaux : les médecins restent payés à l'acte mais facturent des tarifs plus bas et acceptent des procédures de contrôle des pratiques. Le patient, lui, est mieux couvert s'il utilise les praticiens du réseau, mais garde la liberté de consulter à l'extérieur, en contrepartie d'une participation financière plus importante.

La croissance du *managed care* s'est effectuée essentiellement à partir des modèles les plus ouverts ou hybrides. Si les MCO ont en commun quelques principes d'organisation de la filière du patient (par exemple le rôle du généraliste - comme filtre – *gatekeeper* - vis-à-vis de l'accès aux soins spécialisés et lourds), les outils de contrôle des pratiques médicales, les méthodes de rémunération des médecins, les systèmes de partage du risque financier sont très variables. La tendance actuelle est au transfert du risque financier vers les groupes de médecins : le plan verse au groupe une fraction de la prime qu'il reçoit par patient, et le groupe de médecins assume le risque d'un dépassement du budget.

Les *Provider Sponsored Organizations*, les PSO constituent une réponse des prestataires : des organisations de producteurs proposent directement aux entreprises un service de prise en charge des patients moyennant capitation, sans passer par un tiers (le *health plan*), et donc en assurant directement les fonctions habituellement prises en charge par cet intermédiaire.

L'expérience américaine semble montrer que le *managed care* permet de résorber des inefficacités à un instant donné (notamment en réduisant la durée de séjour et la fréquence des hospitalisations) même si son développement n'a pas, aux Etats-Unis, réussi à freiner durablement l'inflation des dépenses de santé.

Au travers de cet exemple, nous avons brièvement montré que l'assureur peut se doter de moyens multiples pour gérer les risques et non se contenter de rembourser les dépenses encourues par les personnes qu'il couvre. Dans un univers concurrentiel, le développement de cette activité lui permet théoriquement de proposer des primes plus faibles et contribue à améliorer l'efficacité de l'allocation des ressources médicales et de la "production" des prises en charges.

Le modèle de *managed care* mis en œuvre par les assureurs fait figure de référence dans de nombreux pays et inspire des réformes qui visent à faire émerger cette fonction d'entrepreneur de soins. Tous s'interrogent cependant sur les incitations qu'il faut mettre en place pour que les opérateurs désignés agissent dans l'intérêt des patients. Les débats autour de la *managed competition* (Enthoven, 1994, par exemple), soulignent notamment la nécessité de contrôler l'activité des assureurs qui, pour faire pression sur les coûts, peuvent être tentés de limiter l'accès aux soins ou la qualité des prises en charge.

3. TARIFICATION DE L'ASSURANCE

Le champ de l'assurance et la nature du service étant définis, l'assureur doit y associer un prix, en l'occurrence, estimer ex ante le niveau de risque qu'il couvre en fonction de ses engagements contractuels. Expliciter les principes de la tarification actuarielle dans le cas particulier de la santé (3.1) permet de caractériser plus précisément la nature du service qui peut être rendu par l'assurance, en adoptant le point de vue des personnes qui cherchent à se couvrir (3.2).

3.1 Tarification actuarielle du risque en santé

Pour élaborer une prime, l'assureur doit identifier clairement les facteurs de risque, estimer les dommages possibles associés et les dépenses qui seront, en moyenne, encourues par la personne si ces risques se traduisent par des dommages. C'est le principe de la tarification actuarielle, ou actuariellement neutre, fondamentalement associée au service d'assurance comme nous l'avons vu dans le premier chapitre. En théorie, la prime d'assurance doit être égale à la moyenne des dépenses que va engendrer un individu pendant la période contractuelle. Or, la tarification du risque en santé est un exercice complexe (3.1.1). Cette difficulté contribue à expliquer la durée relativement courte des contrats (3.1.2).

3.1.1 Le principe

Nous avons montré que les sources d'hétérogénéité entre l'état de santé des individus sont multiples. La tâche de l'assureur est donc complexe. Partons de l'hypothèse (Newhouse, 1996) que les dépenses d'un individu suivent le modèle :

$$D_{it} = \alpha + X_{it}\beta + \mu_i + \varepsilon_{it}$$

où D_{it} est la dépense de l'individu i pour la période t , X_{it} est un vecteur de facteurs de risque qui caractérise l'individu i à la date t , μ_i est une caractéristique invariante de l'individu, ε_{it} est un terme d'erreur individuel qui reflète la variance pure et qui varie avec le temps mais qui peut éventuellement être auto-corrélé.

La variance des dépenses se décompose en une variance pure et une variance explicable :

1. la variance pure résulte de la réalisation d'événements aléatoires, par exemple le fait d'avoir un accident. Assumer cette variance pure est le rôle de l'assureur ;
2. la variance explicable dépend des caractéristiques individuelles et devrait être entièrement intégrée dans la prime. Elle se décompose à son tour en plusieurs composantes :
 - certaines variables sont objectives et communes à des sous-groupes d'individus (le sexe, l'âge, la région...). Elles sont prises en compte dans le vecteur X , dont les coefficients sont identiques pour tous les individus ;

- les autres sont les caractéristiques intrinsèques d'un individu précis (μ_i), qu'il est impossible de déterminer objectivement, comme, pour prendre un exemple simple, le fait d'être hypocondriaque. Toutes choses égales par ailleurs, une personne hypocondriaque aura tendance à consommer plus de soins qu'une autre. Il est aussi probable qu'une part de ce μ_i reflète des caractéristiques que l'état de la science ne permet pas d'objectiver et d'intégrer au vecteur X^{14} ;
- enfin, on peut prédire une partie de l'évolution inter-temporelle des variables caractérisant les dépenses de santé. L'âge est l'un de ces facteurs, mais on peut aussi pronostiquer pour partie l'évolution de certaines maladies (guérison progressive ou aggravation) sur plusieurs périodes.

La prime idéale d'assurance devrait prendre en compte l'ensemble de ces caractéristiques. L'assureur doit objectiver le plus grand nombre possible de facteurs de risque X , ou plus précisément à trouver ceux qui permettent d'expliquer la plus grande part de la variance. L'assureur associe à chaque sous-groupe défini par une combinaison différente des caractéristiques X , une prime qui est égale à l'espérance de dommage. Dans les faits, la variable individuelle intrinsèque ne pourra que difficilement être prise en compte dans la tarification au risque puisque l'assureur doit, par construction, déterminer des groupes homogènes afin de mutualiser le risque. Si l'assureur parvient à estimer le μ_i de chaque individu, il peut cependant théoriquement envisager de constituer, au sein de chaque cellule de risque définie par la combinaison des critères X , des groupes de personnes qui ont des niveaux de μ_i identiques et leur appliquer un tarif plus ou moins élevé. La faisabilité d'une telle tarification est assez improbable. En revanche, on peut penser que la part de cette variable sera amenée à diminuer au fur et à mesure que la connaissance des déterminants de l'état de santé augmentera.

¹⁴ Si par exemple, on découvre que toutes les personnes qui ont une caractéristique génétique précise ont, toutes choses égales par ailleurs, des dépenses plus élevées, on pourra ajouter une variable indicatrice de la présence du gène dans le vecteur X et le μ_i de toutes les personnes qui présentent cette caractéristique génétique sera plus faible.

3.1.2 La périodicité des contrats

Compte tenu de l'importance de la dimension temporelle dans la détermination de l'état de santé, mais aussi de la nature de l'activité d'assurance, la périodicité du contrat est un élément central de caractérisation du service. L'assureur doit estimer ex ante une prime telle qu'il pourra, pendant la période contractuelle, couvrir en moyenne tout ou partie des dépenses de santé des assurés. Traditionnellement, il n'a pas la possibilité de se retourner contre les assurés (indépendamment des cas de fraude) si les sommes collectées ne lui ont pas permis à échéance du contrat de couvrir ses engagements¹⁵. Par rapport à de nombreux biens et services, il y a donc inversion du cycle de production puisque le service est payé avant d'être rendu. Un tel mécanisme de pré-paiement a pour conséquence de le rendre "*peu capable de faire face à des risques qui n'ont pas été anticipés*" (Blanchet, 1996).

La difficulté d'estimer les risques explique, pour une large part, la durée des contrats d'assurance privée qui ne dépasse quasiment jamais un an. Si le contrat était conclu pour une durée longue, l'assureur pourrait commettre des erreurs de tarification :

- le risque moyen que présente une personne ou un groupe peut évoluer au fil du temps. Estimer non seulement les risques à l'entrée dans le contrat mais leur évolution possible à long terme, relève de la gageure ;
- quand bien même les risques seraient évalués correctement, le coût de prise en charge des dommages peut évoluer, du fait de l'apparition de nouvelles technologies ou de modifications des modalités de traitement. Pour s'engager à long terme, l'assureur devrait non seulement connaître le risque pour une personne de souffrir d'une affection mais aussi anticiper le coût de la prise en charge au moment où le dommage surviendra ;
- enfin, il existe des formes d'aléa trop radicales pour que l'assureur les anticipe et soit en mesure d'en évaluer les coûts. L'émergence de nouveaux risques, comme le SIDA, ne peut être anticipée et, a fortiori, tarifée.

¹⁵ En revanche, il peut lui même souscrire un contrat de réassurance pour le cas où une telle situation se produirait.

A l'inverse, un contrat plus court permet à l'assureur d'utiliser l'information acquise durant la période où il court pour tarifier le contrat à la période suivante. Les contrats que l'assureur va proposer sont donc, en règle générale, circonscrits dans le temps.

3.2 Interprétation et conséquences pour les assurés

Compte tenu des éléments qui précèdent, le champ des risques pesant sur la santé pouvant être pris en charge par une assurance est précisé. Il apparaît alors clairement que le mécanisme d'assurance n'est pas pertinent pour couvrir tous les risques en santé (3.2.1) et que les possibilités de se couvrir à long terme sur le marché privé sont insuffisantes (3.2.2).

3.2.1 La pertinence limitée du mécanisme d'assurance

Le résultat de l'opération d'assurance peut être interprété, pour la personne qui se couvre, comme un transfert de la variance du revenu sur une entité extérieure. L'espérance de richesse de l'individu est identique avec et sans assurance. Sans assurance, la variance de la richesse est positive, avec assurance, elle est nulle. Le but de l'assurance n'est donc pas de prendre en charge les dépenses de santé d'un individu mais de le protéger contre une variation autour de la moyenne de ses dépenses.

Ainsi, la dégradation inéluctable de l'état de santé associée au vieillissement ne peut pas être analysée comme provenant de la réalisation d'une variation aléatoire. Toute personne doit s'apprêter à voir ses dépenses moyennes augmenter en fonction de son âge et doit allouer ses ressources en conséquence, en ayant recours à l'épargne.

Toutefois, la détermination du rôle que doit, ou devrait jouer l'assurance, dépend du point de vue que l'on adopte pour appréhender le risque. En naissant avec un handicap, avec la quasi certitude de développer une maladie ou d'être exposé à un environnement lui-même générateur de risque, une personne pourra avoir, au long de sa vie, des besoins médicaux plus élevés qu'une autre. Ces événements défavorables peuvent être analysés selon deux points de vue :

- ex ante, on peut les considérer comme des formes d'aléa contre lesquelles une assurance devrait pouvoir être souscrite. Il s'agirait alors de se protéger contre le fait de présenter un facteur de risque particulier ;

- ex post, on observe une hétérogénéité des risques. Les besoins de santé étant, en moyenne, différents entre les individus, l'assureur les regroupe par niveau de risque homogène et mutualise leurs risques pour éliminer la variance de leurs dépenses.

Dans la réalité, il n'est pas possible de souscrire un contrat contre le fait de naître avec un niveau de risque élevé. Il s'agit là d'une première limite à l'adéquation du service d'assurance aux besoins de couverture des individus.

Mais le paradoxe entre les deux approches du risque (ex ante / ex post) est probablement plus frappant pour une personne qui, à un moment de sa vie, voit son état de santé se dégrader et donc son niveau de risque augmenter au delà de ce qui est prévisible en moyenne pour une catégorie d'âge donnée. Considérons le cas d'une personne qui souffre d'une affection chronique lourde et dont les besoins médicaux sont élevés (SIDA, insuffisance rénale chronique, démence sénile, maladie cardio-vasculaire, ...).

Si la personne est assurée au moment la maladie survient, l'assurance doit en prendre en charge les conséquences. En particulier, elle couvrira les dépenses de santé jusqu'à échéance du contrat puisque cela correspond à son engagement contractuel. Au moment où la personne doit à nouveau souscrire un contrat d'assurance, si son état de santé n'est pas rétabli à son niveau initial, elle relève d'une catégorie de risque différente. Pour une personne souffrant d'une telle maladie, la différence entre le fait de payer une assurance actuarielle ou directement les soins (si tant est qu'elle ait des ressources suffisantes) est probablement infime. La couverture de certaines maladies pose donc un problème au regard de la nature même de la notion d'assurance qui consiste à mutualiser une variable aléatoire.

La première question soulevée ici est celle de la pertinence du concept d'assurance face à des dommages longs qui, comme le souligne Arrow ne serait pas cruciale s'il était possible de s'assurer dans le long terme. *"On a lifetime insurance basis, insurance against chronic illness makes sense, since this is both highly unpredictable and highly significant in costs. Among people who already have chronic illness, .. insurance in the strict sense is probably pointless"*¹⁶ (1963, pp. 963).

¹⁶ Sur la durée d'une vie l'assurance contre les maladies chroniques est pertinente, dans la mesure où elles sont très aléatoires et très coûteuses. Parmi les malades chroniques, l'assurance, dans l'acception stricte du terme, n'a pas de sens.

3.2.2 La question de l'assurance de long terme

La seconde question est donc celle du rapport entre, d'une part, l'évolution dans le temps des risques et, d'autre part, la durée du contrat d'assurance qui permet de se prémunir contre ces risques. Les attentes des individus à cet égard sont légitimes (a) mais le marché ne répond pas pleinement à leurs besoins (b).

a) Les attentes des individus

En dehors du vieillissement, le niveau de risque d'un individu peut évoluer au fil du temps pour des raisons aléatoires, ou, ce qui revient au même, de l'information peut être révélée qui conduit à réévaluer ce niveau de risque. Si l'on éprouve de l'aversion pour le risque, il est légitime de souhaiter s'assurer contre la réalisation d'un tel "dommage". Ceci se traduit par la nécessité, à partir d'un instant donné, de consacrer plus de ressources qu'un autre individu à se soigner (où à s'assurer).

En théorie, un individu devrait pouvoir souscrire un contrat d'assurance de long terme ou se procurer de l'assurance contre le risque de devenir un "haut risque" (Dowd et Feldman, 1992).

- Le premier type de contrat devrait se traduire par un engagement réciproque et irréversible sur la durée de la vie, dont la faisabilité est douteuse. Il présente en outre l'inconvénient de rendre la personne captive d'une compagnie d'assurance, gommant ainsi largement les incitations concurrentielles à l'efficacité et mettant l'assuré dans une position de faiblesse par rapport à l'assureur. Une assurance de cohorte dans laquelle des personnes, qui auraient des niveaux de risques initialement identiques, s'engageraient à s'assurer de façon périodique mais toujours solidaire, si certaines d'entre elles voyaient leur risque évoluer, est une solution qui procède du même principe. Elle repose aussi sur l'hypothèse que des personnes sont capables de maintenir un engagement volontaire et irréversible, alors qu'elles auraient la possibilité, pour celles dont le risque s'avère faible, de souscrire, à titre individuel, une assurance beaucoup plus avantageuse. Enfin, si le risque moyen de la cohorte évolue de façon très défavorable, le problème de la pertinence et de la faisabilité de l'assurance peut se reposer.
- Le second type de contrat peut être assimilé à une option permettant à l'assuré de faire face à l'augmentation de la prime dans le cas où il deviendrait un haut risque, voire dans le cas où il naîtrait avec un niveau de risque élevé. Ceci n'implique pas

nécessairement de choisir, au moment de la souscription de l'option, la compagnie d'assurance auprès de laquelle on s'assure effectivement. Cochrane (1995) propose une solution théorique séduisante qui procède de ce principe. Elle consiste à introduire une clause de paiement de séparation (*severance payment*). Si le risque d'un individu a augmenté pendant la durée du contrat, l'assureur lui verse une somme qui correspond à l'augmentation de la valeur actualisée de ses coûts. Cette somme permet à l'assuré de souscrire une assurance actuarielle à la période suivante. Si le risque de l'individu a au contraire diminué, celui-ci paie une somme qui correspond à la diminution de son espérance de dommage. Cette dernière contrainte étant difficile à imposer, Cochrane propose d'établir un compte spécifique, abondé par chaque agent, sur lequel l'assureur prélève ou verse des sommes qui correspondent aux paiements de séparation et prélève les primes actuarielles de court terme. Ce type de solution n'a, pour l'instant, jamais émergé spontanément du fonctionnement du marché. Il est vraisemblable que l'incertitude sur le coût futur de la prise en charge d'une pathologie et sur l'évolution des risques (sans mentionner la possibilité d'apparition d'un nouveau risque comme le SIDA) est trop radicale pour permettre la mise en œuvre et la tarification de telles polices d'assurance.

b) Les solutions de marché partielles

Il existe cependant des formes d'organisation concurrentielles qui apportent une solution à ce problème. Elles reposent sur la constitution de groupes de personnes qui présentent des niveaux de risque différents et qui se couvrent collectivement sur des périodes longues. L'économie du système repose sur un principe fondamental et différent de celui de l'assurance de cohorte : le groupe constitué se renouvelle et en particulier des personnes qui présentent des niveaux de risque faible le rejoignent régulièrement.

1. L'assurance de groupe est un premier exemple. Le fonctionnement par répartition du risque au sein de l'entreprise parie sur le renouvellement, voire l'expansion de la main d'œuvre de cette entreprise. Certes, chaque risque individuel évolue et vieillit, mais le recrutement assure que des personnes présentant de meilleurs niveaux de risque, plus jeunes, vont venir compenser cette évolution naturelle du risque collectif. Cette solution reste partielle car l'individu n'est protégé que dans la mesure où son niveau de risque n'est pas dégradé au point qu'il ne peut plus travailler.

2. La constitution de centrales d'achat est aussi une solution qui émerge depuis quelques années aux USA (Newhouse et Chairman, 1998). Des personnes ou des entités qui ont du mal à s'assurer à titre individuel se regroupent et créent une structure intermédiaire qui élargit la péréquation des risques en dehors de la structure légale de l'entreprise. La création de ces structures intermédiaires a été autorisée, voire encouragée dans un grand nombre d'Etats américains depuis le début des années quatre-vingt-dix et semble rencontrer un succès important auprès des PME américaines. En effet, en se regroupant, elles peuvent obtenir un pouvoir de marché face aux assureurs et négocier des tarifs avantageux.
3. On peut enfin remarquer que certaines mutuelles françaises qui limitent la tarification au risque cherchent à jouer un rôle similaire : les personnes jeunes ou qui présentent un niveau de risque faible paient des primes relativement plus élevées que si elles s'assuraient auprès de compagnies d'assurance privées. Elles financent de fait les dépenses des personnes couvertes par la même mutuelle qui présentent un niveau de risque plus élevé ou qui sont plus âgées.

On mesure bien que la réussite de tels projets repose sur la capacité de l'opérateur à assurer le renouvellement des risques. Il faut, en outre, que les assurés acceptent de participer à un tel système de répartition privé et qu'ils aient confiance en sa pérennité afin de pouvoir bénéficier de la forme de couverture de risque à long terme qu'elle génère.

CONCLUSION

La déclinaison des principes assurantiels au domaine de la santé nous a conduit à montrer en quoi la nature du risque couvert conditionne celle du service rendu.

En premier lieu, l'assurance porte exclusivement sur les soins et non sur la santé, comme on aurait pu l'envisager a priori. Les personnes qui souscrivent une assurance santé se protègent en réalité contre le risque de faire face à des dépenses de santé imprévues ou plus précisément, contre une variation de leurs dépenses de santé autour de la moyenne.

Comme la réparation des dommages en santé passent par la consommation par l'assuré de biens et services médicaux, et que l'assurance diminue le prix de ces services, les assureurs santé sont confrontés au risque moral *ex post*. Plus généralement, ils subissent indirectement les conséquences du fait que le fonctionnement des marchés de biens et services médicaux est imparfait. Dans un tel univers, ils peuvent surmonter ces difficultés en ajustant les niveaux de couvertures, voire en adoptant un rôle plus actif dans la gestion du risque de type *managed care*. En ce sens, les assureurs peuvent devenir des intermédiaires entre les patients et les prestataires de services.

Ceci dit, le service rendu par les assureurs reste limité dans le temps, parce que la nature des risque est difficile à estimer, particulièrement dans le long terme. Des solutions partielles à la difficulté qu'ont les assurés à se couvrir contre l'évolution de leur niveau de risque émergent sur les marchés, mais elles sont fragiles.

CONCLUSION DE LA PARTIE I

Le rôle des assureurs privés consiste essentiellement à mutualiser le risque individuel d'avoir à faire face à des dépenses de soins, au sein d'un groupe de personnes qui sont dans un état de santé comparable.

La nature du service proposé par les assureurs en santé ne diffère donc pas fondamentalement de celle du service qu'ils rendent pour couvrir d'autres dommages, si ce n'est que la tarification du risque en santé est un exercice probablement plus difficile que dans d'autres domaines. Ceci dit, il n'y a pas a priori de raisons pour que l'assureur ne puisse pas, pour un niveau de risque donné, proposer un contrat associé à une prime, contrat que les personnes choisissent ou non de souscrire. Le fait que certaines personnes n'aient pas les moyens de le faire, tant que la prime proposée est concurrentielle (actuariellement juste), n'est pas, en soi, constitutive d'un échec de marché mais relève d'un problème d'équité (Dowd et Feldman, 1992) qui peut être traité comme tel, par des mécanismes redistributifs spécifiques. Dans le même temps, en encourageant et en pérennisant les structures qui proposent des formes de couverture contre le risque de long terme, l'Etat peut contribuer à améliorer le fonctionnement du marché.

Au terme de cette partie, on pourrait donc envisager que la couverture du risque maladie soit essentiellement confiée à des assureurs privés, l'Etat se réservant la possibilité de procéder à une redistribution des ressources si cela s'avère nécessaire. Comme le suggère la théorie standard, l'allocation des ressources serait efficace et les assureurs, afin de limiter les comportements stratégiques des assurés comme des prestataires de soins, développeraient la gestion des risques, contribuant ainsi à améliorer l'efficacité productive du système de santé.

Cette approche est pourtant naïve, et nous étudions, dans la seconde partie de ce travail, en quoi la concurrence entre assureurs, notamment dans le domaine de la santé, n'est pas systématiquement synonyme d'efficacité.

PARTIE II

DE L'ANTISELECTION

A LA SELECTION :

LES LIMITES A L'EFFICACITE

PARTIE II

DE L'ANTISELECTION A LA SELECTION : LES LIMITES A L'EFFICACITE

L'analyse économique standard conduit à penser que la concurrence permet d'allouer un bien ou un service de manière efficace. Or, en observant les grandes réformes envisagées ou entreprises dans différents systèmes de santé ces dernières années, on constate une évolution qui peut sembler paradoxale :

- de nombreux pays se tournent vers un système d'assurance de type universel et semblent, par là, s'éloigner d'un modèle libéral où l'assurance privée concurrentielle jouerait un rôle prépondérant et dans lequel les individus seraient libres de choisir de s'assurer ou non ;
- dans le même temps, les vertus de la concurrence sont mises en avant de façon systématique.

L'objectif de cette seconde partie est de montrer progressivement en quoi il nous semble que cette double évolution dénote une focalisation croissante des débats autour de la nécessité de gérer la sélection pour améliorer l'efficacité des systèmes de santé.

Dans la mesure où l'on étudie l'allocation d'un service d'assurance, le terme de sélection renvoie naturellement, en premier lieu, à la notion de sélection adverse ou d'antisélection. En effet, la théorie économique suggère que si les assureurs ne sont pas en mesure de distinguer les niveaux de risque des personnes qui souhaitent souscrire un contrat, ce phénomène menace l'existence même du marché. L'Etat peut alors légitimement intervenir pour pallier ce dysfonctionnement majeur et rétablir l'allocation optimale des ressources. Nous démontrerons dans le quatrième chapitre que l'Etat peut atteindre cet objectif en laissant à des assureurs en concurrence le soin d'offrir des contrats. Cette première analyse pourrait donc a priori offrir une voie de résolution du paradoxe précédemment évoqué : une intervention calibrée du régulateur permet de garantir que toute personne, qui le souhaite, obtiendra un contrat d'assurance sur un marché concurrentiel, même si l'existence de l'asymétrie d'information interdit d'atteindre

l'optimum de premier rang. Dans le cinquième chapitre, nous recentrerons l'analyse sur le marché de l'assurance santé. Plus précisément, nous chercherons à déterminer si l'antisélection est effectivement constituée sur ce marché. Pour ce faire, nous nous appuyerons sur une analyse critique de la littérature empirique qui tend à indiquer que les prédictions du modèle théoriques sont assez largement corroborées par les travaux disponibles. Dans le même temps, il apparaîtra que le véritable enjeu de la réglementation des marchés n'est pas tant le fait que certaines personnes ne parviennent pas à obtenir autant de couverture qu'elles pourraient le souhaiter, comme le suggère le modèle canonique de Rothschild et Stiglitz, mais plutôt le fait que certaines personnes n'ont pas accès à l'assurance. Nous verrons dans le chapitre six que le concept de sélection, par le revenu et par le risque, permet de rendre compte des explications de ce phénomène. Nous concluons en montrant que l'évolution constatée des systèmes de santé traduit un consensus croissant autour de l'idée qu'il faut gérer, au nom de l'efficacité productive, la sélection dans ses différentes dimensions.

CHAPITRE 4

ASSURANCE PRIVEE ET ANTISELECTION : APPROCHES THEORIQUES

L'analyse de la nature des risques en santé nous a conduit à souligner la multiplicité des facteurs qui contribuent à déterminer l'état de santé et des raisons qui font qu'il évoluera de façon plus ou moins favorable au cours du temps. A cette approche longitudinale, on peut opposer une approche transversale et s'interroger sur le fait de savoir si, au moment où l'individu cherche à souscrire une assurance, l'assureur est en mesure de déterminer le niveau de risque de la personne. Dans l'optique du contrat de court terme qu'il va proposer, l'assureur ne cherche pas tant à connaître les raisons pour lesquelles l'état de santé d'un individu se situe à un niveau précis mais à estimer ce niveau ainsi que les risques qui lui sont associés. Or, compte tenu de la complexité d'une évaluation objective de l'état de santé, il est tout à fait envisageable que l'assureur ne puisse avoir accès à toute l'information dont l'individu dispose sur celui-ci.

Cette préoccupation fait naturellement écho à l'étude des dysfonctionnements des marchés résultant d'asymétries d'informations entre des parties qui cherchent à échanger un produit ou un service. Akerlof (1970) étudie le cas du marché des voitures d'occasion. Le vendeur connaît la qualité du produit qu'il met sur le marché mieux que l'acheteur potentiel. Pour un niveau de prix théoriquement associé à un niveau de qualité donné, l'acheteur court le risque que seuls les vendeurs qui proposent des produits de qualité inférieure se manifestent et Akerlof montre que l'existence même du marché est menacée. Rothschild et Stiglitz (1976) ont étendu la problématique au cas où, sur le marché de l'assurance, les consommateurs connaissent leur niveau de risque et les assureurs ne sont pas en mesure de distinguer les personnes qui présentent un niveau de risque élevé ou faible. Si on établit un parallèle avec la situation décrite par Akerlof, pour un prix donné (une prime), seules les personnes qui présentent un niveau de risque identique ou supérieur à celui qui est associé au niveau de prime proposé vont souscrire le contrat. Les assureurs sont victimes d'un phénomène d'antiselection ou de sélection adverse.

L'expression de sélection adverse décrit une situation générique dans laquelle un assureur couvre *"un grand nombre d'agents hétérogènes dans leurs probabilités de subir un dommage. (Il) propose un prix unique qui reflète la probabilité moyenne de perte de l'agent représentatif de cette économie.. et il devient inintéressant pour les agents dont la probabilité de subir un accident est faible de s'assurer. Il s'opère donc un phénomène de sélection par les prix et on dit qu'elle est adverse parce que ce sont les mauvais agents qui demeurent"* (Chassagnon, 1996, pp.21). Les assureurs victimes de ce phénomène risquent de se retirer du marché.

Nous étudions dans ce chapitre le fonctionnement du marché de l'assurance en présence d'antisélection, car, y compris dans le cadre d'un contrat de court terme, la nature du risque santé laisse envisager que ce phénomène pourrait entraver le fonctionnement du marché particulier que nous étudions, voire en menacer l'existence.

Les questions de l'existence d'un équilibre, d'un optimum et de leurs formes respectives en présence d'antisélection sont à l'origine d'une vaste littérature de l'économie de l'assurance qui constitue une référence incontournable. Nous ne cherchons pas tant ici à présenter formellement l'ensemble de cette littérature qu'à donner une intuition aussi rigoureuse que possible de ses principaux résultats (Section 1). Dans cette perspective, nous privilégions une approche graphique qui permet une présentation plus simple que celle qui en est généralement effectuée. Cette démarche nous conduira à souligner les limites du fonctionnement des marchés. Puis, nous prolongerons l'analyse en nous interrogeant sur les interventions possible de l'Etat pour pallier les défauts mis en évidence (Section 2).

1. OPTIMUM ET EQUILIBRE AVEC ANTISELECTION

Nous avons vu dans le premier chapitre qu'un équilibre optimal au sens de Pareto peut être atteint sur les marchés d'assurance lorsque l'information est parfaite. Si les individus présentent des niveaux de risque plus ou moins élevés, traditionnellement représentés par une probabilité différente de subir un dommage dont le montant est connu, chaque individu obtient un contrat d'assurance complète et le prix unitaire de sa couverture est égal à sa propre probabilité de dommage. Toutefois, la situation est différente lorsque l'information est imparfaite et nous décrivons tout d'abord le cadre théorique de cette analyse (1.1). Puis nous présentons les conditions que doivent remplir les allocations des individus afin d'être optimales dans le cas où il y a deux types de risques dans l'économie (1.2). Enfin, nous étudions les modèles décrivant le fonctionnement d'un marché concurrentiel entre les compagnies d'assurances, c'est-à-dire d'un marché dans lequel l'assurance est fournie intégralement par des compagnies à but lucratif, qui se font concurrence sur les niveaux des primes et sur les taux de couverture (1.3).

1.1 Cadre général de l'analyse

La formalisation de l'allocation des ressources en présence d'antisélection repose sur un petit nombre d'hypothèses qui permettent de décrire les conditions d'apparition du phénomène (1.1.1). Nous les présentons dans le cadre de l'analyse graphique que nous utilisons tout au long de ce chapitre (1.1.2).

1.1.1 Les principales hypothèses

Les hypothèses utilisées sont très proches de celles que nous avons présentées pour décrire le fonctionnement du marché de l'assurance.

- a) L'allocation des individus est égale à W dans le cas où ils ne subissent pas de dommage (état du monde 1). Dans le cas où ils subissent un dommage d'un montant x (état du monde 2), cette allocation est égale à $W-x$.

- b) La population est divisée en deux groupes : les hauts risques, dont la probabilité de subir le dommage est π_h et les bas risques, auxquels on associe la probabilité π_b avec :

$$0 < \pi_b < \pi_h < 1$$

Ces hypothèses correspondent à un cas où les individus ont tous la même richesse initiale et peuvent subir un dommage monétaire dont le montant est connu. Ils se différencient uniquement par leur probabilité de subir ce dommage.

La portée de ces modèles doit être relativisée par rapport à la nature du risque qui nous intéresse : la situation décrite correspondrait typiquement à une maladie qui peut être soignée par un traitement dont le coût est connu avec certitude et que les individus, pour des raisons exogènes (une prédisposition génétique), ont une probabilité différente de subir. Cette formalisation du risque maladie peut sembler réductrice, une assurance dans le domaine de la santé couvrant l'individu pour l'ensemble de ses dommages liés à la santé et non pour des maladies précises. Pour autant, on peut considérer que chaque personne est caractérisée, à un instant donné, par une probabilité plus ou moins élevée de subir un certain nombre de dommages dont le montant, étant donné un mode de prise en charge standard, est globalement connu. Si tant est que l'on compare des personnes qui ont le même niveau de richesse initiale, la formalisation proposée ici des différences entre niveaux de risque est une simplification plausible.

- c) La proportion de hauts risques dans la population est de λ . La probabilité moyenne de réalisation d'un sinistre sur la population est donc :

$$\pi_m = \lambda\pi_h + (1 - \lambda)\pi_b$$

- d) L'espérance d'utilité d'un individu est :

$$(1 - \pi_i)U(W_{i,1}) + \pi_i U(W_{i,2}) \quad (1)$$

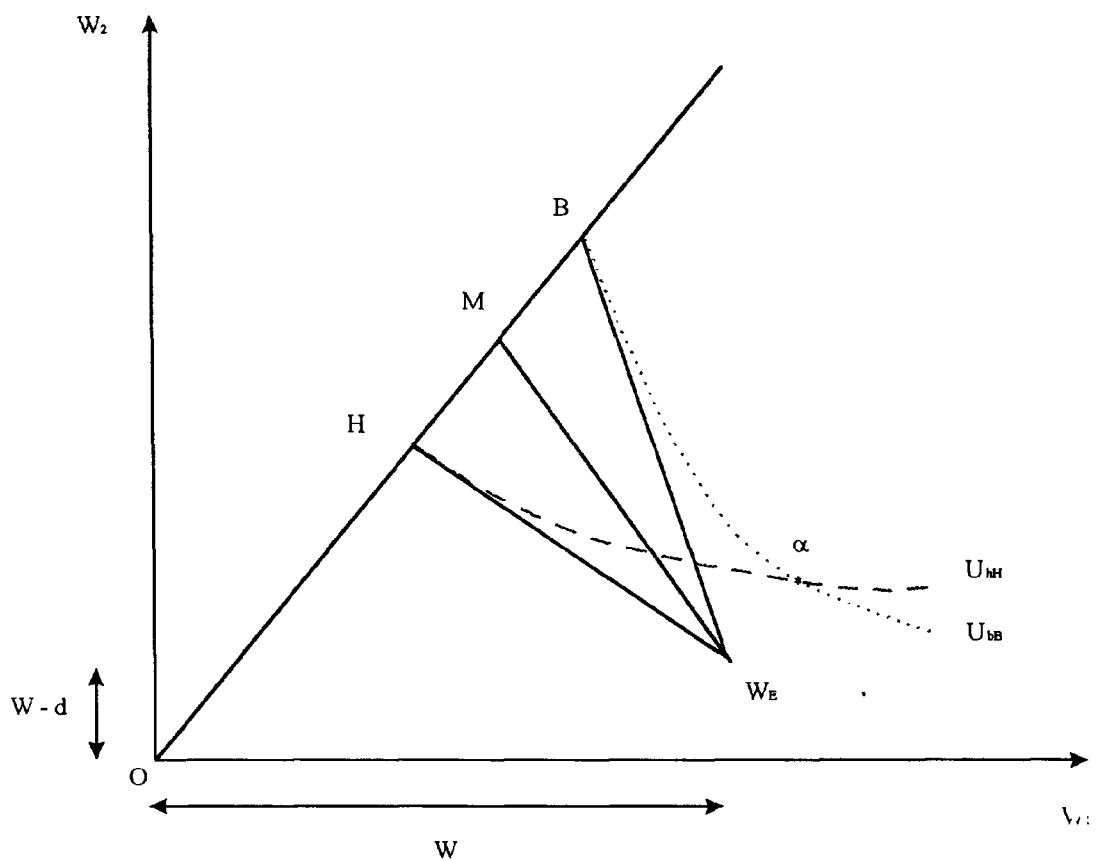
où $W_{i,1(2)}$ est la richesse de l'individu i dans l'état du monde $1(2)$. U ne dépend pas de i .

- e) Contrairement à la situation que nous avons étudiée dans le premier chapitre, on considère désormais que l'information sur le "type" de l'assuré est connue de lui seul. L'assureur ne peut distinguer qui présente un risque élevé (π_h) de qui présente un risque faible (π_b).

1.1.2 Les principes de l'analyse graphique

La figure 1 présente le cadre général de cette analyse. Les richesses des états du monde 1 et 2 sont représentées respectivement en abscisse et en ordonnée. L'allocation initiale des individus est représentée par le point W_E .

Figure 1
Cadre Général
Equilibre et optimum en information parfaite



a) Les droites actuarielles

Les droites qui partent du point W_E sont des droites d'assurance actuarielle. Elles représentent les opportunités de transfert de richesse pour les individus entre les deux états du monde.

Etude d'une droite

Lorsque p est le prix de l'assurance et que les individus achètent une quantité X de couverture, ils reçoivent $X-pX$ (remboursement net de prime) dans l'état 2 et versent $-pX$ dans l'état 1. Le taux d'échange entre les états du monde, qui représente la pente de la droite étudiée est donc :

$$-\frac{1-p}{p}$$

Pour l'assureur, la droite de prix actuariel (*fair-odds line*) est une droite de profit nul : S'il vend l'assurance au prix $p=\pi$, l'ensemble des primes qu'il reçoit lui permet de couvrir l'ensemble des sinistres. Remarquons que l'intersection entre chaque droite de prix actuariel et la droite de pente 1 représente une situation d'assurance actuarielle complète puisque $W_1=W_2$.

Les trois droites représentées

La droite la moins pentue (W_EH) représente la droite actuarielle des hauts risques ($p=\pi_h$). Un contrat d'assurance représenté sur cette droite rapporte un profit nul à l'assureur: s'il est souscrit par des hauts risques. (W_EB) est la droite actuarielle des bas risques ($p=\pi_b$) : les contrats représentés sur cette droite rapporteront un profit nul s'ils ne sont souscrits que par les bas risques.

La droite intermédiaire (W_EM) représente la droite actuarielle moyenne de la population ($p=\pi_m$). Pour réaliser un profit nul, il faut qu'un contrat sur cette droite soit souscrit par les deux types d'individus dans les mêmes proportions que leur représentation dans la population.

Si un individu obtient une allocation à droite (resp. à gauche) de la courbe actuariellement juste qui correspond à son propre niveau de risque, il génère une perte (resp. un gain) pour l'assureur. Si par exemple, tous les individus ont une allocation sur la droite actuarielle moyenne, les bas risques génèrent un gain, compensé exactement par la perte réalisée sur les hauts risques.

b) Les préférences des individus

Les courbes d'indifférences

La courbe d'indifférence représente l'ensemble des combinaisons (W_1, W_2) qui confèrent le même niveau d'espérance d'utilité à un individu. Les courbes d'indifférence des individus dans le repère (W_1, W_2) sont convexes, ce qui reflète le fait que les individus éprouvent de l'aversion pour le risque. Plus la courbe d'indifférence d'un individu est éloignée de l'origine, plus son espérance d'utilité est élevée.

Le taux marginal de substitution (TMS)

Le taux marginal de substitution entre l'état 2 et l'état 1 en un point du plan (W_1, W_2) donne la quantité que l'individu i doit recevoir dans l'état 2, si sa richesse disponible dans l'état 1 diminue à la marge, pour que son espérance d'utilité reste constante, soit :

$$-\frac{dW_2}{dW_1}$$

Graphiquement, le TMS est égal à la valeur absolue de la pente en chaque point d'une courbe d'indifférence.

Calcul du TMS

L'annulation de la différentielle totale de l'équation (1) au point (W_1, W_2) permet de calculer le TMS :

$$(1 - \pi_i)U'(W_1)dW_1 + \pi_i U'(W_2)dW_2 = 0$$

Soit :

$$-\frac{dW_2}{dW_1} = \frac{1 - \pi_i}{\pi_i} \frac{U'(W_1)}{U'(W_2)}$$

Par construction, en un point quelconque du plan, la pente de la courbe d'indifférence des bas risques est plus élevée en valeur absolue que celle des hauts risques puisque $\pi_h > \pi_b$ (point α).

Pour rappel, on représente sur la figure 1 (pp. 110) l'équilibre et l'optimum en information parfaite. Chacun obtient l'allocation qui correspond à son niveau d'utilité le plus élevé (situé sur la courbe d'indifférence la plus éloignée de l'origine) et qui, dans le

même temps, est faisable. Cette seconde condition se traduit par le fait que l'allocation de l'individu est sur la droite actuarielle qui représente ses opportunités d'assurance en fonction de son niveau de risque. Les hauts risques obtiennent donc H et les bas risques B .

1.2 Caractérisation des optima en présence d'antisélection

Conformément à la méthode que nous avons adoptée dans le premier chapitre, nous commençons par présenter, dans le contexte où l'information est asymétrique, les caractéristiques de l'optimum. Il constitue une norme à l'aune de laquelle on peut évaluer les résultats qui émergent de la concurrence. L'asymétrie d'information sur le niveau de risque modifie la problématique de détermination de l'optimum (1.2.1) ainsi que les allocations optimales (1.2.2).

1.2.1 Problématique

Déterminer l'optimum revient, dans le cas général, à maximiser la somme pondérée des utilités des individus en respectant les contraintes de ressources globales de l'économie.

Dans le contexte où les individus sont les seuls à connaître leur propre niveau de risque, les allocations constitutives de l'optimum de premier rang ne peuvent être optimales. En effet, les hauts risques ont intérêt à choisir l'allocation qui correspond à l'assurance complète des bas risques. Dans ce cas, la contrainte de ressources globale de l'économie n'est pas respectée. Il faut donc introduire dans le programme de recherche de l'optimum des contraintes de révélation : les hauts (resp. bas) risques doivent préférer les allocations qui leur sont destinées à celles qui sont proposées aux bas (resp. hauts) risques. La notion d'optimalité est donc altérée par l'asymétrie d'information. On se réfère dans ce cas au concept d'optimalité de second rang qui caractérise une situation où il n'existe pas de mécanisme révélateur qui donne plus d'utilité à un ou des agents, sans diminuer l'utilité d'autres.

La contrainte de ressources globale de l'économie est la suivante :

$$\lambda[(1 - \pi_h)W_{h1} + \pi_h W_{h2}] + (1 - \lambda)[(1 - \pi_b)W_{b1} + \pi_b W_{b2}] = W - \pi_m x$$

Cette équation peut être interprétée comme une contrainte de profit global nul pour un assureur qui offrirait des contrats actuariels ($p_i = \pi_i$) en jouant simultanément sur deux variables : le prix et la quantité de couverture. Il vend un couple prix-quantité, qu'il détermine en sachant que les individus peuvent cacher leur type.

On peut considérer que l'assureur joue sur deux critères (Chassagnon, 1996, pp.33) :

- un critère de séparation : l'assureur offre des contrats moins attrayants (assurance incomplète) aux bas risques afin de limiter l'intérêt des hauts risques pour ces derniers.
- un critère de mutualisation : il fait payer aux hauts risques un prix moins qu'actuariel et aux bas risques un prix unitaire plus élevé que leur probabilité de dommage. Il compense les pertes sur le premier contrat par les gains sur l'autre. Nous avons signalé dans le premier chapitre que le terme de mutualisation comportait une ambiguïté. Nous préférons donc lui substituer celui de péréquation ou de subvention croisée (sous entendue des hauts par les bas risques). Nous verrons par la suite que cette péréquation nécessaire à l'obtention de l'optimum peut s'avérer incompatible avec un mécanisme d'allocation marchand.

1.2.2 Résolution

On détermine les allocations optimales de second rang (Crocker et Snow, 1985, Pannequin, 1992) selon la méthode suivante : on fixe pour les bas risques un niveau minimum d'utilité (U_{hmin}) puis on maximise l'utilité des bas risques en prenant en compte l'ensemble des contraintes de sélection et de ressources.

a) Conditions nécessaires

Si on se limite¹ aux cas où $U_{hmin} < U_{hM}$ (niveau maximal d'utilité des hauts risques au point M), les solutions de ce problème remplissent les conditions suivantes :

- Condition 1 : les hauts risques reçoivent une assurance complète, autrement dit leur richesse est la même dans les deux états du monde. Sur la figure 2, leur allocation se situe sur le segment $[OM]$.

¹ Si on admet que U_{hmin} peut être supérieure à U_{hM} , les conditions 1 et 2 "s'inversent". L'optimum est tel que les hauts risques ont une allocation au dessus de la bissectrice (ils sont sur-assurés, ce qui économiquement n'est pas acceptable) et les bas risques reçoivent une assurance complète située sur la courbe d'indifférence des hauts risques. Dans ce cas, les contrats des bas risques rapportent un profit positif et ceux des hauts risques un profit négatif. Pour plus de détail, voir Crocker et Snow pp.211.

Quand l'allocation des hauts risques s'éloigne de M vers O , l'assureur offre deux contrats tels que la perte sur l'un soit compensée par le gain sur l'autre². Pour un point H^* qui correspond à une assurance complète des hauts risques (condition 1), le contrat offert par l'assureur aux bas risques B^* doit se situer sur la parallèle à la droite actuarielle des bas risques passant par D , point d'intersection entre la droite actuarielle moyenne et la parallèle à la droite actuarielle des hauts risques passant par H^* .

Plus précisément, pour un point H^* qui correspond à une assurance complète des hauts risques, l'assureur détermine le contrat offert aux bas risques de la façon suivante :

- comme H^* est entre M et H_{RS} , le contrat H^* détermine une perte (puisque le prix payé par les hauts risques est inférieur au prix actuariel) qu'il faut donc que l'assureur compense par un gain strictement équivalent sur les bas risques ;
- tous les points de la parallèle à $(W_E H_{RS})$ passant par H^* réalisent le même niveau de perte. Intuitivement, on peut l'expliquer de la façon suivante : tous les points de $(W_E H_{RS})$ amènent par définition un profit de même niveau à l'assureur, profit nul puisque c'est la droite actuarielle des hauts risques. Une droite parallèle à $(W_E H_{RS})$ donne donc un niveau de perte (ou de gain) constant à l'assureur quand il est souscrit par les hauts risques ;
- au point d'intersection D de la droite parallèle à $(W_E H_{RS})$ passant par H^* et de la droite actuarielle moyenne, on trouve un contrat qui, s'il est aussi choisi par les bas risques, compense la perte réalisée par le contrat déficitaire H^* . En effet, le point D étant sur la droite actuarielle moyenne, il réalise un profit globalement nul s'il est souscrit par tous. Autrement dit, il permet à l'assureur de réaliser un gain sur les bas risques qui compense exactement la perte sur les hauts risques au point D ;
- si on suit un raisonnement équivalent sur les bas risques, on voit que tous les points de la parallèle à la droite actuarielle des bas risques passant par D amènent un gain à l'assureur qui permet de compenser la perte sur H^* .

² Pour $U_{hmin} = U_{hM}$, le seul contrat possible est le contrat M d'assurance complète qui réalise par construction un profit nul sur l'ensemble de la population puisqu'il est sur la droite d'assurance actuarielle moyenne.

Le contrat B^* qui correspond à la condition 2 énoncée ci-dessus se situe donc à l'intersection entre la courbe d'indifférence des hauts risques la plus élevée passant par H^* (U_{hH^*}) et la parallèle à la droite actuarielle des bas risques passant par D . A ce point, on a, par construction, indifférence des hauts risques entre les deux contrats proposés. Remarquons que si H^* est entre H_{RS} et O , on construit de la même façon les couples de contrats (H^*, B^*) mais on aboutit à une situation où les contrats sur les hauts risques réalisent un profit qui compense une perte réalisée sur les contrats des bas risques.

La courbe (MI) représente l'ensemble des contrats offerts aux bas risques qui, associés à un contrat de hauts risques entre O et M , aboutissent à un profit global nul pour l'assureur et pour lesquels la contrainte de sélection des hauts risques est saturée.

c) Détermination des optima de second rang

Tous les couples associant un contrat pour les hauts risques sur $[OM]$ et un contrat pour les bas risques sur (MI) ne sont pas des optima. On peut établir deux séries de résultats selon le niveau de U_{hmin} qui est fixé arbitrairement.

1. Considérons le point B^* pour lequel la courbe d'indifférence des bas risques et la courbe (MI) sont tangentes. Ce point B^* est le contrat préféré des bas risques sur la courbe (MI) . Il est apparié à un contrat H^* qui apporte U_{hH^*} aux hauts risques. Si le niveau minimal d'utilité des hauts risques est fixé en dessous de ce seuil, le couple de contrats à profit globalement nul qui correspond au niveau d'utilité fixé est Pareto dominé par (B^*, H^*) ³. Dans ce cas, il existe un niveau minimal d'utilité pour les hauts risques, U_{hH^*} , en deçà duquel l'optimum de second rang n'est pas atteint. Ceci signifie que les bas risques ont intérêt subventionner les hauts risques.
2. Si U_{hmin} est fixé arbitrairement entre U_{hH^*} et U_{hM} , le meilleur contrat que peuvent obtenir les bas risques est situé entre B^* et M .

³ Par exemple, si U_{hmin} est fixé à U_{hRS} le couple (H_{RS}, B_{RS}) n'est pas un optimum. En passant à (B^*, H^*) , on augmente l'utilité des deux catégories d'individus.

Finalement, les contrats optimaux sont des couples associant pour les hauts risques un contrat entre H^* et M et, pour les bas risques, le contrat de (MI) pour lequel les hauts risques sont indifférents (situé plus précisément entre B^* et M). Parmi ces optima, celui qui est préféré par les bas risques est B^{*4} .

Remarquons enfin que le lieu des point H^* et B^* dépend de la proportion de hauts risques dans la population. Quand elle diminue, M se déplace vers B , le point B^* se déplace vers le point B_{RS} et H^* se déplace vers H_{RS} . Il existe donc un seuil limite λ_0 pour lequel aucune subvention des hauts risques par les bas risques n'est nécessaire pour atteindre l'optimum de second rang. Si la proportion λ de hauts risques diminue sous ce seuil, certains optima de second rang correspondent à une situation où les hauts risques subventionnent les bas risques puisque l'allocation des hauts risques est à gauche de H_{RS} et celle des bas risques à droite de B_{RS} ⁵.

1.3 Equilibre des marchés d'assurance

On étudie désormais l'existence et des propriétés de l'équilibre. L'existence d'un équilibre en présence d'antisélection sur les marchés d'assurance est-elle menacée comme le laisse entendre l'analyse l'Akerlof? Existe-t-il un mécanisme qui permette aux assureurs en concurrence de surmonter l'imperfections de marché liée à l'asymétrie d'information? Dans l'hypothèse où un tel mécanisme existe, permet-il d'atteindre l'optimum de second rang?

Rothschild et Stiglitz montrent qu'un l'équilibre peut émerger mais l'existence de cet équilibre n'est pas garantie et il n'est pas nécessairement optimal (1.3.1). D'autres auteurs ont proposé des solutions qui permettent de pallier ces défauts. Selon les hypothèses posées sur le comportement stratégique des acteurs (assurés et assureurs), l'existence et la nature de l'équilibre varient. Nous présentons les principales contributions et voyons

⁴ Considérons un instant la méthode (non développée ici car moins intuitive sur le plan graphique) qui permet la détermination analytique de l'optimum en maximisant la somme pondérée des utilités des individus. On fixe pour chaque catégorie d'individu un poids arbitraire entre 0 et 1 de façon à ce que la somme des poids soit nulle. On maximise alors la somme des utilités pondérées sous contrainte de ressource globale et de sélection. (H^*, B^*) est la solution du problème quand le poids alloué aux hauts risques est nul. M est obtenu quand le poids des hauts risques est égal à λ . Si le poids arbitrairement attribué aux hauts risques est supérieur à λ (jusqu'à 1), les allocations des hauts risques sont au-dessus de la droite d'assurance complète. Ils obtiennent une allocation plus élevée en cas de dommage (sur-assurance), solution que l'on exclut en général de l'analyse (Pannequin pp.109 pour une discussion plus détaillée). (B^*, H^*) correspond donc à une situation où l'on "favorise" implicitement les bas risques.

⁵ La valeur de λ_0 dépend de l'écart relatif entre π_b et π_h .

dans quelle mesure les solutions de marché qui émergent sont optimales de second rang (1.3.2).

1.3.1 Le modèle de Rothschild et Stiglitz

Rothschild et Stiglitz (1976, noté ci-après RS) ont montré que si les assureurs offrent des couples de contrats qui se différencient par la quantité d'assurance proposée, les assurés sont conduits à révéler leur niveau de risque et un équilibre de marché peut émerger.

a) Principe

Définition⁶ : *Un équilibre de RS est un ensemble de contrats définis en deux dimensions (prix, quantité) tel que, lorsque les différentes catégories de consommateurs choisissent un contrat particulier de façon à maximiser leur utilité, chaque contrat réalise un profit non négatif et il n'existe pas d'autre contrat en dehors de cet équilibre qui, s'il était offert, réaliserait un profit.*

L'élément crucial de cette définition est que l'équilibre est atteint *lorsqu'il n'existe pas de contrat hors de l'équilibre qui, s'il était offert, réaliserait un profit*. Cette stratégie est construite selon le modèle de Cournot-Nash : chaque assureur considère l'offre de ses concurrents comme donnée lorsqu'il résout son propre programme de maximisation du profit. On peut décrire la séquence des décisions de la façon suivante :

1. dans un premier temps, les assureurs offrent des contrats ;
2. dans un second temps, les assurés choisissent les contrats qu'ils préfèrent parmi ceux qui sont offerts.

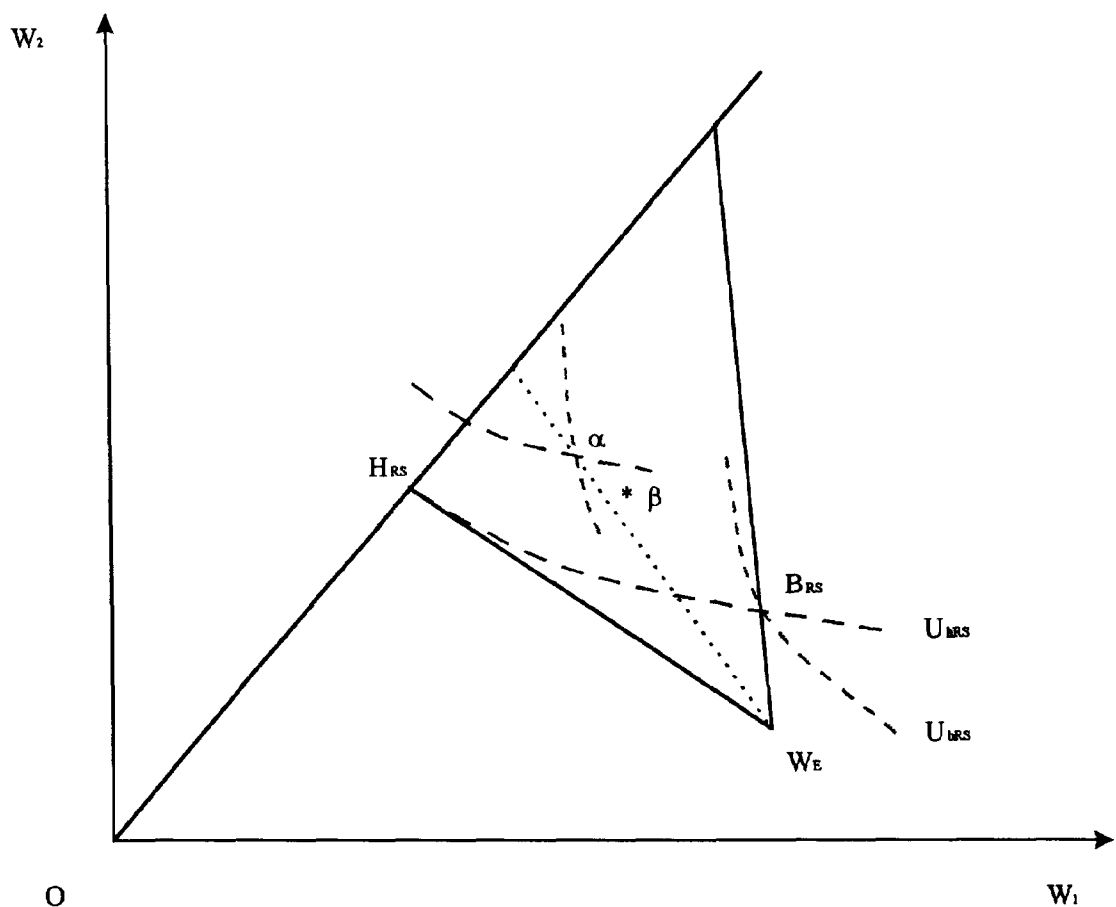
Si le ou les contrats offerts sont à profit nul, on peut a priori envisager deux types de solutions :

1. un contrat poolé ou mélangeant qui offre la même quantité d'assurance à tous et qui est situé sur la droite actuarielle moyenne. Il doit être souscrit par les deux catégories d'individus ;
2. un équilibre séparateur où deux contrats différents sont offerts et dans lequel les contraintes de révélation mentionnées plus haut sont respectées.

⁶ Cette définition de l'équilibre, comme les suivantes est inspirée de celle proposée par Crocker et Snow (1985).

Le premier type de solution ne peut constituer un équilibre de RS. Si par exemple le contrat poolé α est offert (figure 3), un assureur peut proposer un contrat β préféré à α par les bas risques mais pas par les hauts risques (en choisissant ce contrat ils sont sur une courbe d'indifférence plus proche de l'origine et donc leur utilité diminue). Les bas risques se regroupent sur β alors que les hauts risques restent sur le contrat poolé α . Le contrat β réalise alors un profit strictement positif puisqu'il est vendu à un prix actuariel plus élevé que celui que devraient payer les bas risques et qu'il n'est choisi que par eux. Il existe donc au moins un contrat *qui, s'il est offert, réalise un profit*. Le contrat poolé ne peut être un équilibre au sens de RS.

Figure 3
Equilibre de Rothschild et Stiglitz



Plus généralement, la contrainte de profit nul sur chaque contrat exclut toute subvention croisée et l'équilibre de RS, quand il existe, est unique (figure 3) : les bas risques obtiennent une assurance partielle B_{RS} à un prix actuariel alors que les hauts risques obtiennent une assurance complète au prix actuariel H_{RS} . Les hauts risques sont indifférents entre le contrat qui leur est offert et celui qui est choisi par les bas risques⁷. L'intuition de ce résultat est simple : les assurés se voient offrir le choix entre un contrat complet et un contrat avec franchise, ceux qui se savent plus exposés optant pour le contrat d'assurance complète. Par rapport à la situation d'information parfaite, les bas risques sont pénalisés dans la mesure où, pour obtenir un contrat actuariel, ils doivent faire des concessions sur le niveau de couverture.

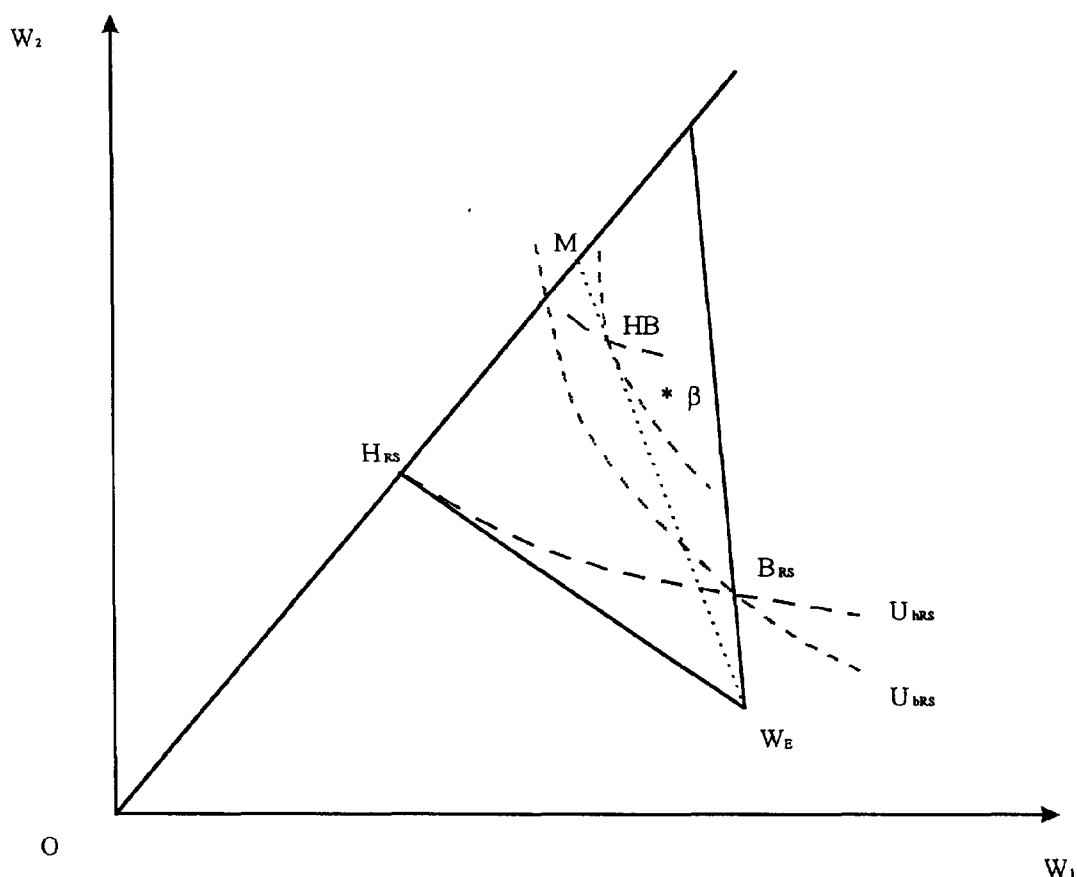
b) Limites

L'équilibre défini par RS pose deux difficultés :

- si la proportion de hauts risques est au dessous du seuil λ_o défini précédemment, l'équilibre de RS n'est pas un optimum de second rang puisque celui-ci ne peut être atteint que par le biais d'un niveau minimal de subvention croisée que RS proscrit. La contrainte de profit nul sur chaque contrat est coûteuse pour les bas risques qui seraient prêts à subventionner les hauts risques relativement peu nombreux dans ce cas de figure pour obtenir une couverture plus complète ;
- en outre, si λ continue à diminuer et passe en dessous d'un seuil $\lambda_{RS} < \lambda_o$, l'équilibre de RS n'existe pas. On peut en effet trouver au moins un contrat (par exemple HB dans la figure 4, pp. 122) qui, s'il est proposé, attire tous les individus et réalise un profit (nul). Ce contrat est un contrat pooling qui ne peut être, comme nous l'avons vu, un équilibre au sens de RS. Il n'existe donc aucune configuration stable de contrats formant un équilibre dans cette situation. Elle correspond graphiquement à l'ensemble des cas où la courbe d'indifférence des bas risques qui passe par le point B_{RS} , coupe la droite actuarielle moyenne. Dans ce cas, il existe au moins un point de cette droite qui est préféré par les bas risques à B_{RS} .

⁷ On ajoute en fait une hypothèse supplémentaire : les hauts risques, indifférents entre les deux contrats, choisissent malgré tout le contrat d'assurance complète.

Figure 4
Absence d'équilibre de Rothschild et Stiglitz
Equilibre de Grossman et Wilson



Face à ces deux difficultés, différentes solutions ont été proposées, que l'on analyse maintenant.

1.3.2 La sophistication du modèle de RS

a) Modification des comportements stratégiques des acteurs

Les principales contributions qui établissent l'existence d'un équilibre, quel que soit le niveau de λ , modifient les hypothèses de comportement des acteurs du marché. Le principe d'un comportement de type Nash est abandonné : les agents anticipent que leurs propres actions auront des conséquences sur les stratégies des autres acteurs du marché. Nous présentons les contributions de Wilson et Grossman qui montrent que dans le cas où l'équilibre de RS n'existe pas, l'équilibre pooling préféré par les bas risques émerge.

Wilson (1977) propose une nouvelle définition de l'équilibre :

Un équilibre est un ensemble de contrats tel que, lorsque les différentes catégories de consommateurs choisissent un contrat particulier de façon à maximiser leur utilité, chaque contrat réalise un profit non négatif et il n'existe pas d'autre contrat ou ensemble de contrats qui, offert, réalise un profit positif même lorsque les premiers contrats existant sur le marché sont retirés.

La définition de l'équilibre adoptée par Wilson intègre donc l'anticipation par un assureur de la réaction de ses concurrents qui retirent du marché des contrats devenus déficitaires, du fait de sa propre entrée dans le marché.

L'hypothèse comportementale de Wilson permet de rétablir l'existence d'un équilibre unique dans le cas où $\lambda < \lambda_{RS}$ ⁸. Dans ce cas, un seul contrat est offert à tous les individus : c'est l'équilibre pooling préféré par les bas risques⁹ (*HB* sur la figure 4, page précédente). Le prix de ce contrat est p_m , le coût actuariel moyen et l'assurance n'est pas complète.

La dimension stratégique ajoutée au comportement de l'assureur aboutit à éliminer la tentation de déstabiliser le pooling en écrémant les hauts risques. Si un assureur offre le contrat β , il attire tous les bas risques et réalise un profit positif. Le contrat *HB*, initialement offert, qui n'est plus souscrit que par les hauts risques, devient déficitaire (puisque'il a un profit nul lorsqu'il est souscrit par tous). Il est donc retiré du marché. Le contrat β qui reste seul offert est souscrit par tous et réalise un profit négatif puisqu'il est à droite de la droite actuarielle moyenne. L'assureur anticipant cette réaction n'offre pas le contrat β et *HB* émerge comme équilibre unique.

L'hypothèse sur le comportement de l'assureur énoncée par Wilson est très forte : elle n'offre pas de contrat qui sera profitable à court terme et qui ne le sera plus quand les autres contrats seront retirés du marché. On peut décrire la séquence des décisions qui aboutissent à l'équilibre de Wilson de la façon suivante :

1. dans un premier temps, les assureurs offrent des contrats ;
2. dans un second temps, les assurés choisissent les contrats qu'ils préfèrent parmi ceux qui sont offerts ;

⁸ Dans le cas où $\lambda > \lambda_{RS}$, l'équilibre qui émerge est celui de RS.

⁹ Un autre contrat pooling ne serait jamais choisi par les bas risques.

3. les assureurs retirent les contrats qui sont déficitaires et les assurés refont un choix parmi ceux qui restent offerts (retour à l'étape 2).

Grossman (1979) propose une autre interprétation des deuxième et troisième étapes qui met l'accent sur un comportement stratégique des assurés. L'existence de ce comportement est justifiée selon Grossman par l'observation du fonctionnement des mécanismes de souscription d'un contrat d'assurance. Les assureurs proposent une gamme de contrats et les individus manifestent le souhait de souscrire telle ou telle option. Ce faisant, ils peuvent être amenés à révéler leur niveau de risque ; les assureurs peuvent alors rejeter leur demande pour les contrats qu'ils ont initialement préféré. Connaissant cette possibilité, les hauts risques évitent de révéler l'information qu'ils détiennent sur leur propre risque et adoptent une stratégie de dissimulation (*dissembling behaviour*). En particulier, ils ne choisissent pas un contrat offert qui deviendra déficitaire s'il est choisi par eux. Si l'on considère que les assurés ont un comportement stratégique de ce type, le contrat pooling préféré par les bas risques émerge comme solution lorsque l'équilibre de RS n'existe pas (figure 4, pp. 122). En effet, lorsque le contrat β est proposé, les hauts risques le souscrivent alors même qu'ils lui préfèrent le contrat HB . Ils imitent en cela les bas risques car ils savent que s'ils restent sur le contrat HB , ils seront "repérés" comme étant des hauts risques, le contrat sera déficitaire et donc disparaîtra. Mais comme en souscrivant β , les hauts risques le rendent déficitaire (puisque β est à droite de la droite actuarielle moyenne), les assureurs se voient obligés de le retirer.

Les anticipations "à la Wilson" et "à la Grossman" aboutissent à faire émerger le même équilibre là où l'équilibre de RS n'existait pas¹⁰. Les secondes proposent probablement une vision plus réaliste des relations entre assureur et assuré. Les souhaits exprimés par des individus révèlent leur niveau de risque. L'assureur utilise cette information et peut refuser à une personne de souscrire un contrat qui ne lui est pas destiné.

¹⁰ D'autres hypothèses d'anticipations ont été proposées. Riley (1979,) définit par exemple une règle d'anticipation des assureurs qui aboutit à l'émergence de l'équilibre séparateur de RS quel que soit λ .

Dans ces deux modèles, l'équilibre existe donc, quel que soit le niveau de λ et il est unique. En revanche, il n'est pas nécessairement optimal. Lorsque l'équilibre pooling émerge, le prix unitaire payé par les bas risques est plus élevé que leur prix actuariel (et vice versa pour les hauts risques). Les bas risques subventionnent donc implicitement les hauts risques puisque ces derniers ont une incidence négative sur le profit globalement nul de cet unique contrat. Mais comme les hauts risques sont relativement peu nombreux, les bas risques y trouvent un avantage en terme de niveau de couverture atteint par rapport à une solution séparatrice qui ne peut être atteinte. Ceci dit, l'équilibre pooling de Wilson n'est généralement pas un optimum de second rang. D'une part, les hauts risques ne bénéficient pas d'une assurance complète. D'autre part, le niveau de subvention croisée imposé par la contrainte de tarification actuarielle moyenne, est plus élevé que celui qui correspond à l'optimum. En effet, les points de (M) sont tous à droite de la droite actuarielle moyenne, ce qui signifie que les bas risques dans le cadre de l'optimum paieraient un prix unitaire moins élevé que sur la droite actuarielle moyenne. Le seul contrat pooling optimal qui pourrait émerger dans cette configuration (quand $\lambda_{RS} < \lambda$) est celui qui offre une assurance complète à tous et qui correspond au point M . Il faudrait pour cela qu'il soit le contrat préféré des bas risques sur la droite actuarielle moyenne.

Le tableau ci-dessous résume les caractéristiques des équilibres de type Wilson ou Grossman.

Tableau 1 : Caractérisation des l'équilibres de Wilson et Grossman

	Forme de l'équilibre	Optimalité
$\lambda > \lambda_0$	Equilibre séparateur de RS	De second rang
$\lambda_0 > \lambda > \lambda_{RS}$	Equilibre séparateur de RS	Non optimalité
$\lambda_{RS} > \lambda$	Equilibre pooling	Non, à moins que M soit le contrat pooling préféré des bas risques (optimalité de premier rang).

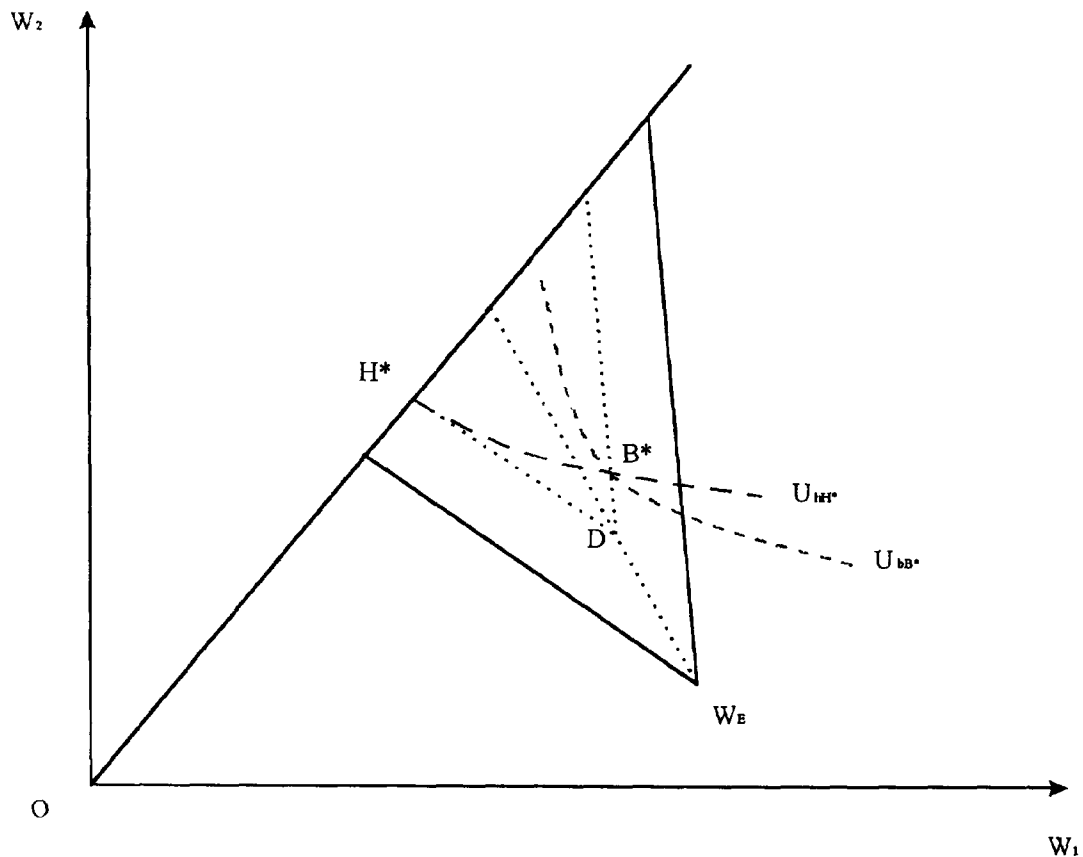
b) La redéfinition de la contrainte de profit

Le concept d'équilibre de Miyasaki (1977), appliqué par Spence (1978) au marché de l'assurance s'appuie sur des anticipations à la Wilson et modifie la nature de la contrainte de profit des assureurs. Dans ce cas, l'équilibre existe, est unique et optimal.

Un équilibre de Miyasaki-Spence-Wilson est un ensemble de contrats qui réalise un profit globalement non négatif tel que, lorsque les différentes catégories de consommateurs choisissent un contrat particulier de façon à maximiser leur utilité, il n'existe pas d'autre contrat ou ensemble de contrats qui, offert, réalise un profit positif même lorsque les premiers contrats existant sur le marché sont retirés.

L'idée est que l'assureur peut offrir un ensemble de contrats rapportant un profit globalement positif et non plus seulement des contrats individuellement profitables. La nouvelle contrainte de profit que doit respecter l'assureur est la même que la contrainte de ressources globales utilisée précédemment pour déterminer les optima de second rang. Les ensembles de contrats qui respectent cette contrainte sont ceux associant un contrat sur $[OM]$ pour les hauts risques à un contrat sur (MI) pour les bas risques (figure 5). Les comportements d'anticipation stratégiques des entreprises sont ceux proposés par Wilson.

Figure 5
Equilibre de Miyasaki-Spence



Dans la configuration proposée, l'équilibre peut être séparateur sans que les personnes paient un prix actuariel. Ceci permet un ajustement plus fin du transfert de richesse des bas vers les hauts risques que dans le cas où ces "transferts" même virtuels, s'effectuent au prix actuariel moyen (contrat pooling). La péréquation mise en place par l'assureur reste néanmoins limitée par le fait que les bas risques doivent y participer volontairement.

Le modèle de Miyasaki-Spence permet d'atteindre un équilibre unique, qui est un optimum de second rang. La solution qui émerge dépend du niveau de λ . On distingue trois cas, détaillés ci-dessous.

1. Lorsque $\lambda = \lambda_0$, aucune subvention croisée n'est nécessaire pour atteindre l'optimum. Le contrat B^* , qui est le contrat répondant à la contrainte de profit global nul préféré par les bas risques, est par définition confondu avec le contrat B_{RS} . Il n'est pas possible de trouver un autre contrat préféré par les bas risques qui respecte la contrainte de profit global nul. Le contrat B_{RS} sera donc offert et choisi par tous les bas risques. Le contrat H_{RS} est offert aux hauts risques. Il sature leur contrainte de sélection et réalise un profit nul. (B_{RS}, H_{RS}) est donc un équilibre et un optimum de second rang.
2. Lorsque $\lambda > \lambda_0$, le contrat à profit globalement nul préféré des bas risques, B^* est situé entre B_{RS} et I . Ce contrat ne peut être offert que s'il est subventionné par les hauts risques. Dans ce cas, ces derniers obtiennent un niveau d'utilité inférieur à celui que leur confère une assurance complète au prix actuariel qui correspondrait à une allocation égale à H_{RS} . Ils n'accepteront pas une telle situation. Le meilleur contrat, que peuvent alors obtenir les bas risques et qui n'attire pas les hauts risques, est B_{RS} . Le couple (B_{RS}, H_{RS}) fait partie des optima de second rang puisque $\lambda > \lambda_0$.

Ainsi, lorsque $\lambda \geq \lambda_0$, la solution de RS émerge et elle est optimale de second rang. Il n'y a pas de subvention croisée.

3. Lorsque $\lambda < \lambda_0$, l'équilibre de RS n'existe pas toujours et il n'est jamais optimal. Dans la configuration où les assureurs peuvent effectuer des subventions croisées entre les contrats, on montre que si le couple (B^*, H^*) est offert, aucune stratégie cherchant à séparer les individus ne peut déstabiliser cette offre. En effet :
 - d'une part, il n'existe pas d'autre contrat préféré par les bas risques qui n'attire pas directement ou indirectement les hauts risques. Si on offre un contrat préféré par les deux catégories d'individu, il réalise un profit négatif. Si on offre un contrat préféré

par les bas risques mais pas par les hauts risques, B^* est délaissé par les bas risques et H^* , déficitaire, est retiré du marché. Le nouveau contrat (selon la règle de Wilson) ne représente donc pas une menace crédible car la disparition de H^* le rendra déficitaire ;

- d'autre part, si un assureur offre un contrat à profit globalement nul autre que (B^*, H^*) il n'arrivera pas à attirer les bas risques qui peuvent s'attendre à obtenir un niveau d'utilité au moins égal à U_{bB^*} .

Sur la base du concept d'équilibre étudié ici, pour chaque niveau de λ , il existe un équilibre unique, optimum de second rang. Lorsque $\lambda < \lambda_0$, les bas risques subventionnent les hauts risques et le niveau minimal de subvention croisée nécessaire à l'obtention de l'optimum s'instaure.

L'équilibre ainsi obtenu nécessite que les assureurs soient suffisamment coordonnés pour accepter de vendre les contrats déficitaires. Les pertes seront compensées par des bénéfices sur d'autres contrats. Il faut donc admettre que :

- la répartition des hauts et des bas risques au sein de chaque firme qui émerge par le jeu de la concurrence est strictement proportionnelle à celle qui prévaut dans la population ;
- aucune firme ne sera tentée de mettre en œuvre une stratégie déstabilisante en cherchant à attirer des bas risques ;

Ces hypothèses sont nécessaires à l'émergence de tous les équilibres pooling et ne suffisent donc pas à justifier en quoi le concept proposé par Miyasaki-Spence-Wilson est moins susceptible que d'autres de décrire l'équilibre des marchés d'assurance. La critique la plus fondamentale qui lui est adressée est qu'il permet aux firmes d'offrir des contrats qui réalisent un profit positif, "*a state of affairs which could not persist in a market with truly free entry*"¹¹ (Crocker et Snow, 1985, pp. 1145). Un équilibre de marché stable ne peut, en concurrence parfaite, reposer sur des contrats qui réalisent des profits positifs et négatifs.

¹¹ Une situation qui ne pourrait perdurer sur un marché où l'entrée est vraiment libre.

Tirons les principales conclusions cette première analyse.

1. L'antisélection sur les marchés d'assurance n'entraîne pas forcément la disparition du marché : les assureurs peuvent offrir des contrats différenciés afin de faire révéler aux assurés l'information qu'ils détiennent sur leur propre niveau de risque. Cette stratégie de réponse consiste plus précisément pour tous les assureurs à proposer un contrat qui couvre l'intégralité du dommage et un autre qui n'en couvre qu'une proportion. Ces contrats sont construits de façon à ce que hauts et bas risques choisissent un contrat différent. L'assureur peut donc facturer un prix actuariel à chacun. L'ensemble de la population est couverte, mais les bas risques n'obtiennent pas une assurance complète.
2. Cependant, l'existence d'un équilibre concurrentiel qui repose sur ce mécanisme séparateur n'est pas garantie si l'on considère que les agents ont un comportement de type Nash. En leur attribuant des comportements d'anticipation plus sophistiqués, il est possible de montrer qu'un équilibre existe en toute circonstance. La nature de cet équilibre dépend toutefois du type d'anticipation qui est postulé.
3. Enfin, l'équilibre qui émerge n'est pas nécessairement optimal de second rang.

2. ANTISELECTION ET INTERVENTION DE L'ETAT

Dans cette section, nous prolongeons l'analyse en cherchant à savoir si l'Etat peut intervenir pour éviter les échecs de marché qui apparaissent lorsque :

- l'équilibre de Rothschild et Stiglitz n'est pas optimal ;
- l'équilibre de RS ne s'instaure pas. Dans ce cas, quand bien même l'équilibre pooling préféré par les bas risques émergerait (hypothèse de Wilson ou Grossman), il n'est pas optimal et l'Etat peut intervenir pour améliorer la situation de tous.

En se replaçant dans le cadre de l'analyse graphique adoptée précédemment, nous chercherons à voir s'il est possible de faire émerger le couple de contrats (H^*, B^*) qui correspond au minimum de subvention nécessaire des bas risques par les hauts risques et qui rétablit l'optimum.

Implicitement, nous faisons l'hypothèse que l'intervention de l'Etat est nécessaire et nous nous plaçons donc dans le cas où la proportion de hauts risques dans la population est faible (inférieure au seuil λ_0 en deçà duquel l'équilibre n'est pas efficient). Choisir une situation de référence dans laquelle les hauts risques sont relativement peu nombreux n'est pas nécessairement réducteur car on sait que les dépenses de santé sont très concentrées sur une faible proportion de la population (Cutler, 1994b, pp.33 pour les Etats-Unis , van Barneveld, et al., 1996, Pays-Bas, Com-Ruelle et Dusmesnil, 1999, France).

Nous articulerons cette présentation autour de celle d'un article récent (Neudeck et Podczeck, 1996) dont l'analyse a pour cadre le marché de l'assurance maladie. Les auteurs étudient l'impact de différentes modalités d'intervention de l'Etat sur le marché de l'assurance. Nous enrichirons cette approche en y intégrant explicitement les solutions proposées par d'autres auteurs et en nous interrogeant à chaque étape sur la nécessité de faire l'hypothèse d'un comportement à la Grossman sur laquelle se fonde le travail de Neudeck et Podczeck afin de généraliser leurs résultats. Les interventions possibles de l'Etat se déclinent selon deux modalités :

- soit l'Etat intervient comme assureur dans le cadre d'un marché pour partie concurrentiel (2.1) ;
- soit il régule la concurrence entre les acteurs privés de ces marchés (2.2).

Enfin, nous dégagerons de cette analyse, les conditions véritablement nécessaires à la gestion de l'antisélection par l'Etat. Contrairement à ce que laissent entendre Neudeck et Podczeck, nous montrerons plus précisément que l'intervention de l'Etat en tant qu'assureur n'est pas indispensable (2.3).

2.1 Intervention de l'Etat comme assureur

Nous étudions deux modalités de participation directe de l'Etat au marché de l'assurance :

1. assurance incomplète de base obligatoire par l'Etat et assurance "complémentaire" offerte par les assureurs privés ;
2. assurance complète non obligatoire offerte et financée par l'Etat et option de sortie (possibilité d'acheter un contrat privé à la place du contrat offert par l'Etat).

Dans cette analyse, nous écartons d'emblée le cas dans lequel l'Etat fournit une assurance complète à tous les individus. Dans ce cas, pour ne pas réaliser de déficit, il doit alors offrir M , assurance complète située sur la droite actuarielle moyenne qui est, par ailleurs, un optimum de premier rang. Par rapport à la problématique définie, ce cas présente un intérêt limité pour deux raisons :

1. en premier lieu, le seul intervenant sur le marché est l'Etat et les compagnies d'assurance privées n'ont pas a priori de rôle à jouer dans une telle configuration ;
2. de plus, cette solution implique une redistribution radicale en faveur des hauts risques et elle n'est pas comparable du point de vue de Pareto avec la solution de marché qui émerge sans intervention de l'Etat.

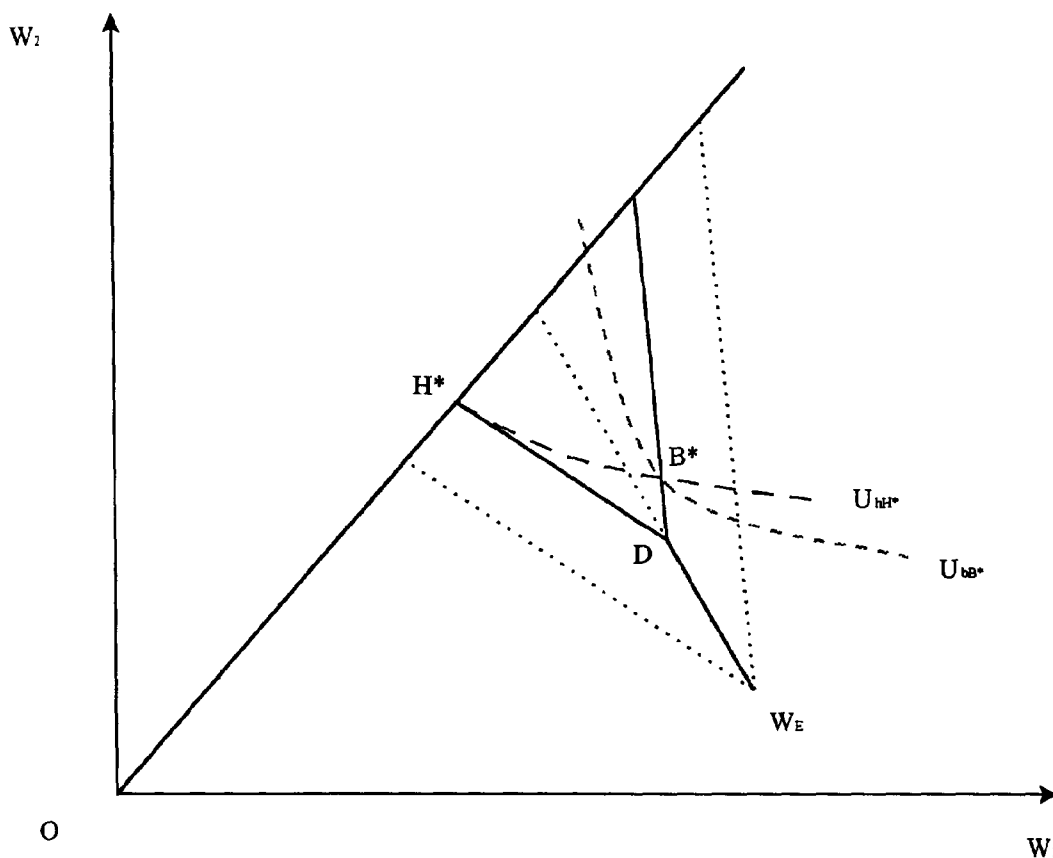
2.1.1 Assurance publique partielle et assurance complémentaire privée

La configuration dans laquelle l'Etat offre une assurance universelle obligatoire partielle et laisse le marché compléter cette offre a été envisagée par Dahlby (1981).

a) Principe

Pour que le contrat d'état soit équilibré, il faut qu'il soit sur la droite actuarielle moyenne de la population. Cette assurance seule, à moins d'être complète, ne peut constituer un optimum. Le résultat est toutefois différent si les assureurs ont la possibilité d'offrir une assurance qui vient compléter l'assurance de base offerte par l'Etat. La solution qui émerge est présentée dans la figure 6 (page suivante).

Figure 6
Assurance publique de base



Si l'assurance publique atteint le niveau D , où D est l'intersection entre la droite actuarielle moyenne et la parallèle à la droite actuarielle des hauts risques qui passe par U_{hH^*} , alors la solution de marché qui émerge est (H^*, B^*) . Tout se passe comme si l'allocation initiale W_E avait été déplacée au point D et que le raisonnement à la RS s'appliquait au nouveau point de dotation des individus. Les deux contrats offerts sont sur les nouvelles droites actuarielles des hauts et bas risques qui partent de D . Les hauts risques obtiennent une assurance complète. Les bas risques souscrivent le contrat d'assurance incomplète B^* , situé sur leur droite actuarielle qui est tel que les hauts risques ne sont pas tentés de le souscrire. L'équilibre qui émerge est donc l'optimum de second rang préféré par les bas risques.

Deux remarques viennent compléter cette analyse.

1. Si le niveau d'assurance publique est plus proche de W_E (autrement dit, s'il couvre une proportion moins élevée du dommage) et si l'équilibre initial est séparable, le nouveau couple de contrat qui émerge est aussi séparable (le contrat obtenu par les bas risques est sur la courbe (MI) que nous avons représentée dans la figure 2, pp. 115); En revanche, l'équilibre n'est toujours pas optimal.
2. Si le niveau d'assurance offert par l'Etat dépasse D (et se rapproche de M), l'équilibre séparable qui s'instaure est aussi un optimum de second rang mais il confère aux hauts risques un niveau d'utilité plus élevé que U_{bH^*} . Ces solutions impliquent une redistribution plus forte en faveur des hauts risques.

b) Robustesse du résultat

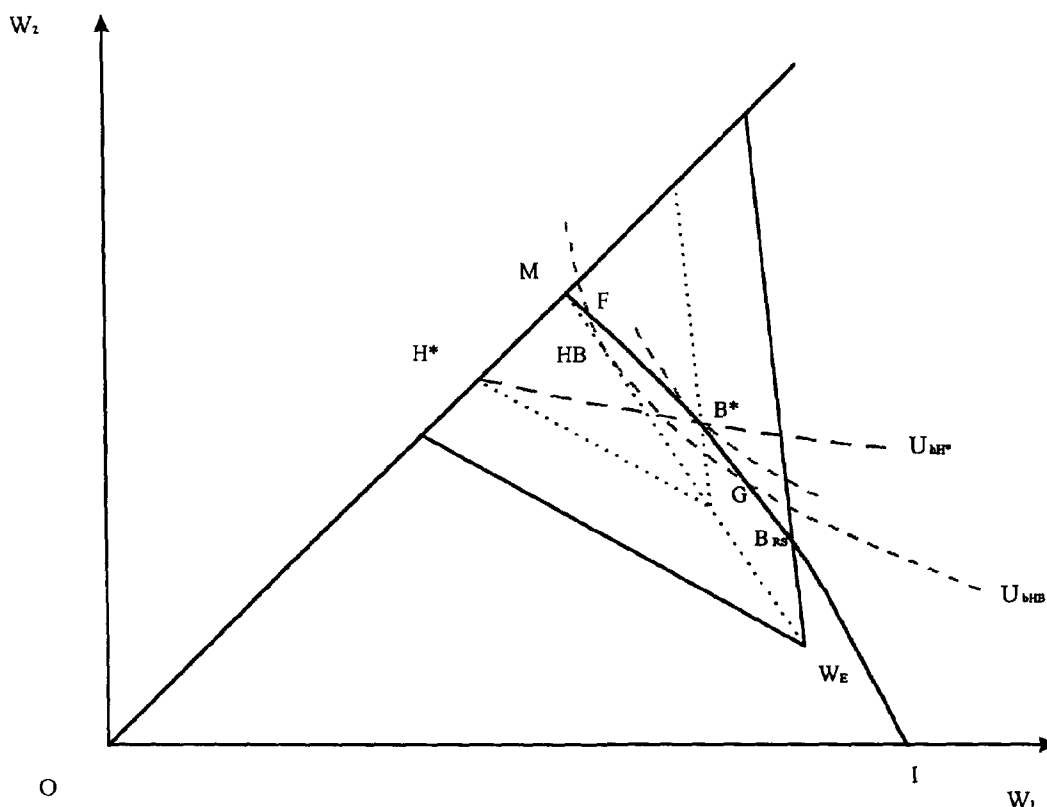
Etudions tout d'abord le cas où la proportion de hauts risques dans la population est trop faible pour que l'équilibre de RS émerge. Dans ce cas, on peut considérer que tous les individus obtiennent, avant intervention de l'Etat, le contrat HB de la droite actuarielle moyenne préféré par les bas risques¹². Cette situation n'est pas optimale. En offrant une assurance de base, l'Etat peut aussi rétablir l'optimum de second rang préféré par les bas risques. Le raisonnement qui sous-tend cette affirmation - et qui n'est pas étudié par Neudeck et Podcizek - est détaillé ici sur la base d'un raisonnement graphique (figure 7, page suivante).

- En HB , il y a tangence entre la droite actuarielle moyenne et la courbe d'indifférence des bas risques U_{bHB} .
- On sait que la courbe (MI) passe par M et B_{RS} (qui, en l'occurrence, sur le marché est initialement dominé par le contrat pooling HB que les bas risques préfèrent).
- Comme les courbes d'indifférences des bas risques sont convexes, la courbe (MI) coupe nécessairement U_{bHB} en deux points F et G . Tous les points de l'arc (FG) confèrent aux bas risques une utilité plus élevée que U_{bHB} . Le point B^* , préféré par les bas risques parmi ceux de la courbe (MI) , fait nécessairement partie de la portion (FG) de (MI) . Si l'Etat offre D , le couple de contrats séparateurs (H^*, B^*) émerge sur le marché. Dans cette nouvelle configuration, l'équilibre initial pooling ne peut plus "menacer" l'équilibre séparable.

¹² Si on reste sur l'hypothèse de RS selon laquelle l'équilibre n'existe pas, le résultat final est identique.

Figure 7

Equilibre pooling et assurance publique de base



L'intervention de l'Etat qui impose une assurance obligatoire de base permet donc d'obtenir le résultat cherché. Ce résultat a été confirmé par Pannequin (1992) qui construit un système d'assurance "à deux étages", semblable à celui que nous venons présenter. Si l'Etat impose la souscription d'un contrat de base, l'équilibre du marché privé existe toujours et l'allocation qui en résulte est optimale. Le choix du niveau d'assurance de base par l'Etat traduit une volonté plus ou moins forte de favoriser explicitement les hauts risques. Dans le cas où la proportion de hauts risques est inférieure au seuil λ_0 , il faut cependant que la couverture de base de l'Etat soit au moins égale à D . Dans ces conditions, Pannequin souligne que "l'équilibre existe en toute circonstance sans qu'il soit nécessaire de prêter des anticipations spécifiques aux compagnies d'assurance". Une fois la dotation initiale "déplacée" au point D , l'allocation de Rothschild et Stiglitz émerge et un comportement de Nash suffit à supporter l'équilibre.

L'obligation d'assurance garantit donc l'accès universel à un minimum d'assurance et permet d'obtenir la redistribution nécessaire à l'instauration d'un optimum (de second rang). Les préférences et les niveaux de risque des individus les conduisent par la suite à des choix différenciés sur le marché privé.

2.1.2 Assurance publique et option de sortie

Le mécanisme que nous étudions ici est plus original. Il consiste pour l'Etat à offrir un contrat d'assurance complète, sans obligation de souscription et à laisser le marché privé compléter cette offre. Nous verrons par la suite qu'un système de ce type existe en Allemagne : les caisses publiques offrent une assurance maladie et les personnes dont le revenu dépasse un certain seuil ont la possibilité de sortir du système public et de souscrire à la place une assurance privée. Ce type de mécanisme est souvent critiqué car on craint que les personnes en bonne santé soient attirées par les assurance privées, et que l'assurance publique ne couvre que les hauts risques.

Neudeck et Podczeck montrent qu'un tel système peut, si certaines conditions de financement sont respectées, permettre d'instaurer un équilibre qui est un optimum de second rang : si l'assurance complète proposée par l'Etat confère aux hauts risques le niveau d'utilité U_{HH^*} , l'équilibre qui émerge est le couple (B^*, H^*) , optimum de second rang préféré par les bas risques.

a) Détermination du montant et de la méthode de financement du déficit public

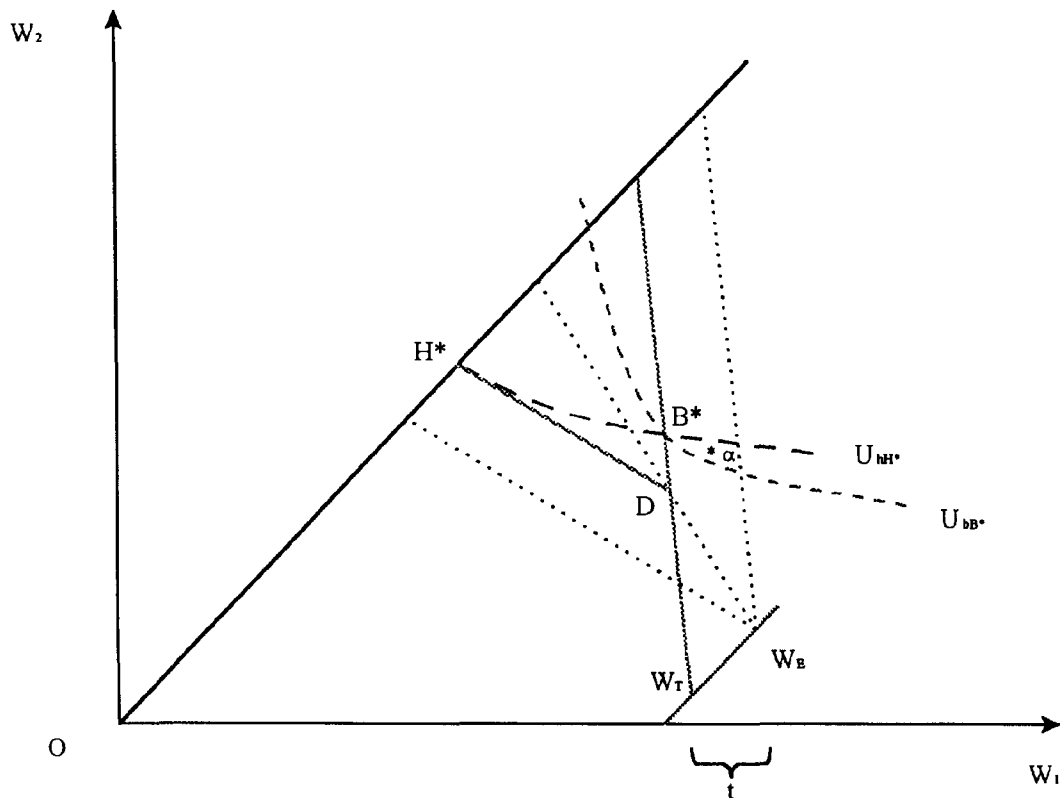
On ajoute l'hypothèse supplémentaire que si le contrat offert par l'Etat est déficitaire, le déficit est intégralement financé par une taxe forfaitaire t payée par l'ensemble de la population.

Si le contrat public est choisi par les hauts risques seulement (et nous montrerons par la suite que c'est le cas), on peut déterminer graphiquement (figure 8, page suivante) le montant de cette taxe.

Remarquons tout d'abord qu'un tel prélèvement conduit à déplacer l'allocation initiale des individus vers un point W_t situé sur une droite de pente 1 passant par W_E puisque la taxe est prélevée ex ante sur toute la population, quel que soit l'état du monde.

La détermination du lieu exact du point W_t sur cette droite repose sur un raisonnement similaire à celui que nous avons suivi pour déterminer les lieux des optima de second rang. En effet, le montant du déficit encouru par haut risque couvert est constant le long de la droite (H^*D) . Ce déficit est financé si les bas risques ont une allocation sur la droite (DB^*) . Le point W_t est défini de façon unique par l'intersection entre cette droite (DB^*) et celle de pente 1 qui passe par W_E .

Figure 8
Contrat public et option de sortie



Notons que si les seuls bas risques ont une allocation ex ante W_t , le déficit moyen par haut risque est financé et les hauts risques peuvent obtenir H^* . Intuitivement, on comprend que les hauts risques n'ont pas besoin de payer effectivement la taxe : qu'ils la paient ou non, ils obtiennent par construction du modèle W_{H^*} ex post. S'ils la paient ex ante, on doit inclure une compensation équivalente ex post pour qu'ils arrivent au niveau de richesse souhaité. Le régulateur a donc le choix entre deux mesures équivalentes sur le plan économique pour financer une telle politique : faire payer la taxe forfaitaire à tous les individus ou faire payer aux seuls bas risques un montant t qui représente un droit de sortie du système public.

b) Comportement des individus et construction de l'équilibre

Montrons que si l'Etat prélève la taxe t et offre le contrat H^* , les assureurs privés offrent B^* et l'équilibre séparateur souhaité émerge.

1. Rappelons que si l'assureur veut pouvoir offrir un contrat spécifique pour les bas risques, il faut que les hauts risques soient indifférents entre le contrat H^* et le contrat des bas risques. Par ailleurs, il faut que ce contrat soit sur la droite actuarielle des bas risques lorsque ceux-ci ont l'allocation initiale après prélèvement de la taxe W_i . Le contrat doit donc, pour réaliser un profit nul, se trouver sur la droite $(W_i D)$. Seul le contrat B^* répond à ces conditions.
2. De façon évidente, les bas risques préfèrent B^* à H^* . Ils ne choisissent donc pas de souscrire le contrat d'assurance sociale.

Montrons qu'aucune autre solution ne peut constituer un équilibre.

Si un contrat est offert qui n'attire que les bas risques (par exemple α), il réalise un profit strictement négatif (puisque'il est à droite de la nouvelle droite actuarielle des bas risques).

On peut aussi montrer qu'aucune offre de contrat qui conduirait les deux catégories de risques à "sortir" du contrat offert par l'Etat ne constitue un équilibre.

1. Si le couple de contrat qui viendrait déstabiliser l'équilibre est séparateur, il faut que le nouveau contrat des hauts risques réalise un profit nul. Or, par construction, H^* , qui est subventionné, est préféré par les hauts risques à tous les contrats tarifés actuariellement sur la base de leur propre risque sur le segment $[W_E H_{RS}]$.
2. Il faudrait alors que le contrat qui incite tous les individus à sortir du contrat offert par l'Etat soit un contrat pooling qui attire les deux catégories de risques. Si ces conditions ne sont pas remplies, le contrat est déficitaire. Or, aucun contrat pooling n'est préféré par les bas risques à B^* . Un contrat pooling n'attirant que les hauts risques serait déficitaire et ne peut donc être offert.

En définitive, il n'existe pas de contrat ou de menu de contrats, qui, s'il était offert, réaliserait un profit¹³: le couple (H^*, B^*) est l'unique équilibre de Nash cette économie.

¹³ On montre aisément que si l'équilibre initial est poolé, le résultat est identique.

Conclusion partielle

L'analyse de cette première série de mesures conduit à penser qu'un partage bien calibré de la couverture du risque maladie entre public et privé peut permettre de résoudre les échecs de marchés liés à la présence d'antisélection telle qu'elle est traditionnellement envisagée dans les modèles d'assurance. Contrairement à ce que Neudeck et Podzeck laissent entendre, ces résultats sont valables lorsque tous les acteurs ont une stratégie de type Nash. Il n'est donc pas nécessaire de postuler un modèle d'anticipation de type Grossman.

Toutefois, on peut objecter que les avantages potentiels de la concurrence en terme d'efficacité sont limités par ce type de solution dans la mesure où une partie des contrats est prise en charge par l'Etat réputé moins efficace. Nous nous tournons désormais vers des solutions dans lesquelles l'Etat se contente d'imposer un certain nombre de règles de fonctionnement à l'assurance sans intervenir lui même comme assureur.

2.2 Réglementation du marché de l'assurance par l'Etat

Nous nous interrogeons maintenant pour savoir si l'Etat pourrait se contenter d'imposer des règles aux assurances sans intervenir directement sur le marché afin de rétablir l'équilibre ou l'optimum.

Plus précisément, on étudie deux cas.

1. Dans le premier, les assureurs doivent offrir un minimum d'assurance à tous, sans contrainte de tarification.
2. Dans le second, obligation est faite aux assureurs d'offrir des contrats standards (assurance complète ou assurance minimale dont la tarification est décidée par l'Etat), sans qu'il soit possible de rejeter la candidature de personnes qui choisissent ce contrat, quel que soit leur niveau de risque.

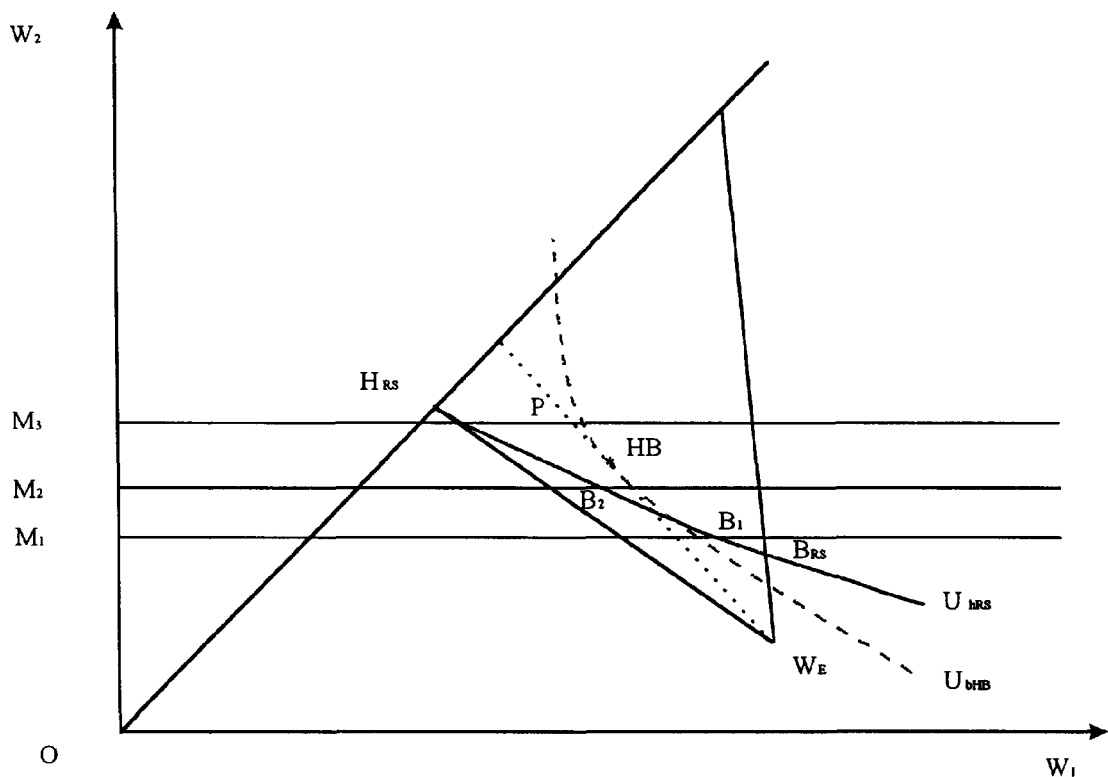
2.2.1 Assurance minimale

Le premier type d'intervention est motivé par l'idée que tout individu doit obtenir une couverture qui lui permette de couvrir un niveau minimum du dommage quel que soit son niveau de risque. L'Etat impose donc un niveau de couverture minimal, obligeant les assureurs à l'offrir et les individus à le souscrire. Etant donnée la configuration du

modèle, ceci revient à dire qu'en cas de dommage, les assurés ne peuvent avoir un niveau de richesse inférieur à un certain seuil.

La solution qui émerge dépend du niveau d'assurance minimal qui est imposé. L'ensemble des cas possibles est représenté dans la figure 9 dans laquelle le point HB est l'équilibre pooling préféré par les bas risques. Dans le cas représenté, l'équilibre avant intervention de l'Etat est (H_{RS}, B_{RS}) .

Figure 9
Couverture minimale universelle



1. Si l'assurance minimale imposée est inférieure à celle qu'obtiennent les bas risques dans l'équilibre initial, la mesure n'a évidemment aucun impact.
2. Si l'assurance minimale est M_1 , les bas risques doivent obtenir plus d'assurance qu'en B_{RS} . Un équilibre séparateur (B_1, H_{RS}) s'instaure, où B_1 donne aux bas risques le niveau minimal d'assurance exigé. Ce contrat est en outre tel que les hauts risques sont indifférents entre B_1 et H_{RS} . Toutes les entreprises réalisent un profit positif puisque B_1 est à gauche de la droite actuarielle des bas risques.

Aucune autre offre profitable ne peut la déstabiliser puisque :

- les bas risques préfèrent B_1 à tout contrat pooling ;
 - un contrat qui attirerait les seuls bas risques devrait offrir un niveau de richesse dans l'état 2 inférieur à M_1 , ce qui est interdit.
3. Si l'assurance minimale est égale à M_2 , le contrat B_2 , qui donne une assurance minimale aux bas risques et qui est tel que les hauts risques sont indifférents entre ce contrat et H_{RS} , ne peut pas faire partie de l'offre d'équilibre. En effet, les bas risques lui préfèrent le contrat HB qui, s'il est proposé, est choisi par tous. Si on reste dans le cadre d'un comportement Nash, cet équilibre ne peut émerger. En admettant des anticipations de type Grossman ou Wilson, on peut à l'inverse montrer que ce contrat HB ne peut être déstabilisé.
4. Si le niveau minimal d'assurance est tel que le point HB ne satisfait plus la contrainte d'assurance minimale (par exemple M_3), seul un contrat pooling peut émerger qui donne à tous le niveau d'assurance minimum (point P). On montre alors aisément que ce contrat ne peut être déstabilisé¹⁴.

Un équilibre stable semble donc pouvoir émerger dans la majorité des cas¹⁵, voire dans tous les cas en fonction des hypothèses posées sur le comportement des agents. La raison principale est que les firmes éprouvent plus de difficulté à déstabiliser l'équilibre en attirant les bas risques car cela les conduirait à offrir des contrats qui ne respectent pas la réglementation. Ceci dit, les contrats qui émergent lorsque ce type de réglementation est en place ne peuvent jamais être optimaux de second rang, et un des objectifs de l'Etat n'est pas atteint.

2.2.2 Définition par l'Etat de contrats standards

Une des solutions envisagées pour limiter l'antisélection sur les marchés d'assurance est d'obliger les assureurs à offrir un contrat standard, dont le prix et le taux de couverture seraient spécifiés préalablement et d'accepter en assurance toute personne qui demanderait à le souscrire, quel que soit son niveau de risque.

¹⁴ Aucun autre contrat attirant les bas risques et respectant l'obligation légale n'existe.

¹⁵ Le cas où il n'y a initialement pas d'équilibre de RS donne des résultats qualitatifs de même nature et nous ne les présentons pas en détail

Dans cette configuration, on ajoute l'hypothèse que les hauts risques font confiance à l'Etat pour faire respecter cette réglementation et souscrivent le contrat standard même s'ils anticipent qu'il sera déficitaire. L'hypothèse de dissimulation inhérente au modèle de Grossman ne s'applique donc pas aux contrats que l'Etat impose aux assureurs d'offrir.

a) Contrat standard d'assurance complète

Dans la mesure où l'objectif de l'Etat est de rétablir l'optimum et de garantir l'existence d'un équilibre, on postule que le contrat complet d'assurance standard proposé est situé entre H^* et M . En effet, s'il était au delà de M sur la première bissectrice, il ne pourrait jamais être profitable et s'il était en deçà de H^* il ne pourrait faire partie d'une offre Pareto optimale - de second rang - pour les hauts risques. Neudeck et Podczeck étudient le cas où H^* doit être offert par tout assureur. Ils montrent que l'équilibre au sens de Grossman ne peut s'instaurer dans cette configuration.

Avant de détailler le raisonnement, notons au préalable que puisque H^* est obligatoirement offert, les hauts risques doivent obtenir au moins U_{hH^*} .

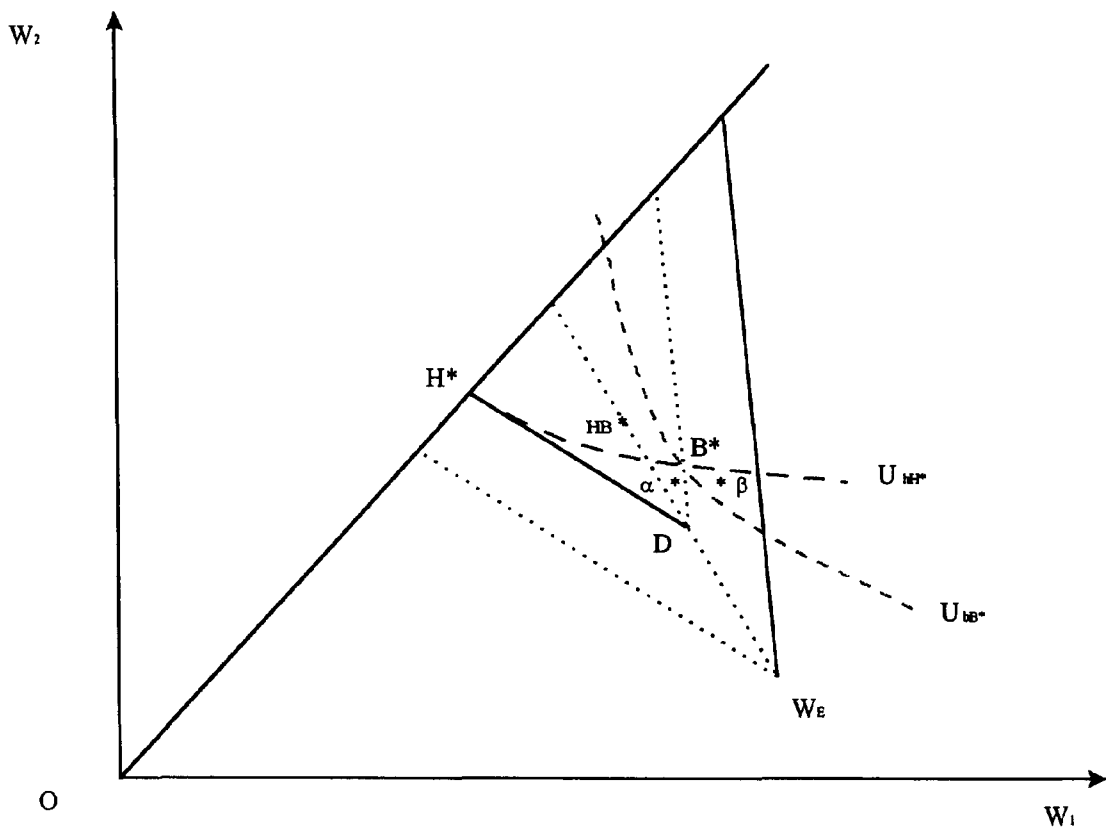
Présentons les étapes de l'argumentation sur la base de la figure 10 (page suivante).

Un équilibre pooling ne peut émerger

Si l'équilibre était constitué d'un contrat pooling HB (parmi ceux préférés à H^* par les hauts risques), un assureur entrant sur le marché pourrait le déstabiliser et réaliser un profit.

Plus précisément, étudions le cas où le nouvel entrant offre le contrat H^* et un contrat α qui est préféré à tout contrat pooling par les bas risques, sans être préféré par les hauts risques au contrat HB . Comme les hauts risques anticipent que l'Etat peut obliger les entreprises à fournir H^* , ils ne cherchent pas à dissimuler leur risque en souscrivant α . Ce contrat n'est donc souscrit que par les bas risques. Le contrat α étant à gauche de la droite actuarielle des bas risques passant par D , il réalise un profit positif, y compris si tous les bas risques souscrivent H^* , puisqu'il suffirait qu'il soit sur (DB^*) pour réaliser un profit nul en accueillant tous les hauts risques. Le contrat pooling ne peut donc être un équilibre.

Figure 10
Contrat standard privé d'assurance complète



L'équilibre séparateur est aussi contestable

Seul un contrat séparateur pourrait constituer un équilibre. Dans ce cas, les hauts risques obtenant H^* , il est naturel d'envisager que les bas risques obtiennent B^* et que chaque firme réalise un profit espéré nul. Cependant, une firme entrant sur ce marché et offrant H^* et β peut déstabiliser cette offre. Le contrat β est construit de façon à apporter un profit positif s'il est choisi par les bas risques. Il est préféré par ces derniers à B^* alors que les hauts risques préfèrent H^* à β . La firme qui offre (H^*, β) attire donc tous les bas risques. Si on fait l'hypothèse supplémentaire que les hauts risques sont ex ante également répartis entre l'ensemble des firmes qui offrent H^* , la firme offrant (H^*, β) réalise un profit globalement positif (le profit sur tous les bas risques compense la perte sur une fraction des hauts risques).

Neudeck et Podczeck concluent alors que (H^*, B^*) n'est pas un équilibre stable puisqu'il existe une possibilité d'écramer les bas risques et de réaliser un profit positif.

La proposition d'Henriet et Rochet

Henriet et Rochet (1987b) proposent un mécanisme de régulation des entreprises d'assurance qui s'approche de celui que nous venons de décrire et permettrait d'établir un équilibre efficient. Dans la configuration qu'ils proposent, l'Etat oblige les assureurs à offrir un contrat d'assurance complète et il impose une prime maximale pour ce contrat. Comme ce contrat est de fait souscrit par les haut risques, un parallèle peut être établi avec la proposition de Neudeck et Podczeck. La prime maximale que peuvent faire payer les assureurs aux hauts risques dans le cas où $\lambda < \lambda_0$ doit être telle que ces derniers obtiennent H^* ¹⁶. Dans ce cas, les assureurs réalisent une perte qui est compensée par le profit réalisé sur les bas risques. Henriet et Rochet insistent sur le fait que l'agence régulatrice doit avoir les moyens de faire respecter l'obligation d'assurance (les assureurs ne peuvent refuser de couvrir un haut risque). Ils postulent en outre implicitement que les hauts risques sont répartis de façon uniforme dans les différentes compagnies d'assurance. Si les assureurs ont un comportement de Nash, hypothèse explicitement posé par les auteurs, le couple (H^*, β) peut théoriquement déstabiliser cet équilibre comme nous l'avons montré dans le cas précédent. L'équilibre optimal n'est donc pas stable, si l'on s'en tient à un comportement strict de Nash.

b) Contrat standard d'assurance partielle

Etudions maintenant le cas dans lequel le contrat standard que toute firme doit offrir est un contrat d'assurance partielle tarifé au prix actuariel moyen (point D , pp. 142). La question est de savoir si ce mécanisme permet de dupliquer la situation où l'assurance partielle obligatoire est fournie par l'Etat.

Si les assurances ont le droit d'offrir d'autres contrats que le contrat standard, le seul candidat à l'équilibre est le couple (H^*, B^*) qui sera offert par les assurances en plus de D . Les trois contrats D , H^* et B^* réalisent un profit nul.

¹⁶ C'est le seuil m^* évoqué page 239 dans Henriet et Rochet 1987b. Si la prime maximale est plus faible, la subvention des bas vers les hauts risques implique une redistribution supérieure à celle qui est strictement nécessaire au rétablissement de l'équilibre.

Sur la base des mêmes arguments que précédemment, on voit que cet équilibre ne peut s'instaurer durablement. Un entrant peut offrir le contrat β (en plus du contrat D). Il attire tous les bas risques (qui délaissent B^* et D) et réalise un profit positif puisqu'il n'a en portefeuille qu'une fraction des hauts risques sur le contrat standard.

c) Analyse critique de ces résultats

Conclusion du raisonnement précédent

L'idée selon laquelle imposer aux assureurs de fournir un contrat standard permet de résoudre les problèmes d'efficience est rejetée par Neudeck et Podczeck. Par rapport aux situations étudiées précédemment où l'Etat offre lui-même un contrat standard et impose les subventions croisées entre les hauts et les bas risques, une réglementation qui tente d'imposer les subventions croisées aux firmes est vouée à l'échec. Si on fait l'hypothèse que les hauts risques sont également répartis entre les différentes firmes, un entrant respectant la lettre de la réglementation peut offrir un nouveau couple de contrat qui déstabilise les candidats à l'équilibre. Implicitement, Neudeck et Podczeck abandonnent les hypothèses de type Grossman pour un raisonnement de type Nash. L'équilibre est impossible puisque les firmes sont tentées de se concurrencer en attirant l'ensemble des bas risques par l'introduction de contrats qu'ils souscriront massivement.

Remise en cause de la robustesse de cette conclusion

Le raisonnement selon lequel le contrat β peut déstabiliser le couple (B^*, H^*) dépend cependant des hypothèses comportementales qui sont posées. Par exemple, dans le cas où l'Etat impose d'offrir un contrat standard d'assurance complète, une fois que le nouvel entrant propose β , tous les assureurs réalisent une perte sur leur contrat H^* , puisque les bas risques ont massivement souscrit β . Il est donc peu vraisemblable qu'ils restent sur le marché. Une fois qu'ils se sont retirés, le couple (H^*, β) devient à son tour déficitaire : pour que la perte liée au contrat H^* soit compensée par les gains réalisés sur le contrat β , il faudrait que ce dernier soit sur la droite (DB^*) . Dans ces conditions, on voit que la menace que représente (H^*, β) n'est pas crédible à moyen terme et qu'un entrant n'a pas intérêt à dévier. Le contrat séparateur à subventions croisées optimal de second rang (H^*, B^*) émerge comme équilibre. Le même type de raisonnement tient dans le cas d'un contrat standard d'assurance partielle.

Discussion

Le raisonnement que nous venons de suivre prolonge celui de Neudeck et Podczeck qui s'arrêtait implicitement à la conclusion que les firmes restent sur le marché alors même qu'elles réalisent un profit négatif. En fait, nous introduisons une hypothèse supplémentaire d'anticipation à la Wilson : *il n'existe pas de contrat en dehors de l'équilibre qui, offert, réalise un profit positif même lorsque les premiers contrats existant sur le marché sont retirés*. Récemment, ce résultat a été confirmé par Feldman, Escribano et Pellisé (1998). Les auteurs étudient le cas où l'Etat impose à tout assureur d'offrir le contrat D et sur la base d'anticipations à la Wilson et ils montrent que l'équilibre optimal est atteint. Ils soulignent en outre que si l'équilibre de RS est optimal, autrement dit, si B^* est, soit confondu avec B_{RS} , soit à droite de B_{RS} (la proportion de hauts risques est assez élevée), l'intervention de l'Etat peut s'avérer néfaste. En effet, dans ce cas, l'Etat impose aux assureurs d'offrir un contrat sur la droite actuarielle moyenne qui attirera les hauts risques dans la mesure où il leur confère une utilité supérieure à celle qu'ils obtiennent en H_{RS} . En revanche, les bas risques n'accepteront pas de souscrire ce contrat standard car il ne pourra jamais être combiné à un contrat privé de façon à ce qu'ils obtiennent une utilité supérieure à celle de B_{RS} . Autrement dit, les assureurs obligés d'offrir un contrat subventionné pour les hauts risques n'auront pas les moyens de compenser la perte qu'il génère et n'auront d'autre solution que de se retirer du marché.

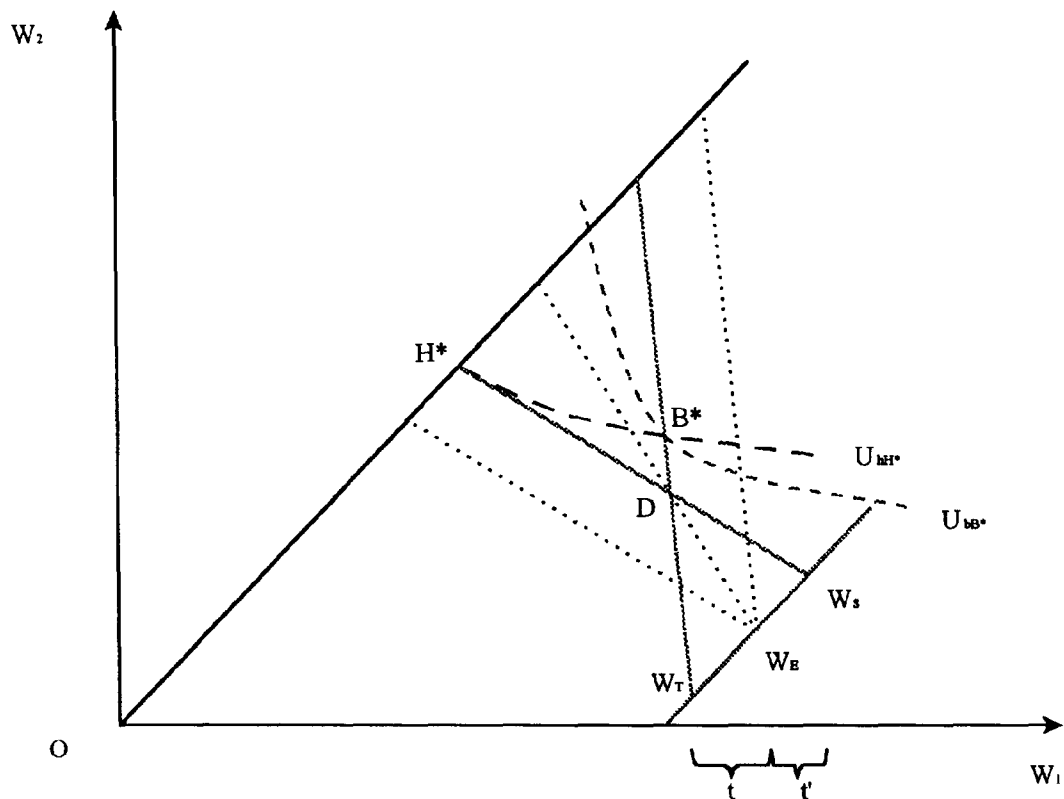
Dans le cas où le couple (H^*, B^*) représenté dans la figure 10 est candidat à l'équilibre, il n'en reste pas moins que, selon les hypothèses que l'on pose sur les anticipations des acteurs (Nash ou Wilson), les conclusions sont inverses. Sans formuler de jugement définitif quant à la plus ou moins grande plausibilité de la réaction des acteurs du marché, on peut admettre que même si l'introduction d'une réglementation de ce type peut établir un optimum de second rang à long terme (le temps que les firmes, en répétant le jeu, comprennent qu'elles n'ont rien à gagner en introduisant des contrats de type β), elle risque de déstabiliser dans un premier temps le marché. En fait, le problème est similaire à celui que nous avons évoqué dans le cas de l'équilibre de Miyasaki, à cela près que les firmes sont ici contraintes d'offrir les contrats déficitaires. La stabilité du marché dépend de leur auto-discipline et de leur résistance à la tentation d'une stratégie gagnante à court terme. Garantir un contrat aux hauts risques ne suffit pas à établir l'optimum, encore faut-il s'assurer que les assureurs ne vont pas se faire concurrence pour attirer plus de

hauts risques que le nombre nécessaire à la compensation du déficit lié au contrat standard.

d) Le mécanisme de Crocker et Snow

Crocker et Snow (1985) proposent une dernière solution de même nature mais qui permet de surmonter cette difficulté.

Figure 11
Contrat standard privé et taxe de Crocker et Snow



Elle consiste à taxer tout contrat qui est tarifé au-dessous du prix actuariel moyen et à subventionner tout contrat tarifé au-delà. L'ensemble des subventions doit être financé par les taxes forfaitaires prélevées. Si l'Etat (figure 11) introduit une subvention d'un montant t' pour les contrats tarifés en deçà du prix moyen et une taxe forfaitaire d'un montant t pour les autres, l'équilibre de Nash (H^* , B^*) qui s'instaure est efficient. Contrairement aux solutions évoquées précédemment, l'équilibre de chaque firme ne dépend pas du nombre exact de hauts et de bas risques qu'elle couvre. Taxes et subventions sont calculées de façon à ce que chaque contrat rapporte un profit net nul.

Un assureur qui couvrirait relativement plus de hauts (bas) risques serait plus subventionné (taxé) que les autres. Tout ce passe comme si les allocations initiales des agents étaient modifiées : les nouvelles contraintes de profit net nul passent par H^* et B^* . Le contrat β attire toujours l'ensemble des bas risques mais réalise un profit net après impôt négatif et ne peut déstabiliser l'équilibre de Nash. Notons au passage qu'aucun contrat pooling (non taxé) ne pourrait attirer les bas risques qui lui préféreront toujours B^* . Un autre avantage de ce mécanisme réside dans le fait que l'Etat n'a pas besoin de connaître ex ante le niveau de risque des individus.

2.3 Conditions nécessaires au succès de l'intervention de l'Etat: une proposition

Si l'on cherche à résumer l'ensemble des résultats de cette seconde section, il semble, en première analyse, que parmi les solutions que nous avons envisagées :

- celles dans lesquelles l'Etat intervient en tant qu'assureur, en imposant un contrat de base ou en offrant un contrat pour les hauts risques, permettent de garantir l'existence d'un équilibre et de rétablir l'optimum ;
- celles dans lesquelles l'Etat cherche à imposer aux assureurs en concurrence d'offrir des contrats déficitaires ou de garantir un minimum d'assurance, ne permettent pas d'atteindre l'objectif choisi. En particulier, lorsque les assureurs conservent la possibilité de réaliser des profits de court terme en attirant les bas risques, l'analyse théorique montre que la réglementation échoue.

L'Etat assureur semble donc avoir plus de chances d'atteindre l'objectif souhaité que l'Etat régulateur, ce qui est la conclusion de Neudeck et Podczech. Néanmoins, cette analyse est trop simplifiée et en rapprochant quelques-uns des cas présentés (tableau 2, page suivante), nous pouvons dégager les conditions théoriques qui garantissent le succès de l'intervention de l'Etat en dehors de toute hypothèse comportementale sophistiquée.

Tableau 2 : Synthèse des principales analyses

	<i>Offre de l'Etat</i>	<i>Offre des assureurs</i>	<i>Réglementation complémentaire</i>	<i>Résultat en comportement Nash</i>
1. Dahlby-Pannequin (2.1.1)	D	(H*, B*)	Obligation de souscrire D	Equilibre optimal atteint
2. Neudeck et Podczeck (2.2.2 cas b)	-	D imposé par l'Etat (H*, B*)	Pas d'obligation de souscrire D	L'équilibre n'est pas stable
3. Neudeck et Podczeck (2.1.2)	H*	B*	Pas d'obligation de souscrire H* Taxe sur B*	Equilibre optimal atteint
4. Crocker et Snow (2.2.2 cas e)	-	H* imposé par l'Etat B* spontané	Taxe sur B* Subvention de H*	Equilibre optimal atteint
5. Neudeck et Podczeck (2.2.2 cas a)	-	H* imposé par l'Etat B* non stable	Pas d'obligation de souscrire H*	L'équilibre n'est pas stable

En analysant ce tableau, on note que les interventions efficaces de l'Etat sont en fait celles qui aboutissent à ce que l'allocation des bas risques, associée à la souscription d'une assurance privée, se situe sur la droite actuarielle fondée sur leur probabilité de dommage π_b qui passe par le point D , autrement dit, la droite (DB^*) .

Dans le cas 1, cette condition est remplie parce qu'on leur impose la souscription du contrat d'Etat D .

Dans les cas 3 et 4, s'ils ne souscrivent pas le contrat H^* , ce qu'ils n'ont pas intérêt à faire, ils paient une taxe qui les positionne sur la droite (DB^*) . Cette taxe est, dans le cas 3, un droit de sortie de l'assurance publique et dans le cas 4, une surprime directement associée au fait de souscrire B^* . Le résultat est qu'aucun assureur ne peut déstabiliser l'équilibre. Pour ce faire, il devrait offrir aux bas risques un contrat à droite de (DB^*) nécessairement déficitaire.

On peut alors pousser le raisonnement plus loin. Dans le cas 2 (figure 10, pp. 142), l'équilibre n'est pas stable parce que les assureurs peuvent attirer les bas risques en leur proposant β . Si l'Etat impose à tous la souscription de D au prix actuariel moyen et laisse à chacun la possibilité de compléter la couverture en souscrivant un second contrat actuariel, l'équilibre constitué par D , B^* et H^* est rétabli. En effet, quand les bas risques ont souscrit D , un assureur qui leur offrirait un contrat β réaliserait un profit négatif sur ce contrat. A contrario, on peut déduire que dans le cas 1, ce qui importe n'est pas le fait

que ce soit l'Etat qui offre le contrat *D*, mais le fait que chacun (les bas risques en particulier) soit obligé de payer une garantie de base au taux actuariel moyen.

Pour ce qui concerne les cas 3, 4 et 5, ce qui importe est le fait que tout contrat *H** sera subventionné et tout contrat *B** sera taxé. Dans ce cas, Crocker et Snow avaient poussé le raisonnement jusqu'au bout en montrant qu'il suffisait à l'Etat d'imposer la subvention croisée (cas 4) et que l'Etat n'était pas obligé d'intervenir comme assureur (comme dans le cas 2).

En dernière analyse, on voit donc que la condition véritablement nécessaire au rétablissement d'un équilibre optimal (de second rang), lorsque celui-ci n'émerge pas spontanément sur le marché, est que l'Etat mette en œuvre un mécanisme de subvention croisée qui ne puisse être contourné par les assureurs ou les assurés. Deux possibilités s'offrent à lui :

- soit il impose l'offre par les assureurs et la souscription par les assurés du contrat *D* dont il définit le niveau et le prix. A ce premier contrat peuvent s'ajouter d'autres contrats, en l'espèce, *B** et *H**, souscrits respectivement par les bas et les hauts risques ;
- soit il impose que le contrat *H** soit offert. Dans ce cas, il doit mettre en œuvre un mécanisme de subvention croisée tel que, pour chaque personne couverte par *H**, une subvention soit versée à l'assureur qui la couvre et pour chaque personne couverte par *B**, une taxe soit prélevée et si nécessaire transférée à d'autres assureurs en fonction du nombre de hauts risques couverts.

L'Etat peut gérer et offrir directement *D* ou *H**, mais ce n'est pas une condition nécessaire.

CONCLUSION

A l'issu de ce chapitre, on peut conclure que si l'antisélection menace a priori le fonctionnement des marchés d'assurance, les assureurs peuvent, dans certaines circonstances, surmonter la difficulté que pose l'asymétrie d'information. En offrant des contrats qui se différencient par la quantité d'assurance, ils amènent les assurés à révéler leur niveau de risque et ils peuvent tarifer des primes actuarielles.

Lorsque la proportion de hauts risques dans la population est faible, cette solution n'est pas optimale de second rang et, si elle diminue encore, l'existence d'un équilibre n'est pas garantie. Dans ce dernier cas, si tous les acteurs du marché adoptent des comportements de type Nash, les assureurs seront contraints de se retirer du marché.

Dans le cadre du modèle standard qui formalise l'existence d'antisélection en assurance, nous avons montré que l'Etat peut se donner les moyens de garantir l'existence de l'équilibre et son optimalité lorsque ces conditions ne sont pas spontanément réunies. Nous avons en outre établi les conditions minimales nécessaires pour que cette intervention soit efficace, dans le cas où les acteurs ont des anticipations de type Nash. Ceci nous a notamment permis, d'une part, de proposer un critère simple d'analyse des différents modèles qui se penchent sur cette question et d'autre part, de souligner que l'Etat n'est pas tenu d'intervenir en tant qu'assureur sur le marché.

La gestion de l'antisélection dans un univers concurrentiel reste cependant essentielle et peut passer par une offre publique d'assurance.

Si cette réflexion théorique doit orienter l'intervention de l'Etat sur le marché de l'assurance maladie, encore faut-il que l'antisélection soit constituée sur ce marché. Nous étudions cette hypothèse dans le chapitre suivant.

CHAPITRE 5

ASSURANCE SANTE ET ANTISELECTION : APPROCHES EMPIRIQUES

L'existence d'une asymétrie d'information en faveur de l'assuré peut théoriquement mettre à mal l'équilibre des marchés d'assurance. En effet, si les assureurs ne parviennent pas à offrir un contrat complet actuariel aux personnes qui ont des probabilités de dommages différentes, ils proposent un prix moyen. Dans ce cas, les bas risques peuvent choisir de ne pas s'assurer. Nous avons toutefois démontré dans le chapitre précédent que cette difficulté peut théoriquement être surmontée, notamment grâce à une intervention de l'Etat.

Dans ce chapitre, nous prolongeons l'analyse en cherchant à déterminer si les prémisses et les conclusions du modèle standard concernant l'antisélection peuvent être confortées par des éléments empiriques qui se rapportent au marché de l'assurance santé. Autrement dit, nous cherchons à savoir si l'antisélection est constituée sur ce marché.

Nous présentons les résultats de travaux qui, dans la lignée du modèle canonique, montrent que les individus qui présentent un niveau de risque faible tendent à s'assurer moins que les autres (Section 1). Au terme d'une analyse critique de ces résultats, nous nous tournons vers l'étude d'éléments qui nous semblent plus clairement relever d'une analyse des prédictions du modèle (Section 2). Enfin, nous complétons l'analyse de la littérature existante par les résultats d'une recherche qui illustre, au travers de l'étude du fonctionnement du marché de l'assurance privée aux Pays-Bas, les résultats théoriques dégagés dans le chapitre précédant concernant l'intervention de l'Etat pour garantir l'existence et l'optimalité de l'équilibre¹ (Section 3).

¹ Cette analyse reprend assez largement, en le complétant, un travail précédemment mené avec Lise Rochaix (Couffignal et Rochaix, 1998).

1. SEGMENTATION DES RISQUES

Pour Rothschild et Stiglitz, l'assureur est en mesure de surmonter l'asymétrie d'information s'il propose des contrats différenciés vers lesquels les individus qui ont des niveaux de risque différents vont se diriger. Les bas risques se reporteront notamment sur des contrats moins couvrants que ceux qu'ils souhaiteraient acquérir pour donner aux assureur un signal sur leur niveau de risque. On considère généralement que : "*adverse selection is the tendency of high risks to be more likely to buy insurance or to buy larger amounts of insurance*"² (Bakker et Van Vliet, 1993). On peut trouver dans la littérature deux types d'études qui cherchent à valider cette hypothèse. Les premières se fondent sur des données d'enquêtes et les secondes sur l'observation "d'expériences naturelles".

1.1 Etudes sur données d'enquête

De nombreux travaux empiriques basés sur des données d'enquêtes transversales en population générale comparent des personnes qui ont des niveaux de couverture différents. Elles doivent s'affranchir d'une contrainte technique liée à la prise en compte explicite de phénomènes éventuels de risque moral.

Nous avons précédemment montré que l'état de santé n'est pas le seul déterminant de la consommation de soins. Pour autant, cette dernière reflète largement le niveau de risque. L'observation à une date t d'une corrélation positive entre consommation et niveau d'assurance peut trouver deux explications :

- d'une part, toutes choses étant égales par ailleurs, si une personne mieux couverte consomme plus qu'une personne non (ou moins bien) couverte, il peut s'agir d'un effet prix qualifié de risque moral : la consommation est plus élevée parce que le coût marginal des soins est plus faible ;
- d'autre part, si, à la période précédant l'observation du fait stylisé, une personne connaissait son niveau de risque et anticipait des dépenses élevées, elle peut avoir décidé de se couvrir ou d'améliorer sa couverture. Dans ce cas, la relation observée

² L'antisélection est la tendance des hauts risques à avoir plus tendance à acheter de l'assurance ou à acheter plus d'assurance.

entre niveau d'assurance et consommation traduit la réalisation de ces anticipations. "*Dans cette hypothèse, ce n'est pas le niveau d'assurance qui explique les dépenses plus importantes mais bien l'inverse*" (Caussat et Glaude, 1993, pp.5) et il y a antisélection.

Parmi les études qui visent à détecter la présence d'antisélection en prenant en compte le risque moral, Price et Mays (1985) adoptent une démarche intéressante qui couple des données exogènes sur l'impact du risque moral et des informations administratives recueillies auprès d'assureurs. Ils étudient la couverture des employés fédéraux qui ont depuis longtemps le choix entre différents contrats d'assurance (plans). En normant les coûts moyens observés par plan pour prendre en compte les différences de taux de couverture et donc les effets du risque moral³, ils estiment le coût moyen de prise en charge de l'ensemble de la population des employés fédéraux, si celle-ci bénéficiait d'une couverture complète. Dans un second temps, pour chaque plan, ils estiment quelle devrait être la prime, étant donné le taux de coassurance, si toute la population le souscrivait. Enfin, ils comparent cette prime avec celle qui est effectivement payée par les assurés. Si la prime effective est plus élevée (faible) que ce qu'elle serait si elle était souscrite par toute la population, le contrat considéré a été souscrit par des personnes qui présentent un niveau de risque élevé (faible). Les auteurs montrent que, par rapport à la prime qui devrait prévaloir si chaque contrat couvrait un échantillon représentatif de la population, la prime du contrat le plus généreux est de 21% plus élevée, et celle du contrat le moins couvrant est de 29% plus faible. Autrement dit, on constate ex post que le mécanisme séparateur de différenciation des contrats fonctionne.

De nombreuses études portent exclusivement sur des données d'enquête dans lesquelles on dispose d'information sur l'état de santé, la couverture assurantielle et la consommation de soins (Cartwright, Hu et Huang, 1992, Caussat et Claude 1993, Genier, 1998, par exemple). Elles s'attachent à estimer ce qui relève de la part respective du risque moral et de l'antisélection et déterminent simultanément :

- d'une part, la couverture que les personnes obtiennent. La variable expliquée est une probabilité d'obtenir une couverture ou un niveau plus ou moins élevé de

³ Sur la base des estimations de la Rand

couverture. Cette équation de demande d'assurance est conditionnée, non seulement par les variables socio-démographiques, mais aussi par l'état de santé ;

- d'autre part, le comportement de recours aux soins, qui est notamment fonction de la couverture assurantielle et de l'état de santé.

L'estimation du modèle simultané est relativement complexe (Caussat et Glaude, p.40). Toutefois, s'il s'avère que les résidus des équations sont indépendants, l'estimation globale du modèle peut être effectuée équation par équation. Ces études économétriques permettent de dégager un certain nombre de résultats sur la demande d'assurance.

1. L'article de Cartwright, Hu et Huang (1989) montre que les personnes qui se déclarent en mauvaise santé ont une probabilité plus faible d'avoir une couverture complémentaire, contrairement à l'hypothèse selon laquelle si on anticipe d'avoir un niveau de risque élevé, on tend à s'assurer plus, toutes choses étant égales par ailleurs. Par contre, parmi les personnes couvertes, moins l'état de santé est bon, plus la couverture est élevée. Cet apparent paradoxe peut s'expliquer par l'existence de sélection par les assureurs (nous y reviendrons par la suite) qui dissuaderaient les plus hauts risques de souscrire une assurance. Une fois ce premier tri effectué, les personnes qui ont un risque élevé achètent plus d'assurance.
2. L'étude de Caussat et Glaude (1993) assez proche dans son esprit de la précédente, a été effectuée sur des données françaises. Elle cherche à déceler des différences de comportement entre les personnes qui ont une assurance complémentaire et les autres. Il apparaît très nettement que le revenu est une variable centrale dans l'accès à l'assurance. Les résultats sur l'état de santé sont globalement difficiles à interpréter puisque les personnes jeunes sont plus susceptibles d'être couvertes (effet de génération) ainsi que celles qui présentent un degré d'invalidité plus élevé. En revanche, le risque vital a plutôt une influence négative sur la probabilité d'être couvert.

3. L'étude de Genier (1998) ne dégage pas non plus de relation très nette entre l'état de santé et le fait d'être ou non assuré. En revanche, contrairement aux deux cas précédents, les données ne permettent pas de rejeter l'hypothèse de non indépendance des équations de demande d'assurance et de demande de soins, rejet qui serait nécessaire pour mesurer équation par équation le risque moral et l'antisélection.

Ettner (1997) adopte une démarche assez originale pour tester la présence d'antisélection. Tout d'abord, elle postule que les personnes qui ont une couverture acquise sur le marché individuel ont de plus grandes chances d'avoir adapté leur niveau de couverture à leur niveau de risque, que celles qui sont couvertes dans le cadre d'un contrat groupe. En effet, les assurés souscrivant des contrats groupes sont moins sensibles aux prix (l'employeur paie une partie de la prime), leurs choix sont plus restreints que sur le marché individuel et la péréquation des risques est globalement plus large au sein de l'entreprise. Ettner segmente la population en quatre sous-groupes en croisant deux critères : le fait que les individus soient couverts ou non (reflétant une propension plus ou moins élevée à consommer des soins) et celui qu'une couverture de groupe leur soit proposée ou non. Elle étudie la différence de consommation entre ces quatre sous-populations. Le fait d'être couvert (par rapport au fait de ne pas l'être) donne l'effet sur la consommation du risque moral. Une fois ceci pris en compte, l'impact de la variable indicatrice "couvert à titre individuel" dénote pour sa part l'antisélection. Après corrections pour les variables socio-démographiques et celles d'état de santé, on voit que la variable indicatrice "couvert à titre individuel" a un effet résiduel significativement et positivement corrélé avec la consommation. Autrement dit, une fois que le risque moral est pris en compte, les personnes qui ont plus de choix (sur le marché individuel) ont une propension à consommer plus élevée que les autres. Sur le marché individuel, le niveau d'assurance augmente donc plus avec le niveau de risque que sur le marché de groupe.

Pour finir, on peut signaler que la difficulté de séparer risque moral et antisélection est atténuée lorsque l'on dispose d'information sur les anticipations que forment les ménages sur leurs dépenses futures. Marquis et Phelps (1987) montrent ainsi que les dépenses futures anticipées par les ménages sont significativement corrélées avec la

probabilité d'acquiescer une assurance et mesurent un effet d'antisélection qui n'est pas biaisé par le risque moral.

1.2 Expériences naturelles

Un certain nombre de travaux se fondent sur l'observation des comportements d'individus qui initialement sont tous couverts par un contrat identique et qui se voient, à un moment donné, offrir une gamme de contrats différenciés. La plupart des résultats empiriques confirment la tendance selon laquelle les bas risques choisissent des contrats moins généreux. Ces travaux sont souvent fondés sur des données américaines mais quelques exemples émanent d'autres pays.

Un observateur privilégié de ce phénomène est la HCFA⁴ : depuis plusieurs années, les bénéficiaires de Medicare peuvent s'inscrire auprès de HMO pour recevoir les soins par le biais de leurs réseaux. On constate toujours (Riley, Lubitz et Rabey, 1991; Riley et alii, 1996 ; Morgan et alii, 1997) que l'état de santé général, l'année précédant le choix, des populations Medicare qui choisissent d'entrer dans un HMO est nettement meilleur que celui des personnes qui restent dans le système traditionnel.

Une revue de littérature très exhaustive (Hellinger, 1995) conclut qu'aux Etats-Unis, tous les plans qui restreignent les choix des assurés à un nombre déterminé de prestataires de soins (les HMO et les EPO⁵) bénéficient d'une segmentation spontanée qui leur est favorable, parmi les personnes âgées mais aussi dans la population générale. Une majorité de ces études se fonde sur l'observation des consommations de soins dans la période qui précède le choix et quelques-unes comparent des données d'état de santé (nombre de maladies ou état de santé déclaré).

Le phénomène n'est toutefois pas limité aux Etats-Unis et si les analyses empiriques sont plus rares, elles convergent vers des résultats similaires. Ainsi Pellisé (1995) observe le comportement des fonctionnaires espagnols, qui ont le choix de s'inscrire auprès de réseaux privés ou de bénéficier du système public de type NHS. Elle montre que les malades chroniques sont beaucoup plus nombreux dans ce dernier. En Suisse,

⁴ *Health Care Financing Administration* : Administration américaine qui gère les programmes Medicare et Medicaid.

⁵ Les Exclusive Provider Organization sont des plans qui ne remboursent les frais que si les personnes se soignent auprès de prestataires désignés.

après la transformation d'une assurance traditionnelle en HMO, on a constaté que les personnes qui choisissaient de quitter le HMO avaient occasionné l'année précédente 39% de dépenses de plus que celles qui acceptaient le passage (Perneger, et al., 1995).

Ces études "avant-après" ou les expériences naturelles sont critiquables à deux égards :

- d'une part, Hellinger souligne que les études surestiment le biais de sélection qui se dissipe probablement au fil du temps, selon le phénomène de régression vers la moyenne⁶ ;
- d'autre part, on peut estimer que le résultat observé provient plus d'une sélection active des risques de la part des assureurs (cf. chapitre 9) que d'une volonté des bas risques d'émettre un signal pour bénéficier d'un prix actuariellement juste. Cet argument ne peut être rejeté a priori, mais on peut noter que le cadre réglementaire, au moins pour les bénéficiaires de Medicare, est très strict : tout HMO qui accepte une délégation de gestion de la HCFA doit couvrir quiconque en fait la demande.

En dépit de ces réserves, les résultats présentés jusqu'ici semble donc confirmer de façon assez systématique l'hypothèse initiale selon laquelle les hauts risques se couvrent plus que les bas risques.

1.3 Analyse critique

En réalité, le fait stylisé que l'on observe dans la totalité de ces études est que si les assurés se voient proposer deux contrats, l'un généreux (généreux pouvant signifier plus libéral dans l'accès aux producteurs de soins) et l'autre moins, les personnes qui présentent des niveaux de risque élevés vont choisir le premier et les bas risques le second. Il y a donc une tendance à la segmentation spontanée si les assureurs proposent des contrats différenciés.

⁶ La consommation a une composante aléatoire. Les personnes qui choisissent l'assurance généreuse en t ont eu une consommation plus élevée en $t-1$. Ceci peut refléter un niveau moyen de consommation toujours plus élevé, ou simplement le fait que la "réalisation" de la variable aléatoire était élevée en $t-1$. Dans ce cas, la consommation des personnes concernées tendra à se rapprocher de sa moyenne par la suite. Les différences observées initialement entre niveaux de consommation peuvent être très élevées et très significatives, mais se dissiper au fil du temps.

Cependant, il nous semble que le fait stylisé mis en évidence ne permet pas en soi de déterminer qu'il y a antisélection ou de montrer qu'un équilibre de Rothschild et Stiglitz s'instaure. Dans le modèle que proposent ces auteurs, la différence de risque est formalisée par une probabilité plus ou moins élevée de subir un dommage et il n'y a pas de frais de gestion. Tous les individus demandent donc la même quantité d'assurance (assurance complète) et paient des prix différents, égaux à leur probabilité de dommage. Si les assureurs sont parfaitement informés, ils offrent un contrat complet et modulent les tarifs en fonction des risques. Dans le cas contraire, ils offrent deux contrats. Toutefois, dans la réalité, quand bien même le tarif retracerait parfaitement les niveaux de risque, il n'est pas évident que les bas risques choisiraient le même contrat que les hauts risques car ils peuvent préférer un contrat moins "couvrant".

On peut construire un exemple pour illustrer cette affirmation. Considérons deux personnes A et B qui ont la même probabilité de subir un dommage. Faisons l'hypothèse que la prise en charge du dommage nécessitera de mettre en œuvre des moyens plus importants pour A que pour B (pour une raison exogène). Si A est une personne âgée et B un enfant casse-cou, on peut imaginer qu'ils ont une probabilité comparable d'avoir une fracture de la jambe. Il est aussi vraisemblable que les moyens médicaux nécessaires pour soigner la personne âgée et réparer son dommage sont plus élevés que ceux qu'il faut mettre en œuvre pour soigner l'enfant. Les deux personnes cherchent à acheter une assurance pour couvrir ce dommage. En information symétrique, la probabilité de dommage de ces personnes étant identique, le prix unitaire de l'assurance est le même. La personne âgée achète plus d'assurance que l'enfant puisque le dommage attendu est plus élevé. La personne mieux couverte présente un niveau de risque plus élevé. Ceci ne signifie pas qu'il y a antisélection, puisque l'information sur l'âge respectif des assurés peut être connue de l'assureur assez facilement.

De façon plus générale, dans le premier chapitre de ce travail, nous avons aussi vu que si les frais de gestion ne sont pas nuls, la prime maximale qu'un individu est prêt à payer augmente lorsque la probabilité ou le montant attendu du dommage augmente⁷, autrement dit lorsque le niveau de risque est plus élevé.

En toute rigueur, la corrélation observée entre niveau de risque et niveau de couverture peut être la simple expression d'une demande d'assurance en information parfaite.

Les éléments empiriques que nous avons présentés confirment surtout l'intuition selon laquelle les individus prennent en compte leur niveau de risque dans leur décision d'assurance (Price et Mays, 1985). Il nous semble donc que le postulat précédemment mentionné selon lequel l'antisélection est la tendance des hauts risques à acheter plus d'assurance que les bas risques n'est pas parfaitement rigoureux. Le fait stylisé mis en évidence par ces études est l'existence d'une segmentation spontanée des risques. On appelle segmentation spontanée le fait que, si l'assureur offre des contrats différenciés à une population hétérogène, les personnes qui présentent des niveaux de risque différents vont être tentées de souscrire des contrats différents⁸.

Il reste néanmoins à souligner que l'existence d'une telle segmentation spontanée est une condition nécessaire pour surmonter l'antisélection :

"adverse selection actually requires that individuals with different risk profiles behave differently in some way which the insurance company can exploit"⁹ (Rothschild et Stiglitz, 1997, pp.75).

En présence d'asymétrie d'information, les assureurs peuvent, en différenciant leurs contrats, amener les assurés à révéler de l'information sur leur niveau de risque. Plus précisément, la stratégie qui permet aux assureurs de surmonter leur manque d'information s'appuie sur le fait que les individus, qui ont des caractéristiques non

⁷ Dans certaines limites pour la probabilité : si elle est trop proche de 1, le bénéfice tiré de l'assurance est trop faible.

⁸ Le terme d'auto-sélection est aussi adapté, néanmoins, comme nous étudions dans cette partie l'antisélection puis la sélection, il nous semble préférable, par souci de clarté, d'adopter une terminologie plus éloignée.

⁹ Le modèle d'antisélection nécessite que les individus qui ont des profils de risque différents se comportent d'une façon différente, diversité que les assureurs peuvent mettre à profit (sous-entendu : pour les différencier).

observables différentes mais liées à leur niveau de risque, vont se segmenter spontanément si l'assureur leur propose des contrats différents.

A ce stade, nous avons montré une convergence des études empiriques sur le fait que les personnes qui se voient proposer des contrats plus ou moins généreux, tendent à se segmenter en fonction de leur risque. Toutefois, comme les risques en santé ne se prêtent pas nécessairement à une formalisation aussi simple que celle sur laquelle se fonde l'analyse de Rothschild et Stiglitz, ce fait stylisé ne suffit pas à confirmer l'existence d'antisélection.

2. PREDICTIONS DU MODELE DE ROTHSCHILD ET STIGLITZ

Afin de déterminer si l'antisélection est bien un problème sur le marché de l'assurance santé, il convient dans un premier temps de s'interroger sur le fait de savoir si l'hypothèse selon laquelle les assurés sont mieux informés de leur propre niveau de risque que les assureurs, est fondée (2.1). L'asymétrie d'information est en effet une condition nécessaire à l'existence d'antisélection. Dans la mesure où elle est vérifiée, nous recensons les éléments empiriques disponibles pour caractériser l'équilibre sur le marché de l'assurance santé. Nous cherchons plus précisément à savoir si, comme le montre l'analyse de Rothschild et Stiglitz, les bas risques, en présence d'antisélection, renoncent à obtenir le service qu'ils souhaiteraient acquérir, afin de révéler une information aux assureurs sur leur niveau de risque (2.2). Puis nous présentons les quelques études empiriques qui analysent plus spécifiquement la nature et la dynamique de l'équilibre (2.3). Enfin, nous voyons si la différenciation des contrats permet effectivement aux assureurs de surmonter l'asymétrie d'information (2.4).

2.1 Asymétrie d'information en faveur de l'assuré

Toute personne, qu'elle cherche ou non à s'assurer, détient une certaine information sur son propre niveau de risque et est en mesure d'anticiper en partie les problèmes de santé auxquels elle risque d'être, à terme, confrontée : *"individuals certainly know something about their projected medical needs. People with a history of heart*

*disease, for example, know they are at above risk for a heart attack*¹⁰ (Cutler, 1994b, pp.30).

Les informations dont dispose l'individu sur son histoire médicale, celle de sa famille, sont de nature privée et ne sont pas a priori accessibles à l'assureur. Plus généralement, un certain nombre de variables qui expliquent les choix d'assurance des individus peuvent ne pas être connues de l'assureur au moment où il propose un contrat et un prix. C'est le cas de l'état de santé auto-évalué, des dépenses anticipées, et des dépenses passées (que l'assureur ne peut connaître que s'il couvrirait l'assuré dans la période antérieure).

La validation de cette intuition, condition nécessaire à l'émergence de l'antisélection, a donné lieu à peu de vérifications empiriques. Pour déterminer si un assuré est mieux informé qu'un assureur sur son niveau de risque, on pourrait montrer :

- d'une part, que l'assuré anticipe raisonnablement ce niveau de risque. C'est probablement le cas si son choix d'assurance en dépend de façon significative ;
- d'autre part, que l'assureur n'a pas accès à l'intégralité de cette information et qu'il ne propose pas un prix actuariel.

Dans ce cas, si les assurés parviennent à souscrire des contrats dont la prime est inférieure à celle qu'ils sont prêts à payer compte tenu de ce qu'ils savent de leur risque, l'asymétrie d'information est très probablement constituée¹¹.

Wolfe et Goddeeris (1991) proposent un test de cette hypothèse dans un modèle général qui présente l'enchaînement des décisions des assurés et des assureurs concernant l'assurance et la consommation de soins sur plusieurs périodes. Dans le modèle qu'ils proposent, les principales variables qui expliquent le comportement de l'individu vis-à-vis de l'assurance sont sa richesse, W_t et son état de santé.

¹⁰ *Les individus ont une certaine connaissance de leurs futurs besoin de santé. Par exemple, des personnes ayant déjà eu une maladie cardio-vasculaire, savent qu'elles courent un risque plus élevé d'être victimes d'une attaque cardiaque.*

¹¹ Dans la perspective d'une analyse du fonctionnement du marché, l'asymétrie d'information peut alors avoir dans ce cas donné lieu à ce qu'on appelle l'antisélection : les assureurs qui proposaient un prix moyen attirent majoritairement des personnes dont la prime est sous-estimée. Elle peut aussi avoir donné lieu à une péréquation entre les personnes qui présentent des niveaux de risque différents (pooling ou Miyasaki). Je remercie le Pr. Pierre Picard d'avoir souligné ce point lors du séminaire du Delta, Juin 1999.

Pendant la période t , ce dernier est égal à H_t , état de santé auto-déclaré duquel on déduit un élément aléatoire ε_t (plus ε_t est élevé, moins la personne est en bonne santé). A l'instant où l'individu prend la décision de s'assurer, il connaît W_t , H_{t-1} et ε_{t-1} mais ni H_t ni ε_t . On considère cependant que ε_t est soumis à un processus stochastique auto-régressif d'ordre 1.

Cette hypothèse est centrale :

- elle traduit le fait que l'état de santé de l'individu contient une composante aléatoire ;
- elle rend compte de l'évolution d'une année sur l'autre de l'état de santé (par exemple un processus de guérison progressive¹²) ;
- enfin, elle permet de postuler une asymétrie d'information : les caractéristiques du processus stochastique sont connues de l'assuré mais pas de l'assureur. Le premier détermine, sur la base de cette information, la prime d'assurance qu'il est prêt à payer P_t^* . Il se fonde notamment sur l'impact d'une variation de son état de santé en $t-1$ sur la demande de soins en t . L'assureur propose pour sa part une prime P_t dont on suppose qu'elle prend en compte l'information disponible sur le risque. On considère que l'assureur est en mesure de traduire une proportion λ de la variation anticipée des dépenses dans la prime. Si $\lambda=1$, il n'y a pas d'asymétrie d'information et donc pas d'antisélection. Si $\lambda < 1$, il y a asymétrie d'information en faveur de l'assuré.

La décision de s'assurer dépend du signe de la différence entre la prime qu'il est prêt à payer et la prime qui lui est proposée :

$$(P_t^* - P_t) = f(W_t, H_{t-1}, \varepsilon_{t-1})$$

? - +

L'individu s'assure si la prime proposée par l'assureur est inférieure ou égale à celle qu'il est prêt à payer en fonction des informations privées dont il dispose. En règle générale, plus le risque est élevé, plus les individus sont prêts à s'assurer¹³.

¹² L'âge a peu d'influence sur une période assez courte.

¹³ A moins que la probabilité de dommage ne soit très proche de 1 et que les frais de gestion soient élevés.

L'effet total de la richesse sur la différence étudiée ou plus précisément sur P_t^* est ambigu car :

- tout d'abord, si l'aversion au risque est une fonction décroissante de la richesse, pour un niveau de risque donné, la prime maximale que l'individu est prêt à payer diminue quand la richesse augmente ;
- à l'inverse, comme les soins sont des biens normaux dans le modèle, la valeur accordée par l'individu à la réduction du prix des soins que représente l'assurance, augmente avec la richesse. Cet effet ne peut toutefois s'exercer que si l'assureur n'est pas en mesure de refléter dans la prime P_t l'augmentation intégrale des coûts anticipée par l'individu.

Au total, si la demande d'assurance augmente avec la richesse et que le second effet l'emporte, on peut penser que $\lambda < 1$.

Par ailleurs, on estime à la période t la demande de soins, une fois le contrat d'assurance souscrit, sur la base de l'équation suivante :

$$M_t = M (\underset{+}{W_t}, \underset{-}{(H_t - \varepsilon_t)}, \underset{+}{I_t})$$

où I_t est une variable indicatrice de la présence d'assurance dont le coefficient (a priori positif) mesure le risque moral.

Le modèle est testé sur données de panel et la variable d'assurance étudiée est la souscription d'un contrat Medigap¹⁴. Les données dont les auteurs disposent sont très complètes et comprennent les dépenses sur différentes périodes, les primes d'assurances payées et des indications sur l'évolution de l'état de santé des individus.

Les résultats de l'étude de la consommation montrent tout d'abord que le coefficient d'auto-corrélation est significatif, lorsqu'on a pris en compte des variables explicatives parmi lesquelles figurent l'état de santé auto-déclaré et l'invalidité. Pour ce qui concerne l'assurance :

- on observe que l'état de santé déclaré en $t-1$ n'a pas d'impact significatif sur la décision d'assurance, ce qui peut laisser à penser qu'une partie des variations d'état de santé est bien prise en compte dans les primes ;

¹⁴ Souscrit en complément de Medicare.

- par contre, on trouve une corrélation significative entre les dépenses *non expliquées* (les résidus de l'équation de consommation de soins d'une période ε) et les décisions d'assurance subséquentes. Autrement dit, les personnes qui ont récemment dû faire face à des dépenses de santé qui ne sont pas expliquées par leurs caractéristiques personnelles "de long terme" achètent plus d'assurance parce leurs besoins persistent dans le temps ;
- enfin, on observe un effet positif, élevé et significatif de la richesse sur les niveaux d'assurance.

Ces deux derniers éléments donnent à penser que λ est inférieur à 1, c'est à dire, qu'il y a asymétrie d'information, puisque des informations auxquelles l'assureur n'a pas accès modifient les choix d'assurance des individus. Pour autant, puisque les personnes qui le souhaitent parviennent à souscrire des contrats, l'antisélection n'est pas constitutive d'un échec de marché majeur. Ceci dit, la portée de ce dernier résultat doit probablement être atténuée pour deux raisons :

- l'assurance ne porte pas sur l'intégralité du risque, puisque l'échantillon est constitué de personnes qui bénéficient de la couverture publique de Medicare ;
- les données sur lesquelles portent les tests couvrent la période 1971-1979. Or, la croissance des dépenses de santé et des primes n'est véritablement devenue un enjeu majeur aux Etats-Unis que dans les années quatre-vingt. La concurrence s'est alors accrue sur les marchés d'assurance et on peut se demander si elle n'a pas, depuis, conduit à accentuer l'antisélection.

2.2 Contrainte pesant sur les bas risques

Sous l'hypothèse que les assurés connaissent mieux leur risque que les assureurs, le modèle de Rothschild et Stiglitz montre que les bas risques ne peuvent obtenir autant d'assurance qu'ils le souhaiteraient. Ils sont les principales victimes de l'antisélection puisqu'ils doivent accepter des contrats moins avantageux que les contrats auxquels ils pourraient prétendre en information complète. Rares sont les tests empiriques de cette hypothèse (2.2.1). En revanche, si l'on a une vision moins restrictive des sacrifices auxquels consentent les bas risques, on peut trouver une certaine pertinence aux résultats théoriques (2.2.2).

2.2.1 Limitation de la quantité d'assurance

Un résultat de l'analyse de l'équilibre de Rothschild et Stiglitz est qu'en présence d'asymétrie d'information, l'équilibre le plus susceptible de s'instaurer est tel que les bas risques sont moins bien couverts que les hauts risques. Mais en réalité, étudier si un équilibre séparateur de type Rothschild et Stiglitz s'instaure, nécessite de démontrer :

- que les haut risques sont assurés complètement et qu'ils paient un prix actuariel ;
- et que les bas risques paient aussi un prix actuariel et qu'ils s'assurent moins que ce qu'ils ne feraient si le marché était en information parfaite.

Dans ce cas, l'assureur est parvenu à mettre en place un mécanisme séparateur et un équilibre de type Rothschild et Stiglitz s'est bien instauré.

Or, il existe très peu d'études qui permettent de confirmer cette hypothèse. Browne (1992) est une exception. En fait, l'auteur teste l'hypothèse selon laquelle les bas risques, dans un marché avec antisélection achètent moins de couverture que sur un marché sans antisélection. Le marché des contrats souscrits par le biais des entreprises est présumé moins soumis à l'antisélection que le marché individuel, comme dans Ettner (1997). Le principe du travail de Browne est assez simple : après avoir estimé la couverture dont bénéficient les ménages couverts dans le cadre de contrats groupes, Browne simule la couverture assurantielle dont bénéficieraient les ménages qui se couvrent sur le marché individuel, s'ils étaient couverts par le biais d'une entreprise. Il étudie ensuite la différence entre la couverture simulée et la couverture effective des personnes. Cette différence est positivement corrélée avec un certain nombre de facteurs de risque que l'assureur utilise pour tarifer les contrats individuels mais pas les contrats groupes : une personne âgée est ainsi mieux couverte sur le marché de groupe que sur le marché individuel¹⁵. Mais le résultat le plus intéressant est que la variable indicatrice d'un bon état de santé, qui est construite sur la base de l'état de santé auto-déclaré (information susceptible de ne pas être connue des assureurs), est positivement corrélée avec la différence étudiée. Autrement dit, toutes

¹⁵ La différence est aussi corrélée positivement avec le revenu : des avantages fiscaux spécifiques font que les personnes qui sont plus riches "gagnent" plus à être couvertes sur le marché de groupe et se couvrent donc moins sur le marché individuel.

choses égales par ailleurs, une personne présentant un faible niveau de risque se couvre moins sur le marché individuel que sur le marché de groupe. Une caractéristique fondamentale de l'équilibre de Rothschild et Stiglitz semble donc confirmée par ce test. Browne souligne toutefois que ce résultat n'est valable que dans la mesure où les bas risques conservent un niveau de choix raisonnable dans le marché de groupe et que leurs décisions sur ce marché s'approchent de celles qu'ils réaliseraient en information parfaite. Dans la mesure où les employeurs offrent souvent un choix entre plusieurs contrats et que l'on peut espérer (selon l'auteur) qu'ils cherchent à prendre en compte les souhaits de leurs employés en déterminant les contrats qu'ils vont offrir, cette hypothèse n'est pas invraisemblable.

S'il nous a paru intéressant de présenter la démarche de Browne, son étude ne permet pas de déterminer la nature de l'équilibre qui s'instaure sur le marché. Ainsi, le fait que les bas risques soient "moins bien couverts" pourrait très bien s'accorder avec un équilibre de Miyasaki où les bas risques, tout en étant incomplètement assurés, subventionnent les contrats des hauts risques.

2.2.2 Un critère de jugement différent ?

Les éléments empiriques dégagés dans l'analyse de la segmentation spontanée, en particulier dans les études qui portent sur les expériences naturelles, peuvent toutefois être interprétés dans la perspective qui nous intéresse. Lorsqu'on constate que les bas risques choisissent massivement les plans les moins généreux et les hauts risques les plus généreux, on ne se fonde pas seulement sur le niveau de couverture. Ces résultats empiriques, qui sont largement probants, se différencient de ceux qui testeraient l'hypothèse de Rothschild et Stiglitz "au pied de la lettre", en ce qu'ils ajoutent une dimension à l'analyse : le plan qualifié de moins généreux est en général un HMO et le plan le plus généreux présente une forme d'organisation plus souple (assurance traditionnelle ou PPO). Le fait que les copaiements sont plus ou moins élevés ou que certaines catégories de soins sont ou non prises en charge, autrement dit que l'assurance est complète ou non, semble importer relativement moins que l'impression que l'accès aux soins est contrôlé. Les hauts risques optent pour des plans dans lesquels ils bénéficient d'une plus grande liberté dans le choix de leurs prestataires (Hellinger, 1995). Les bas risques acceptent plus facilement de renoncer à cette liberté

et ils peuvent même y gagner en termes de remboursement global. Ainsi, pour reprendre l'exemple de la délégation de gestion de Medicare aux HMO, ceux-ci couvrent parfois les médicaments (qui ne sont pas couverts par la HCFA) sans exiger de paiement de prime supplémentaire. Les conditions de remboursement sont, de fait, plus favorables que dans le système traditionnel et les bas risques choisissent massivement cette option.

Contrairement au modèle de Rothschild et Stiglitz, dans une grande majorité des cas que nous avons étudiés, ce n'est pas la quantité d'assurance qui discrimine mais le fait d'avoir accès à un service différent. Quant à savoir si le renoncement à une grande liberté de choix du prestataire pour entrer dans un réseau peut être considéré comme un *sacrifice* auquel les bas risques consentent pour ne pas risquer de payer pour les hauts risques dans le cadre d'une assurance traditionnelle, la question reste ouverte. En tout état de cause, il s'agit là d'une spécificité de l'assurance maladie qui reflète indirectement l'importance du "colloque singulier" entre le patient et le médecin. S'il peut sembler naturel qu'un assureur automobile propose une liste de garages agréés, il est probablement plus difficile de convaincre un patient suivi très régulièrement par le même médecin de l'abandonner au profit d'un autre praticien agréé.

En élargissant encore le propos, on peut se demander si l'assureur, pour segmenter les risques, est réellement obligé de faire systématiquement peser une contrainte sur les bas risques : en offrant des contrats couvrant les services préventifs et d'autres couvrant largement les dépenses hospitalières, l'assureur pourra obtenir des personnes qu'elles choisissent des contrats différents (et donc qu'elles révèlent leur niveau de risque), tout en leur proposant des contrats adaptés à leur besoins. Etant donné la nature du risque que l'on étudie, et quand bien même l'asymétrie d'information serait constituée, l'assureur peut la surmonter en partie en différenciant son offre.

2.3 Equilibre du marché avec antisélection

La nature de l'équilibre qui s'instaure sur un marché avec antisélection dépend, selon la théorie, des anticipations des agents présents sur le marché. Malgré cette difficulté, quelques travaux empiriques cherchent à déterminer la nature de l'équilibre (2.3.1).

Par ailleurs, le résultat selon lequel la présence d'antisélection pourrait déstabiliser le marché semble permettre d'expliquer certains phénomènes observés (2.3.2).

2.3.1 Nature de l'équilibre

Les études cherchant à déterminer le type d'équilibre qui prévaut sur des marchés d'assurance santé sont très rares. La complexité du sujet, les biais introduits par les différentes réglementations et caractéristiques des institutions et des agents (employeurs, ménages, assureurs...), ainsi que le manque de transparence des marchés expliquent probablement cette désaffection. S'ajoute à cela la difficulté posée par le choix des concepts d'équilibre et des hypothèses d'anticipation des individus. D'après Hellwig (1987, pp.325) :

- tout d'abord, il est important de déterminer lequel des différents modèles d'anticipation est approprié pour le marché considéré ;
- d'autre part, cet exercice peut s'avérer impossible parce que l'enchaînement des décisions, qui est crucial, n'est pas nécessairement observable ni même stable pour le marché considéré.

De leur côté, Rothschild et Stiglitz rejettent l'idée que des anticipations sophistiquées (de type Wilson) puissent gouverner les actions des acteurs : "*does it make sense for a small firm to worry that its actions will bring about a collapse of the entire market ?*"¹⁶(Rothschild et Stiglitz, 1997, pp.78).

Tout au plus peut-on espérer observer localement des phénomènes et les analyser à la lumière des modèles. Ainsi, on peut chercher à déterminer si, et dans quelle mesure, les assureurs tarifent incorrectement les risques :

- si tout le monde obtient la même assurance et paie le même prix (les bas risques payant alors un prix trop élevé), on est plutôt dans une configuration d'équilibre *pooling* ;
- si les contrats sont différenciés, il peut malgré tout y avoir un transfert de richesse des bas vers les hauts risques si les prix payés par les assurés ne sont pas actuariellement justes. On est alors plutôt dans une configuration de type Miyasaki.

¹⁶ *Est-ce vraiment réaliste d'imaginer qu'une petite entreprise s'inquiétera de savoir si son action entraînera l'effondrement du marché ?*

Sur la base de données issues de l'enquête précédemment citée, Browne (1992) montre que, sur le marché individuel, le profit des assureurs (toutes choses égales par ailleurs) est positivement corrélé avec les variables dénotant un bon état de santé et négativement corrélé avec celles dénotant un état de santé dégradé. Ces variables d'état de santé sont construites sur la base de la déclaration subjective de l'état de santé. Elles rendent donc compte d'une information qui n'est pas directement accessible à l'assureur. Pour l'auteur, ce résultat confirme l'idée que l'équilibre qui émerge se traduit par un certain degré de subvention croisée (équilibre *pooling* ou de type Miyasaki) et qu'en tout état de cause, l'équilibre d'information parfaite ne prévaut pas sur le marché de l'assurance santé. En approfondissant son analyse, Browne (1993) montre aussi que, sur le marché individuel qu'il considère, le montant des franchises, des plafonds de participation financière des assurés, celui des plafonds d'engagement des assureurs, ainsi que les prix de l'assurance à couverture donnée, ne sont pas différents, toutes choses égales par ailleurs, entre hauts et bas risques. Ceci penche en faveur d'un équilibre *pooling* pour le marché étudié.

2.3.2 Absence d'équilibre

Ceci dit, de façon un peu paradoxale, l'enseignement principal de Rothschild et Stiglitz est, selon les auteurs eux-mêmes, que : "*Insurance markets differ from most markets because in insurance markets competition can destroy the market rather than make it work better*"¹⁷ (1997).

Un certain nombre d'études empiriques ou de faits stylisés appuient cette idée et nous avons choisi d'en présenter deux dont les méthodologies sont radicalement différentes mais dont les résultats convergent.

¹⁷ *Les marchés d'assurance diffèrent de la plupart des marchés parce que sur les marchés d'assurance la concurrence peut détruire le marché plutôt qu'améliorer son fonctionnement.*

a) Le cas de Harvard

Cutler et Reber, (1996) développent une analyse de cas. Ils étudient les conséquences d'un changement de politique de financement des achats d'assurance des employés de l'université de Harvard. Dans l'ancien système, les subventions de l'employeur pour les personnes qui choisissaient les plans les plus généreux (en termes de niveau d'assurance mais aussi d'accès aux soins) étaient les plus élevées. Dans le nouveau, la contribution de l'employeur est identique pour tous les ménages. Ce changement vise à promouvoir la concurrence entre les assureurs qui feront désormais face à des assurés plus sensibilisés aux prix. Ce changement de mode de financement conduit dans un premier temps des personnes à abandonner le plan le plus généreux¹⁸. Une étude plus détaillée montre que les personnes qui quittent le plan généreux sont plus jeunes et en meilleure santé que celles qui restent. Le plan généreux est donc obligé d'augmenter ses primes l'année suivante pour compenser l'augmentation du risque moyen, entraînant une nouvelle sortie des bons risques. Indépendamment du fait que le plan généreux soit bien géré ou non, son prix augmente de plus en plus. Théoriquement, deux événements peuvent se produire :

- soit un équilibre est atteint, tel que la personne qui est en meilleure santé dans le plan généreux est indifférente entre le bénéfice qu'elle tire des avantages accordés par le plan généreux et la diminution du prix qu'elle pourrait obtenir en changeant ;
- soit une "spirale de la mort" s'engage et le plan est finalement retiré du marché. En l'occurrence, deux ans après la réforme engagée, l'option généreuse proposée au personnel de Harvard a disparu.

Les auteurs démontrent que ce phénomène entraîne une diminution du bien-être total. Mais elle est plus que compensée par le fait que, parallèlement, on constate une diminution sur l'ensemble des plans du prix de l'assurance (y compris par rapport à la région). La conclusion des auteurs est que la diminution de bien-être des hauts risques pourrait être compensée par l'instauration d'un paiement de l'université qui serait

¹⁸ les informations disponibles permettent aux auteurs d'estimer une élasticité-prix moyenne de la demande d'assurance.

différencié selon les catégories de plans pour tenir compte des facteurs de risque mais qui resterait uniforme au sein d'une catégorie pour stimuler la concurrence marginale.

b) Une analyse de micro-simulation

Marquis (1992) démontre la possibilité qu'un tel phénomène puisse se produire dans le cadre d'un modèle beaucoup plus complexe qui rend compte de l'interaction des décisions des agents. Sur la base des données de la Rand, elle simule :

- en premier lieu, le choix des ménages entre quatre options d'assurance. La décision des ménages dépend en particulier des primes, des dépenses qu'ils anticipent et de leur réponse à une question qui leur était posée sur le fait qu'ils choisiraient ou non une couverture complétant celle dont ils bénéficient. La relation entre ces deux dernières variables permet d'estimer leur propension à intégrer leur propre risque anticipé dans leur demande d'assurance ;
- puis, les dépenses des ménages, une fois ce choix réalisé en prenant en compte l'effet du risque moral et celui des dépenses anticipées¹⁹ ;
- enfin, une prime et une part de marché sont estimées par contrat et selon la taille du ménage (méthode de tarification courante dans les contrats groupes aux Etats-Unis). Lorsque l'assureur utilise l'information dont il dispose sur les coûts de l'année précédente (*experience rating*) pour tarifer ses contrats, on montre que le prix des options se stabilise au bout de quatre périodes et qu'à l'équilibre, un certain nombre de contrats (généreux) ont disparu.

Les deux études que nous avons présentées illustrent un phénomène souvent observé aux Etats-Unis : *"almost every large firm that has encouraged employee choice has found that the cost of the most generous policies increases sufficiently rapidly that those plans are no longer viable (often termed a death spiral)"*²⁰ (Cutler, 1994b, 30).

¹⁹ On demande au début de l'enquête à chaque ménage d'estimer ses dépenses anticipées, Marquis étudie alors la différence entre cette estimation et les dépenses observées par la suite pour évaluer dans quelle mesure les anticipations du ménage rendent compte d'une information dont l'assureur ne dispose pas et se réalisent. Cette mesure permet ensuite d'intégrer les dépenses anticipées par le ménage dans l'estimation des dépenses futures.

²⁰ *Presque toutes les grandes entreprises qui ont encouragé le choix de l'assurance par leurs employés se sont rendu compte que le coût de la police la plus généreuse augmentait suffisamment rapidement pour qu'elle ne soit plus viable (ce qu'on appelle une spirale de la mort).*

Comme le souligne Pauly (1986, pp.651) cette disparition de certains contrats et l'instabilité manifeste des marchés n'est pas sans rappeler le résultat de Rothschild et Stiglitz sur la possibilité que l'équilibre n'existe pas. En tout état de cause, on montre que l'antisélection est suffisante pour éliminer les contrats généreux si les assureurs font payer une prime basée sur les dépenses passées (*experience rating*). Cette "spirale de la mort", qui entraîne l'éviction de contrats, alors même qu'ils répondent à un besoin objectif d'une partie de la population, est constitutive d'un échec de marché : "*adverse selection can lead to the disappearance of the market for some types of insurance hence to market failure*"²¹ (Poterba, 1994, pp. 10).

2.4 L'asymétrie d'information est elle surmontable ?

En s'éloignant d'une vision réductrice des enseignements du modèle de Rothschild et Stiglitz, on peut, à ce stade, remarquer que si les assurés détiennent des informations sur leur niveau de risque auxquelles les assureurs n'ont pas accès, ces derniers parviennent, en proposant différents contrats, à une certaine segmentation des risques. Ce résultat importe plus que le fait de savoir si les bas risques obtiennent effectivement le même service qu'en information symétrique, puisque le problème initial de la séparation des différentes populations est résolu. Une dernière question demeure cependant : en offrant des contrats différenciés et en segmentant les risques, les assureurs parviennent-ils à réduire le différentiel d'information et à offrir des prix actuariellement justes ?

Bakker et Van Vliet (1993), puis van de Ven et van Vliet (1995), cherchent à estimer dans le cadre d'un exemple précis l'ampleur du différentiel d'information entre assureur et assuré et s'interrogent sur son caractère irréductible. Les deux études sont basées sur la même enquête dans laquelle une question était posée aux ménages concernant leur préférence à l'égard de contrats avec ou sans franchise. Cette enquête contient en outre des informations démographiques et sur l'état de santé des assurés, ainsi que leurs dépenses antérieures et postérieures au moment où on leur propose un contrat avec franchise.

²¹ *L'antisélection peut entraîner la disparition du marché pour certains contrats d'assurance et donc un échec de marché.*

Bakker et van Vliet montrent que la différence de consommation entre les personnes qui ont choisi (de façon fictive) la franchise et les autres représente 53% du coût moyen total²². Autrement dit, les personnes qui choisissent la franchise ont un niveau de risque moins élevé que les autres. Si la réduction de prime, accordée par les assureurs aux personnes qui optent pour le contrat avec franchise, ne prend en compte que la diminution des dépenses basée sur le coût moyen de la population, elle est en conséquence largement trop élevée. En ajustant la diminution de prime en fonction de l'âge et du sexe de l'assuré, l'assureur ne parvient cependant pas à offrir aux personnes qui ont choisi la franchise un niveau de prime ajusté à leur risque. Ce premier travail souligne que l'ampleur du surplus informationnel de l'assuré dépend des facteurs de risque, des informations, que l'assureur utilise effectivement pour déterminer la prime. Il est maximal si l'assureur n'utilise aucune information mais il peut être réduit.

L'étude de van de Ven et van Vliet (1995) vise à estimer dans quelle mesure ce surplus est irréductible, autrement dit à déterminer si, en utilisant des informations disponibles sur le niveau de risque des assurés, les assureurs parviennent à évaluer le risque des personnes qui choisissent des contrats plus ou moins généreux.

Pour mesurer le différentiel d'information inévitable, les auteurs estiment la mesure suivante :

$$[C_{NF} - C_F] - [E^x(C_{NF}) - E^x(C_F)]$$

- C_{NF} représente les coûts totaux engendrés en $t+1$ par les personnes qui ne choisissent pas le contrat avec franchise et C_F , les coûts observés pour l'autre groupe.
- $E^x(C_{NF})$ est une estimation des coûts réalisée ex ante pour le groupe qui ne choisit pas la franchise. Elle prend en compte les facteurs de risque x et représente une estimation de la prime que facturerait l'assureur sur la base de cette information²³.

²² [(coût moyen des personnes qui expriment le souhait de ne pas choisir la franchise)-(coût moyen de celles qui choisissent la franchise)]/(coût moyen total).

²³ $E^x(C_F)$ est l'estimation sur la base des mêmes facteurs de risque pour les personnes qui choisissent la franchise.

Si la mesure est élevée, cela signifie que les facteurs de risque x ne reflètent pas les différences de niveaux de risque entre les deux groupes. A l'inverse, si elle est nulle, l'intégralité des différences de coût a été anticipée par les assureurs.

En $t+1$, les dépenses des personnes qui ne choisissent pas la franchise (C_{NF}) sont environ de 60% plus élevées que celles des autres (C_F). Cette différence représente une mesure du surplus informationnel "brut" qui prévaut si l'assureur ne détient aucune information sur les individus. Le tableau ci-dessous synthétise les résultats de l'étude. La tarification âge sexe permet de réduire le surplus de 40%. En utilisant l'ensemble des informations disponibles sur l'état de santé (au demeurant assez grossières) et sur les consommations antérieures, le surplus est diminué de 80%.

Importance du surplus informationnel

	Part de la variance expliquée	Valeur relative du surplus ²⁴	Indice
Aucun ajustement	0%	43%	100
Age sexe	1.9%	26%	61
+ indicatrice d'assurance complémentaire et région	2.1%	23%	53
+ 5 variables d'état de santé <i>nombre de consultations de généraliste l'année précédente, montant des prescriptions pharmaceutiques, nombre de jours de maladie, le fait d'être en invalidité, le fait de déclarer avoir un état de santé plus ou moins bon qu'il y a deux ans</i>	4.9%	12%	28
+ Les coûts de l'année précédente <i>indicatrice de dépenses positives et log des dépenses +1</i>	6.5%	7%	17

Source : van de Ven et van Vliet (1995, pp.161)

Cette étude ne prétend pas déterminer de façon définitive les facteurs de risque que l'assureur doit utiliser pour éliminer le surplus informationnel de l'assuré, mais elle montre que la tâche est difficile et que le fait de séparer les hauts et les bas risques ne suffit pas à garantir que l'assureur offre un prix actuariellement juste. Toutefois, lorsque l'ensemble des informations disponibles est utilisé, ce surplus est largement réduit jusqu'à devenir non significatif, au sens statistique du terme (c'est le cas dans la

²⁴ La valeur relative du surplus est calculée de la façon suivante : [(Coûts observés pour les personnes refusant la franchise - coûts estimés en prenant en compte des facteurs de risques du modèle décrit dans la ligne correspondante) - (Coûts observés pour les personnes choisissant la franchise - coûts estimés en prenant en compte des facteurs de risques du modèle)] / la moyenne des coûts totaux sur la population.

dernière ligne du tableau). Les individus continuent à détenir plus d'information que les assureurs sur leur niveau de risque, mais ils en tirent un avantage considérablement réduit en terme de prime. Un tel résultat nécessite qu'une quantité d'information substantielle soit disponible et l'intégration de ces facteurs de risque dans la prime suppose d'estimer la demande de soins. Pour ce faire, le meilleur modèle économétrique doit être sélectionné et estimé sur un échantillon assez large. En outre, le travail effectué par van de Ven et van Vliet s'affranchit d'un grand nombre de difficultés. Tout d'abord, les données disponibles concernent des personnes qui ont toutes la même couverture sur l'ensemble de la période. On leur demande si elles choisiraient une franchise, mais ce choix reste fictif. Ceci facilite l'évaluation du différentiel car en réalité, la détermination des primes des contrats avec et sans franchise devrait prendre en compte, outre le niveau de risque étudié ici :

- la variation directe des coûts pour l'assureur (qui débourse un montant moindre, toutes choses égales par ailleurs, dans le contrat avec franchise) ;
- la diminution des frais de gestion liée au fait qu'il rembourse moins de sinistres ;
- et enfin les modifications de comportement de consommation liées au risque moral.

Au terme de cette seconde section, il nous semble donc que :

- les analyses empiriques rigoureuses des prédictions du modèle d'assurance avec antisélection restent assez rares dans le cadre de la santé ;
- l'affirmation de Cutler selon laquelle "*almost all health insurance systems where individuals are allowed choice over insurance have experienced adverse selection*"²⁵ (1994b, pp. 30) semble assez largement corroborée par les éléments empiriques disponibles estimant le niveau d'information des parties au contrat, le comportement des individus et la dynamique des marchés ;
- toutefois, excepté dans les cas où des spirales de la mort se mettent en place, la segmentation des risques peut permettre de surmonter l'asymétrie d'information. La tarification des risques reste cependant difficile, y compris si l'on parvient à segmenter les populations.

²⁵ la quasi-totalité des systèmes dans lesquels l'assurance santé est volontaire sont confrontés à l'antisélection.

En définitive, la critique principale adressée au modèle de Rothschild et Stiglitz selon laquelle "*les applications pratiques de ce modèle sont limitées*" (Masson, 1995, pp.139) semble excessive, même si ce modèle ne permet pas de refléter pleinement une réalité très complexe.

3. INTERVENTION DE L'ÉTAT : L'EXEMPLE DES PAYS-BAS

L'antisélection semble donc bien être un phénomène qui touche aussi les marchés d'assurance santé. Parmi les travaux présentés, rares sont ceux qui analysent le fonctionnement du marché dans une perspective globale et vont au delà de l'observation des comportements ou des réactions de tel ou tel acteur. Cette dernière section vise, au travers d'exemples issus de l'observation du fonctionnement des marchés, à alimenter cette analyse dans une perspective réglementaire. Nous étudions plus précisément le fonctionnement du marché de l'assurance privée aux Pays-Bas et les interventions successives de l'Etat sur ce marché à travers le prisme des enseignements du chapitre précédent. Le choix des Pays-Bas a été dicté par plusieurs critères : il s'agit d'un des pays où l'assurance privée joue le rôle le plus important²⁶ ; en outre, contrairement aux Etats-Unis, où la réglementation peut varier d'un Etat à l'autre, elle est uniforme aux Pays-Bas.

3.1 L'exemple des régimes à adhésion optionnelle

Une segmentation explicite en fonction des revenus prévaut pour les soins courants aux Pays-Bas. Les personnes dont le revenu est inférieur à un certain seuil sont couvertes par une assurance obligatoire, les autres ayant la liberté de s'assurer sur un marché privé. Cette segmentation a toutefois évolué vers un cloisonnement de plus en plus étanche au fil du temps.

Avant 1986, l'affiliation à une caisse publique était déjà obligatoire pour les salariés en deçà d'un seuil de revenu. Mais les caisses publiques géraient en outre des régimes à adhésion optionnelle. Au cours des années 1970, les assureurs privés adoptent tour à tour la différenciation des primes en fonction de l'âge et mettent progressivement en

²⁶ Nous analyserons ce point en détail dans le chapitre 7.

place des techniques destinées à améliorer le profil de risque des personnes qu'ils couvrent (sélection à l'entrée, exclusion de certaines pathologies du champ de l'assurance...). En conséquence :

- les personnes présentant des niveaux de risque bas sont progressivement attirées par le privé, où elles bénéficient de primes plus attrayantes que dans les régimes publics volontaires ;
- à l'inverse, un nombre croissant de personnes rencontrant des difficultés pour s'assurer auprès de compagnies privées parce que leur niveau de risque est élevé et remplissant les conditions d'adhésion aux régimes publics volontaires, se tournent en dernier ressort vers ces régimes gérés par les caisses publiques qui ne peuvent légalement refuser de les couvrir.

Les conséquences de cette évolution s'avèrent désastreuses pour l'équilibre financier des régimes à souscription volontaire gérés par les caisses publiques. Pour couvrir des déficits structurels, ces dernières augmentent les taux de cotisation au fur et à mesure que le risque moyen des personnes couvertes s'aggrave. Le système est alors remis à plat en 1986 et le régime d'assurance obligatoire uniformisé. Il absorbe certaines catégories de personnes auparavant couvertes par l'assurance volontaire et considérées comme particulièrement vulnérables (les personnes âgées bénéficiant de revenus faibles).

Les autres assurés des régimes volontaires démantelés sont obligés de s'assurer sur le marché privé. L'expérience des Pays-Bas sur le marché des soins courants tendrait donc à invalider le résultat théorique de Neudeck et Podczeck, selon lequel la coexistence de contrats privés et d'une assurance publique à adhésion volontaire non discriminatoire destinée aux hauts risques peut rétablir l'optimum de second rang. Cette forme d'intervention de l'Etat comme assureur semble en effet avoir été mise en échec par les pratiques d'écémage des assureurs privés. Elle s'est soldée, de fait, par un cloisonnement rendu étanche pour les personnes qui auparavant avaient le choix de passer d'un régime à l'autre. Il faut cependant rappeler que, d'après les prédictions du modèle, l'optimum n'est rétabli que si les bas risques sont explicitement taxés au profit des hauts risques. Or, rien de tel n'était mis en place avant 1986 aux Pays-Bas, les déséquilibres structurels des régimes volontaires étant épongés par des augmentations de cotisations. On peut alors se demander si la mise en place d'une taxe sur les bas

risques (les assurés du privé) aurait pu assurer le succès de cette coexistence. Il n'est évidemment pas possible de répondre directement à la question mais il est intéressant de constater que la réglementation a finalement évolué en ce sens : peu après la séparation des deux populations, un prélèvement spécifique a été instauré pour tous les assurés du privé qui perdure²⁷. Il se présente comme une compensation démographique, ce qui n'est pas incompatible avec l'esprit de la taxe sur les bas risques évoquée dans le modèle théorique, dans la mesure où risque et âge sont fortement corrélés.

3.2 Le contrat standard du marché privé

Le second exemple que nous développons ici s'analyse comme une tentative d'imposer aux assureurs privés de fournir un contrat standard. Ce contrat a été créé en 1986 au moment où certaines personnes auparavant inscrites auprès des caisses publiques se sont vues obligées de souscrire des contrats auprès des assureurs privés. L'Etat a alors imposé aux assureurs d'offrir à celles de ces personnes qui avaient du mal à se couvrir sur le marché privé un contrat appelé WTZ²⁸ dont le prix et le niveau de couverture (proche de celui du contrat public d'assurance obligatoire) étaient réglementés. Cette mesure, destinée à être transitoire, non seulement perdure mais a vu son champ d'application élargi. Après la remise à plat du système en 1986, il apparaît progressivement au régulateur que de nouvelles catégories de personnes se trouvent dans l'impossibilité de souscrire un contrat d'assurance privé. Les conditions d'accès au WTZ sont donc élargies, en 1986, à l'ensemble des personnes âgées et, en 1991, *"aux personnes dont la prime d'assurance santé auprès d'une compagnie privée est supérieure au montant maximal prévu par le contrat standard pour leur catégorie d'âge"* (selon les termes de la loi). Ce contrat couvre à l'heure actuelle environ 15% de la population qui s'assure auprès d'une compagnie privée et génère 45% des dépenses de santé du privé²⁹. On peut donc estimer ex post qu'il répond à un besoin des hauts risques. Toutefois, peu après l'instauration du WTZ, les assureurs privés qui le gèrent réalisent que la prime fixée par l'Etat ne permet pas de couvrir les dépenses des

²⁷ La prime prélevée au titre du MOOZ s'élève environ à 350 francs par an et par contrat privé.

²⁸ WTZ : Wet op de Toegang tot Ziektekostenverzekeringen.

²⁹ Ministère de la santé, entretien avec Mme. OKMA, 1998.

bénéficiaires. Les assureurs tentent d'élaborer un mécanisme privé de financement solidaire du déficit, mais les négociations échouent et des conflits apparaissent à la fois entre les compagnies d'assurance et avec l'Etat (Okma, 1997). Le parallèle avec le résultat selon lequel l'obligation faite aux assureurs privés d'offrir un contrat standard ne permet pas de rétablir l'optimum et déstabilise le marché est là encore intéressant.

Le financement du WTZ pose tant de problèmes aux assureurs privés que son existence se trouve menacée. L'Etat se voit alors contraint d'intervenir à nouveau pour transférer sur les assurés eux-mêmes la responsabilité du financement du déficit. Il instaure une surprime payée par tous les assurés du privé qui ne bénéficient pas du contrat standard. La surprime est déterminée par une commission indépendante en fonction des dépenses engendrées par les assurés du WTZ. Elle s'élève à environ mille francs par contrat privé pour 1997. L'équilibre financier est donc atteint au moyen d'une subvention des hauts risques par les bas risques qui impose la solidarité entre assurés au sein même du marché privé. La solution est tout à fait comparable à celle proposée par Crocker et Snow et rejoint la conclusion selon laquelle il n'est pas nécessaire que l'Etat intervienne comme assureur, mais il suffit qu'il mette en œuvre lui-même un mécanisme de subvention croisée tel que les assureurs n'ont pas intérêt à chercher à attirer les bas risques.

CONCLUSION

L'analyse standard développée dans le précédent chapitre a donc été illustrée ici par un certain nombre d'éléments empiriques. Plus précisément, nous avons montré que les éléments constitutifs de l'antisélection et que certaines prédictions sur la dynamique des marchés issues de la théorie peuvent être confortées par des faits stylisés que l'on observe sur le marché de l'assurance santé. En dernier lieu, nous avons présenté deux exemples d'intervention de l'Etat qui semblent illustrer les conclusions normatives du chapitre précédent.

Il n'en reste pas moins que l'interprétation proposée de certains de ces résultats n'est pas univoque. Ainsi, il est apparu que la dynamique du marché où l'antisélection prévaut pouvait conduire à la disparition des contrats que les hauts risques souhaitent souscrire, par un mécanisme de spirale de la mort. Ce résultat est conforme au sentiment assez largement partagé que dans les faits, les hauts risques rencontrent plus de difficultés que les bas risques à obtenir le niveau d'assurance qu'ils désirent, alors que le modèle de Rothschild et Stiglitz prédit, sauf exception, que ce sont les bas risques qui sont pénalisés par rapport à la situation d'information parfaite.

Atkinson (1995, pp. 209) critique plus sévèrement ce modèle quand il affirme : *"it is the high-risk group that has generated most of the concern about the shortcomings of market insurance: the problems they face in obtaining any insurance. The adverse selection model...does not explain these problems"*³⁰.

Dans le même ordre d'idée, les attendus qui fondent l'intervention de l'Etat sur le marché de l'assurance privée aux Pays-Bas relèvent d'une problématique d'accès à l'assurance pour des catégories de population qui, compte tenu de leur niveau de risque, ne parviennent pas à souscrire un contrat d'assurance santé.

³⁰ C'est le groupe des hauts risques qui a généré le plus d'inquiétudes par rapport aux limites de l'assurance de marché : les difficultés qu'ils éprouvent à obtenir une quelconque assurance. Le modèle d'antisélection ...n'explique pas ce problème.

Dans la lignée de l'analyse d'Akerlof, la préoccupation centrale du modèle canonique est de savoir comment une personne qui présente un bon niveau de risque (ou qui cherche à vendre une voiture d'occasion de bonne qualité) va pouvoir émettre un signal sur ce niveau de risque (de qualité). Si le mécanisme révélateur mis en place par les assureurs ne fonctionne pas, ils peuvent être amenés à sortir du marché. Le modèle décrivant les phénomènes d'antisélection dans le cadre de l'assurance s'interroge donc fondamentalement sur la possibilité pour les assureurs de garantir leur propre pérennité, tout en réalisant des transactions avec des assurés hétérogènes. Or, en assurance santé, la préoccupation fondamentale est plutôt celle de savoir si l'ensemble des risques et toute personne qui le souhaite peuvent être couverts. Arrow dans son analyse du marché de l'assurance santé (1963) insistait abondamment sur cet aspect : *"a great many risks are not covered and indeed the markets for the services of risk coverage are poorly developed or nonexistent (pp.945)...Certain groups - the unemployed, the institutionalized, and the aged- are almost completely uncovered"*³¹ (pp. 964).

Nous allons donc, dans la suite de ce travail, nous recentrer sur une analyse dans laquelle nous cherchons explicitement à déterminer les raisons qui font que certaines personnes n'auront pas accès à l'assurance.

³¹ *Un grand nombre de risque ne sont pas couverts et en fait les marchés de couverture des risques sont peu développés ou inexistant. Certains groupes, les chômeurs, les personnes vivant en institution et les personnes âgées sont presque totalement non couverts.*

CHAPITRE 6

LA SELECTION EN ASSURANCE SANTE

Lorsqu'on observe la dynamique générale des marchés d'assurance, il s'avère, comme le souligne Arrow dès 1963, que l'enjeu majeur est celui de l'accès à l'assurance.

Les Etats-Unis constituent l'exemple le plus connu du phénomène d'exclusion de l'assurance santé. L'assurance privée y constitue l'unique source de couverture pour la grande majorité de la population et l'accès à l'assurance est loin d'être universel : à l'heure actuelle plus de 15 % de la population totale ne bénéficie d'aucune couverture maladie. Cette proportion augmente au fil du temps (elle était de 13,5% en 1989), alors même que celle de la population couverte par Medicaid¹ a, elle aussi, augmenté sur la période (Hellander *et al.*, 1995). On sait, par ailleurs, que le débat politique sur l'introduction de l'assurance universelle est récurrent aux Etats-Unis (Diamond, 1992) et que malgré l'échec de la réforme Clinton qui visait à l'imposer aux Etats-Unis, la question reste d'actualité.

Nous nous proposons ici d'approfondir l'étude du phénomène de non assurance et de l'analyser à la lumière de la théorie économique. En effet, la théorie de la décision en incertain ne permet pas d'expliquer ce phénomène car elle prédit que, confronté à des primes d'assurance concurrentielles, tout individu qui éprouve de l'aversion pour le risque choisira de se couvrir complètement. Les américains sont certes libres de s'assurer, mais il est peu vraisemblable que 40 millions d'entre eux choisissent délibérément de ne pas se couvrir par manque d'aversion au risque. Dès lors que l'on réintroduit l'aversion pour le risque, l'assurance redevient optimale, même si elle est partielle quand les prix ne sont pas actuariels.

La théorie économique suggère alors, en s'écartant du modèle de base, deux principales explications au phénomène de non assurance.

¹ Programme public qui couvre certaines catégories de la population très défavorisées et définies selon des critères stricts. Medicaid couvre des personnes jugées particulièrement vulnérables (de jeunes mères isolées et leurs enfants pendant la période périnatale).

La décision de ne pas se couvrir peut, tout d'abord, s'expliquer par le fait que la prime d'assurance actuarielle représente une ponction trop élevée par rapport au revenu disponible de certains ménages. Cette hypothèse d'un choix contraint par le niveau de revenu est rarement prise en compte dans les modèles d'assurance qui postulent le plus souvent, à l'instar de celui de Rothschild et Stiglitz, une seule hétérogénéité des risques. Or, le revenu est une variable fondamentale d'explication de l'accès à l'assurance. Aux Etats-Unis toujours, on constate que les personnes non couvertes appartiennent majoritairement à des ménages dont les revenus sont faibles : 66% des personnes non assurées appartiennent à des ménages dont le revenu annuel est en dessous de 25.000 \$. Les personnes appartenant à des ménages dont le revenu est intermédiaire (<50.000 \$) semblent de plus en plus touchées et elles représentent une proportion croissante des personnes non couvertes (21% en 1989 et 24% en 1993, Hellander *et al.*, 1998). En France, le constat est identique pour l'assurance complémentaire, quoiqu'avec des conséquences moins dramatiques puisque l'essentiel du risque est couvert par la Sécurité sociale. Les personnes appartenant au décile des plus faibles revenus représentent 30% des personnes non assurées et seulement 6% des personnes assurées². Au travers de ces deux exemples, le revenu apparaît bien une variable explicative centrale pour l'accès à l'assurance.

La seconde explication de l'exclusion d'une partie de la population de la couverture maladie fait peser plus largement sur les assureurs la responsabilité du phénomène : ils éviteraient ou refuseraient de couvrir certains risques ou certaines personnes qui présentent des niveaux de risque élevé. Cette question de la sélection par les assureurs a suscité un regain d'intérêt dans le domaine de la santé depuis le début des années quatre-vingt-dix. Cependant, l'analyse rigoureuse des circonstances qui expliquent l'apparition de la sélection et ses conséquences reste incomplète.

Nous nous penchons successivement dans ce chapitre sur ces deux hypothèses. Tout d'abord, nous étudions l'exclusion de l'assurance dans le cadre d'un modèle simple qui prend en compte explicitement les hétérogénéités de risque et de revenu et dans lequel certaines personnes ne sont pas en mesure de souscrire une assurance (Section 1)³. Puis,

² A revenu équivalent prenant en compte la taille du ménage (calcul à partir des enquêtes SPS 1994 et 1995 du CREDES).

³ Ce premier travail a été réalisé en collaboration avec Jean-Charles Rochet, que je tiens à remercier ici très vivement. Il a donné lieu à la rédaction d'un document, "Couverture maladie ou accès gratuit aux soins", GREMAQ, Cahier N°9824-511.

nous proposons une analyse détaillée et critique du phénomène de la sélection différentielle qui a pour première conséquence l'exclusion de certains individus de l'assurance (Section 2). Ces deux phénomènes nous semblent en effet pouvoir cerner l'essentiel des explications de la non assurance. Nous nous attacherons à souligner en quoi cette double problématique nous semble fédératrice dans l'explication des réformes récentes des systèmes de santé des pays développés qui seront évoquées en dernière partie.

1. ACCES A L'ASSURANCE ET REVENU

Cette première section vise à étudier la question de l'exclusion de l'assurance maladie par le revenu, dans un marché où la souscription est volontaire. Une telle démarche prend tout son sens lorsqu'on compare un système de ce type avec un système dans lequel l'assurance maladie est universelle. Ce travail n'a pas de visée normative mais cherche à souligner les différences entre ces deux types de modèles. Nous présentons la problématique générale (1.1) et le modèle de base qui permet d'en rendre compte (1.2). Puis, nous élargissons le champ de l'analyse pour prendre en compte la décision de participer au marché du travail (1.3), avant de montrer rapidement en quoi les réformes récentes des systèmes de santé traduisent un arbitrage en faveur du modèle universel (1.4).

1.1 Accès aux soins, assurance volontaire et filet de sécurité

Le fait de n'être pas assuré a naturellement des conséquences significatives en termes d'accès aux soins. Une étude empirique américaine (Berk et Schur, 1998) portant sur des données de 1994 met clairement en lumière les enjeux de l'accès à l'assurance. Les auteurs observent trois sous-populations âgées de moins de 65 ans⁴ : les personnes qui ne bénéficient d'aucune assurance, celles qui sont couvertes par Medicaid et celles qui sont couvertes par une assurance privée. Trois indicateurs d'accès aux soins sont étudiés : le fait de déclarer avoir une "source de soins habituelle", le fait d'avoir dû renoncer à des soins au cours de l'année précédente et le nombre moyen de consultations chez un médecin de ville pour la même période. Les personnes non assurées ont une probabilité

⁴ Les personnes de plus de 65 ans sont couvertes par Medicare.

nettement plus faible d'avoir un lieu de recours habituel que les personnes couvertes et ont quasiment deux fois moins de consultations lorsqu'elles sont en mauvaise santé. Les résultats sur le renoncement à des soins vont globalement dans le même sens, même s'ils révèlent que l'amélioration de l'accès que permet Medicaid n'amène pas ses bénéficiaires au niveau des personnes couvertes de façon privée. Une autre étude (Weinick *et al.*, 1997) montre que les personnes qui n'ont pas d'assurance ont une probabilité trois fois plus élevée que les autres d'avoir des difficultés d'accès aux soins et de devoir attendre pour pouvoir bénéficier de soins dont elles perçoivent le besoin. Dans le cas de la France, les travaux issus de l'enquête santé INSEE concluent dans le même sens (Genier, 1998, Raynaud, 1998).

L'accès à l'assurance conditionne donc largement l'accès régulier aux soins. Pour autant, les personnes qui ne sont pas couvertes par une assurance ont, dans les pays développés, toujours accès à des soins en cas de grande nécessité (matérielle et médicale). Les hôpitaux publics et les médecins prennent en charge les personnes démunies et les Etats américains, pour leur part, mettent souvent en place des structures spécifiques pour faire face aux besoins de santé les plus aigus des populations non couvertes. Le système sanitaire, y compris aux Etats-Unis, finit donc généralement par prendre en charge les patients qui ne bénéficient pas d'assurance (Baxter et Mechanic, 1997). Un tel mode de prise en charge est critiqué en France car, "*la mise en place de lieux de traitement des plus défavorisés, palliatifs utiles, maintient l'exclusion du droit commun*" (Rapport Boulard, 1998).

Ce mode de prise en charge des soins soulève d'autres difficultés : les personnes non assurées n'étant pas suivies de façon régulière, leurs besoins sont exprimés plus tardivement. Lorsqu'elles sont finalement prises en charge par la collectivité, ces besoins risquent de se traduire par une prise en charge plus coûteuse. Ainsi, en France⁵, les personnes qui ne bénéficient pas d'assurance complémentaire semblent avoir des niveaux de dépense et des modes de recours aux soins différents de ceux des personnes couvertes. La dépense moyenne, en 1995, des personnes non couvertes par une assurance est de 17% plus élevée que celle des personnes couvertes. De plus, alors que le taux d'hospitalisation par année est comparable entre les populations couvertes et non couvertes, la dépense par personne hospitalisée couverte est quasiment deux fois moins

⁵ Enquête appariée SPS-EPAS 1995, échantillon de 4505 personnes (calcul de l'auteur).

élevée que celle d'une personne non couverte. Ce différentiel peut s'expliquer par le recours plus tardif à l'hospitalisation qui entraînera des dépenses plus élevées. Ce report n'a pas en France pour seule cause l'absence d'assurance : l'avance de frais pour les soins de ville constitue une barrière largement aussi importante pour les personnes dont le revenu est faible. Ceci dit, le renoncement aux soins pour l'optique, les prothèses dentaires ou auditives s'explique, lui, avant tout par l'absence d'assurance complémentaire en raison des taux de couverture très bas de la Sécurité sociale. Comme le souligne Boulard⁶ "*Les restrictions d'accès à ces soins apparaissent d'autant plus préoccupantes qu'elles touchent fortement à l'insertion sociale, professionnelle et tout simplement humaine*".

Outre une augmentation probable du coût de la prise en charge, un tel système introduit des distorsions de prix. Aux Etats-Unis, la pratique est connue sous le terme de "*cost shifting*" : les tarifs payés par les assureurs privés aux prestataires de soins sont surévalués car ces derniers reportent sur leur clientèle couverte et/ou aisée le coût de la prise en charge des personnes non solvables. Remarquons toutefois que le contrôle croissant des coûts par les assureurs rend cette stratégie de plus en plus délicate. En conséquence, médecins et hôpitaux sont peu encouragés à soigner les personnes non solvables et l'équilibre des filets de sécurité est précaire. "*The competitiveness of local markets greatly affects the ability to cross subsidize indigent care and other safety net functions*"⁷ (Baxter et Mechanic, 1997).

L'ensemble de ces préoccupations a marqué les débats autour de la réforme Clinton. Elles font aussi écho à un débat français d'actualité, de nature comparable, mais dont les enjeux sont moindres et qui concerne l'introduction d'une couverture maladie universelle. Citons le rapport Boulard : "*Avec la couverture maladie universelle, il ne s'agit pas de consolider l'existence d'une filière sanitaire des pauvres s'articulant autour de l'hôpital public et des centres de santé mais d'ouvrir aux personnes en difficulté le système de soins de tous*"⁸.

⁶ Boulard, J.-C. 1998.

⁷ *Le degré de concurrence des marchés locaux affecte profondément la capacité d'effectuer des subventions croisées au bénéfice des soins proposés aux indigents et des autres éléments des filets de sécurité.*

⁸ Ibid.

Le présent travail permet de souligner les enjeux de l'introduction d'une telle couverture, et nous permettra de proposer une explication du fait que certains pays sont parvenus à introduire l'assurance universelle.

1.2 Modèle de base

Dans une première étape, nous faisons l'hypothèse que les revenus sont exogènes et qu'il existe une prise en charge ex post par le système sanitaire des dommages de personnes non couvertes par une assurance privée. De plus, cette prise en charge n'est pas plus coûteuse que la prise en charge du dommage pour une personne qui bénéficie d'une assurance ex ante. Nous montrons alors que l'existence d'un tel "filet de sécurité" conduit certaines personnes à ne pas s'assurer, alors même qu'elles en ont les moyens. Pour reprendre les termes de Besley et Gouveia (1994, pp. 228) : *"If some individuals in society (for example the poor) anticipate that the rich will provide them with care whether or not they insure, this will undermine their incentive to buy insurance"*⁹.

1.2.1 Assurance et contrainte budgétaire

Les individus ont deux caractéristiques exogènes : un revenu R (interprété comme le revenu disponible au-delà d'une consommation minimale incompressible) et une probabilité p de subir un dommage D , dont le montant est très élevé par rapport à la moyenne des revenus de la population. On peut par exemple considérer que ce montant correspond au coût d'une hospitalisation de longue durée.

Les individus maximisent l'espérance d'une fonction d'utilité de type von Neumann-Morgenstern, notée U , supposée croissante ($U' > 0$) et concave ($U'' < 0$).

Le marché de l'assurance privée est supposé parfait et chaque individu peut acquérir la quantité d'assurance q qu'il souhaite à un prix unitaire actuariel p .

Si un individu n'est pas assuré de façon privée, il est pris en charge par l'Etat dans le cas où il subit un dommage et sa consommation est alors normée à zéro. Autrement dit, l'Etat prend en charge la part du dommage que l'individu ne peut couvrir par son revenu disponible.

⁹ Si certains individus dans la société, par exemple les pauvres, anticipent que les riches leur fourniront des soins qu'ils s'assurent ou non, cela diminue leur incitation à souscrire une assurance.

Considérons, dans un premier temps, le cas où il n'y a pas de contrainte financière à l'accès à l'assurance.

Chaque individu maximise son espérance d'utilité V par rapport au niveau de couverture q :

$$V(p, R, q) = (1 - p)U(R - pq) + pU(R - pq + q - D)$$

La solution est évidemment l'assurance complète. La consommation finale de l'individu est $R - pD$ et l'utilité individuelle $U(R - pD)$.

Considérons maintenant le cas où le revenu de certains individus ne leur permet pas de souscrire une assurance complète.

Si $R < pD$, les individus pris en charge par l'Etat en cas de dommage n'achètent pas d'assurance (même partielle) et ont alors l'espérance d'utilité suivante :

$$V(p, R) = (1 - p)U(R) + pU(0)$$

Nous noterons $R_0(p)$ le revenu minimum permettant à un individu de risque p de se couvrir : $R_0(p) = pD$.

R_0 est une fonction croissante de p dans le repère (p, R) : plus la probabilité de dommage est élevée, plus la prime augmente, et l'individu souhaitant se couvrir doit disposer d'un revenu plus important. Quand p tend vers 1, aucune personne ayant un revenu inférieur à D ne peut s'assurer.

1.2.2 Renoncement à l'assurance

Montrons que cette possibilité de prise en charge ex post des soins par l'Etat incite certains individus à ne pas s'assurer sur le marché privé, alors même qu'ils en ont les moyens.

En effet, cette prise en charge implique que la fonction d'utilité des agents soit en réalité :

$$V(p, R) = \text{Max}\{U(R - pD), (1 - p)U(R) + pU(0)\}$$

Comme les primes sont actuarielles, la quantité optimale d'assurance est toujours 0 ou D .

Sans perte de généralité, nous poserons $U(0) = 0$.

Le choix des individus est alors dicté par le signe de la fonction :

$$\Phi(R, p) = (1 - p)U(R) - U(R - pD)$$

Plus précisément, l'individu de caractéristiques R (revenu) et p (risque) s'assure si et seulement si :

$$\Phi(R, p) \leq 0.$$

a) Existence de trois catégories de ménages

Proposition 1:

Dans un système d'assurance maladie privée, la prise en charge par l'Etat des soins des ménages les plus démunis désincite certains ménages à s'assurer. Plus précisément, il existe, pour tout niveau de risque p un revenu critique $R^(p)$ (supérieur au coût de l'assurance $R_0(p) = pD$ et éventuellement infini) en deçà duquel les ménages renoncent à s'assurer.*

Si $D < \lim_{R \rightarrow +\infty} \frac{U(R)}{U'(R)}$, $R^(p)$ est fini pour tout p .*

On distingue alors trois catégories de ménages :

- *les ménages aux ressources insuffisantes pour s'assurer. Ils sont caractérisés par un revenu R tel que :*

$$R \leq R_0(p) = pD,$$

- *les ménages renonçant à l'assurance. Ils sont caractérisés par un revenu R tel que :*

$$R_0(p) < R < R^*(p),$$

- *les ménages qui s'assurent, dont le revenu est tel que:*

$$R \geq R^*(p).$$

Démonstration:

Notons $\Phi(R, p)$ la différence entre l'utilité d'un ménage non assuré et celle d'un ménage assuré :

$$\Phi(R, p) = (1 - p)U(R) - U(R - pD)$$

Remarquons tout d'abord que si $R = pD$, alors $\Phi(R, p) = (1 - p)U(R) > 0$. Autrement dit, un ménage qui a tout juste les moyens de payer son assurance préfère recourir à l'aide de l'Etat.

De plus, si $R = 0$, $\Phi = U(0) - U(-pD) > 0$. Un ménage dont le revenu est nul choisira toujours l'aide de l'Etat.

D'autre part,

$$\frac{d\Phi}{dR} = (1-p)U'(R) - U'(R-pD) < U'(R) - U'(R-pD) < 0$$

Cette dernière inégalité vient du fait que U' est décroissante (aversion au risque).

La fonction Φ est donc décroissante de R . Pour un niveau de risque donné, la tentation de ne pas s'assurer diminue quand le revenu augmente.

Enfin :

$$\frac{\Phi(R, p)}{U(R)} = \frac{U(R) - U(R-pD)}{U(R)} - p = pD \left[\frac{U'(R_1)}{U(R)} - \frac{1}{D} \right] \quad (2)$$

où R_1 est dans l'intervalle $[R-pD, R]$. La deuxième égalité découle du théorème des accroissements finis appliqué à la fonction U entre $R-pD$ et R .

U étant croissante et concave, la fonction $\frac{U'(R)}{U(R)}$ est décroissante positive. Elle a donc

une limite ≥ 0 quand R tend vers l'infini. Cette limite est aussi celle de $\frac{U'(R_1)}{U(R)}$ quand R tend vers l'infini.

- Si cette limite est supérieure à $\frac{1}{D}$, alors la formule (2) montre que Φ est positive pour tout R . Le comportement de renoncement à l'assurance est total : aucun ménage ne s'assure. Toutefois, ce cas extrême ne peut se produire que si $\frac{U'(R)}{U(R)}$ a une limite non nulle en $+\infty$ et si D est assez élevé ;
- Si cette limite est $< \frac{1}{D}$ alors il existe un revenu R^* critique (dépendant de p) tel que $\Phi(R, p) > 0 \Leftrightarrow R < R^*(p)$.

b) Comportement de renoncement

Étudions maintenant les propriétés de la fonction $R^*(p)$. On suppose donc que

$$D < \lim_{R \rightarrow +\infty} \frac{U(R)}{U'(R)}.$$

Proposition 2

La fonction $R^*(.)$ est croissante sur $[0,1]$. Ses valeurs extrêmes sont caractérisées comme suit :

$$R^*(0) \text{ est la solution de } \frac{U}{U'}(R) = D ;$$

$$R^*(1) = D.$$

Démonstration

En $R^*(p)$, la valeur de la fonction Φ est :

$$\Phi(R^*(p), p) = (1-p)U(R^*(p)) - U(R^*(p) - pD) \quad (3)$$

La dérivée de (3) en $R^*(p)$, multipliée par p , vaut :

$$p \frac{\partial \Phi}{\partial p} [R^*(p), p] = [DU'(R^*(p) - pD) - U(R^*(p))]p \quad (4)$$

Par définition, en R^* , le ménage est indifférent entre le fait de s'assurer ou non :

$$\Phi(R^*(p), p) = 0 \text{ et donc :}$$

$$pU(R^*(p)) = U(R^*(p)) - U(R^*(p) - pD)$$

Par substitution dans (4) on obtient :

$$p \frac{\partial \Phi}{\partial p} [R^*(p), p] = pDU'(R^*(p) - pD) - U(R^*(p)) + U(R^*(p) - pD)$$

La concavité de U implique que :

$$U'(R^* - pD) > \frac{U(R^*) - U(R^* - pD)}{pD}$$

On a donc :

$$p \frac{\partial \Phi}{\partial p} [R^*(p), p] > 0$$

R^* est une fonction croissante de p . Plus leur niveau de risque est élevé, plus les ménages arbitrent en faveur d'une prise en charge ex post par l'Etat.

En $p = 1$, on voit que $\Phi(R, p) = 0$ si et seulement si $R^*(1) = D$.

La limite de la fonction en 0 se calcule à partir d'un développement limité.

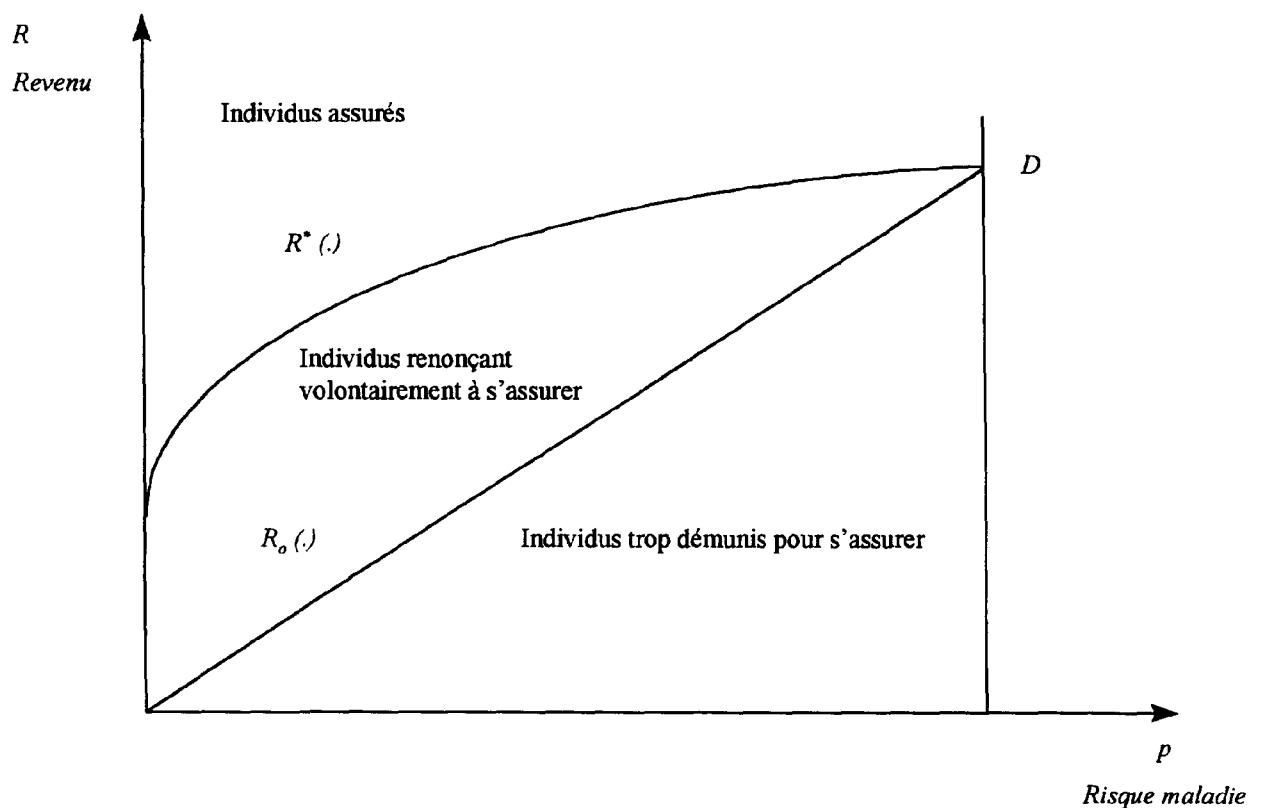
En $p = 0$:

$$\Phi(R, p) \sim p[DU'(R) - U(R)] = 0 \Leftrightarrow \frac{U}{U'}(R^*(0)) = D$$

Il existe donc trois catégories de ménages : les ménages insolubles dont les soins sont pris en charge ex post par le système public, les personnes qui adoptent un comportement stratégique et ne s'assurent pas alors qu'elles en ont les moyens et enfin les personnes qui s'assurent dans le privé.

La figure 1 illustre la proposition 2.

Figure 1 :
Décision d'assurance en fonction du niveau de risque et de revenu



Une parade classique au comportement de renoncement est l'obligation d'assurance, appliquée à la responsabilité civile en assurance automobile. Bien que des infractions à cette loi soient constatées régulièrement dans la plupart des pays, il n'y a pas a priori d'obstacle financier à ce que cette réglementation soit appliquée en matière automobile, dans la mesure où les personnes ayant fait l'acquisition d'une automobile ont, en principe, les moyens de payer aussi leur assurance responsabilité civile. Il en va tout autrement pour l'assurance maladie : la plupart des sociétés développées défendent le principe d'un accès universel aux soins de base alors même qu'une fraction non négligeable de la population n'a pas les moyens de faire face au coût de son assurance maladie. Autrement dit, même si les individus renonçant volontairement à s'assurer pouvaient théoriquement être forcés à le faire, cette obligation d'assurance ne résoudrait pas le problème des ménages les plus démunis.

Nous allons désormais comparer deux systèmes de façon plus formelle :

- dans le premier, l'assurance est volontaire et il existe un filet de sécurité au terme duquel les personnes non couvertes sont prises en charge par le système sanitaire ;
- dans le second, l'assurance est obligatoire et universelle. Nous introduisons un élément fondamental du débat, à savoir la présence de désincitations à participer au marché du travail.

1.3 Participation au marché du travail

La mise en place d'un système d'assurance maladie universelle peut en effet avoir des effets sur la participation au marché du travail. Plusieurs économistes américains ont avancé des éléments empiriques en ce sens. Par exemple, pour Cutler (1994, pp.20) *"empirical estimates suggest that up to one-quarter of the approximately 4 million welfare recipients would enter the labor force if health insurance were available continuously"*¹⁰. De même Newhouse (1994, pp.9) considère que *"the loss of Medicaid benefits [for welfare mothers who start working] is a disincentive to work"*¹¹.

¹⁰ Les estimations empiriques suggèrent que jusqu'à un quart des 4 millions environ de personnes bénéficiant de prestations sociales entreraient sur le marché du travail si l'assurance maladie y était disponible en continu.

¹¹ La perte du bénéfice de Medicaid [pour les mères qui bénéficient de l'assistance sociale] est une désincitation à travailler.

1.3.1 Comportement des individus

Nous allons maintenant endogénéiser la participation au marché du travail : chaque individu peut renoncer à travailler, auquel cas son revenu devient nul. R s'interprète désormais comme un revenu potentiel. On doit donc comparer trois niveaux d'utilité :

- $U_A = U(R - c_I - pD) - \gamma$ pour un actif assuré où c_I est une cotisation (supposée indépendante du revenu) destinée à financer l'aide médicale gratuite, et γ représente la désutilité du travail ;
- $U_N = (1 - p)U(R - c_I) - \gamma$ pour un actif non assuré ;
- $U_I = 0$ pour un inactif.

Ces trois niveaux d'utilité correspondent à trois zones du plan (p, R) que nous allons maintenant caractériser. Nous supposons que la désutilité du travail n'est pas trop forte $\gamma < U(R^*(0))$, où $R^*(0)$ est donné implicitement par $\frac{U}{U'}(R^*(0)) = D$.

Proposition 3

Dans un système d'assurance privée avec aide médicale gratuite des plus démunis, la population se répartit en trois catégories :

- *Les actifs assurés (zone A) caractérisés par un revenu net $R - c_I$ supérieur à $\max(R^*(p), pD + U^{-1}(\gamma))$;*
- *Les actifs non assurés (zone N) dont le revenu net est compris entre $U^{-1}\left(\frac{\gamma}{1-p}\right)$ et $R^*(p)$;*
- *Les inactifs (zone I) , correspondant au reste de la population.*

Démonstration

La zone A correspond aux individus qui sont caractérisés par les deux inégalités :

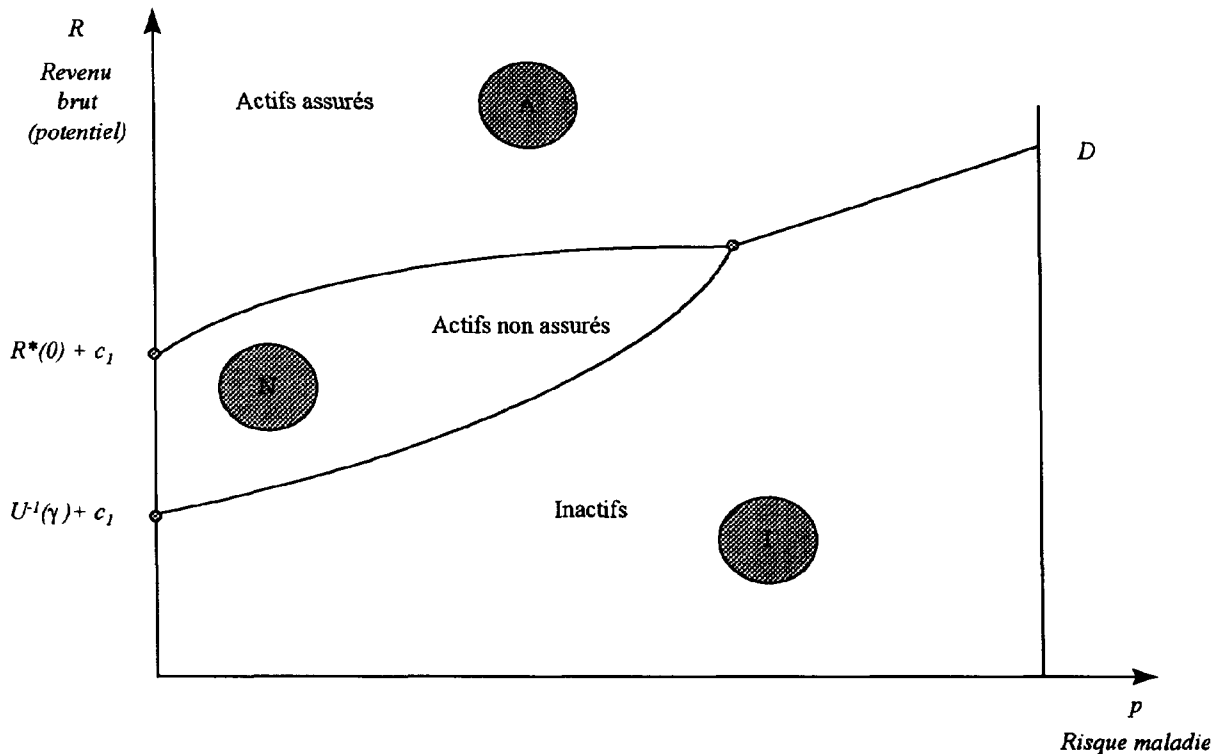
$U_A \geq U_N$ (caractérisée par $R - c_I \geq R^*(p)$, voir proposition 1) et

$U_A \geq U_I$ qui équivaut à $R - c_I \geq pD + U^{-1}(\gamma)$.

Le reste de la démonstration découle de façon évidente de la comparaison des équations définissant les niveaux d'utilité des différentes catégories d'individus.

Les trois zones sont représentées dans la figure 2 ci-après.

Figure 2 :
Assurance et activité en fonction du niveau de risque et de revenu
(Système privé)



1.3.2 Introduction de l'assurance universelle

Nous allons maintenant analyser l'impact qu'aurait une réforme introduisant une couverture maladie universelle, c'est à dire une assurance de tous les ménages, financée par une cotisation c_2 , supposée, là encore, indépendante du revenu¹² et prélevée sur l'ensemble des actifs.

L'espérance d'utilité devient alors indépendante de la probabilité de dommage :

$$U = \max(U(R - c_2) - \gamma, 0)$$

¹² Nous n'étudions pas ici la dimension redistributive des cotisations d'assurance maladie, qui dépendent en général du revenu. Cette question est examinée dans Henriet et Rochet (1998).

Le niveau de c_2 est déterminé par l'équilibre général du système. Il correspond au niveau pour lequel le coût unitaire moyen de la couverture est couvert par le produit du montant de la cotisation c_2 et de la proportion d'assujettis (les actifs) :

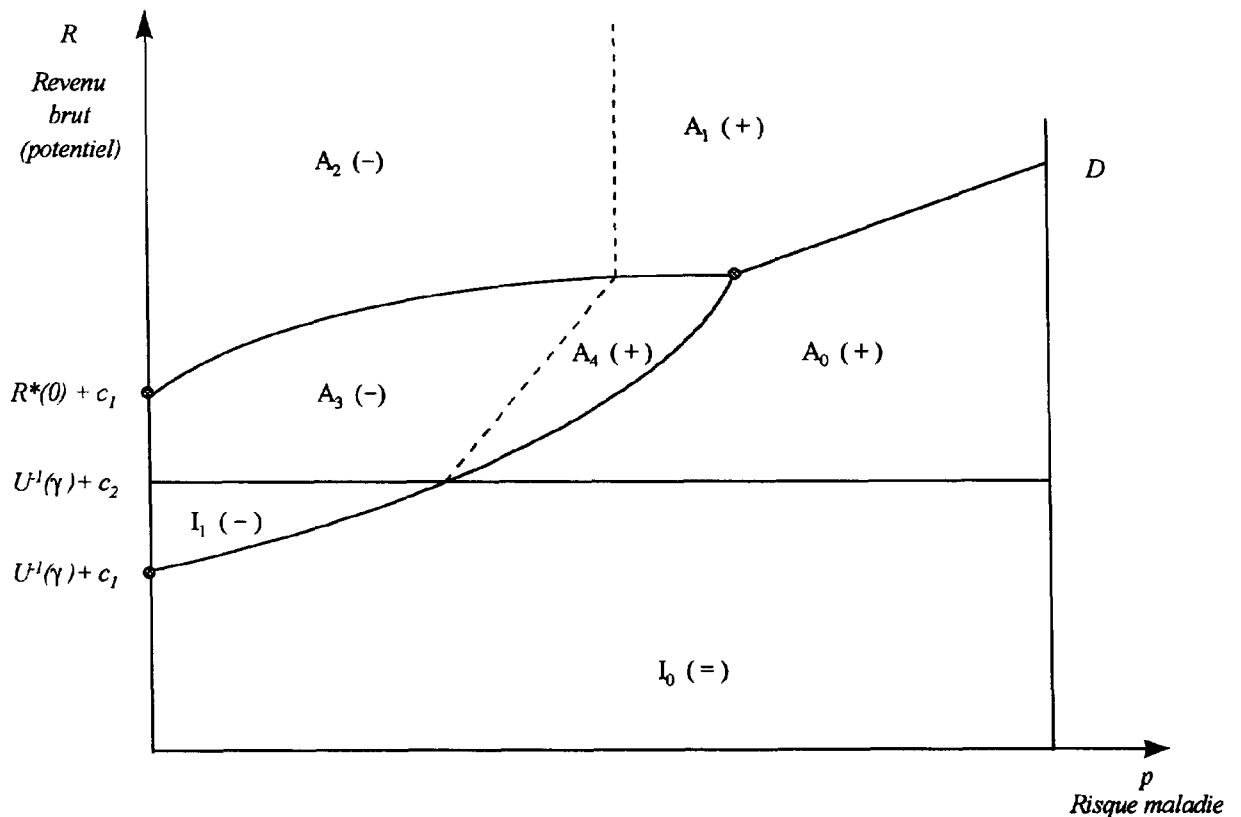
$$p_m D = c_2 (1 - G(c_2 + U^{-1}(\gamma)))$$

où p_m est la moyenne de p dans la population et G la fonction de répartition de R .

c_2 est donc vraisemblablement plus élevé que c_1 qui ne correspond qu'à la couverture du coût beaucoup plus faible de l'aide médicale gratuite. Par contre, on peut légitimement supposer que D est suffisamment élevé pour que $R^*(0) + c_1 > U^{-1}(\gamma) + c_2$.

L'impact de la réforme est représenté sur la figure 3 ci-après.

Figure 3 :
Les conséquences de l'introduction d'une assurance maladie universelle
(+ désigne les ménages gagnants, - les perdants)



Comme nous l'avons remarqué, la séparation entre actifs (A) et inactifs (I) devient indépendante du niveau de risque p : la zone A est désormais caractérisée par :

$$R > U^{-1}(\gamma) + c_2$$

A l'intérieur de la zone I, on distingue les individus que la réforme a désincités à travailler (I_1) et les autres (I_0), dont la situation est inchangée.

Au sein de la zone A, les arbitrages sont plus complexes :

- la région A_0 correspond aux personnes qui, à la différence de I_1 ne travaillaient pas, par peur de perdre leur droit à l'aide médicale gratuite, mais que la réforme incite maintenant à travailler,
- la région A_1 correspond aux personnes actives et assurées dont la cotisation globale a diminué $c_2 < c_1 + pD$,
- la région A_2 correspond aux personnes actives et assurées dont la cotisation globale a augmenté $c_2 > c_1 + pD$,
- enfin les régions A_3 et A_4 correspondent aux personnes qui n'étaient pas assurées mais qui le deviennent d'autorité. Certaines y gagnent (A_4), d'autres y perdent (A_3) la frontière entre les deux régions (en pointillé) étant déterminée par l'égalité :

$$U(R - c_2) = (1 - p)U(R - c_1)$$

A l'aide d'un modèle très simple, nous avons analysé l'impact qu'aurait, dans un pays où l'assurance privée concerne la majorité de la population comme les Etats-Unis, le passage à une assurance maladie universelle financée par une cotisation forfaitaire payée par tous les actifs. Comme l'on pouvait s'y attendre, les ménages qui auraient à perdre d'une telle réforme sont ceux dont le revenu est élevé et le risque faible. En fait, notre analyse permet d'affiner ce résultat intuitif en distinguant trois types de profils perdants :

- les ménages à revenu élevé et risque faible (zone A_2) dont la cotisation d'assurance augmente ;
- les ménages à revenu moyen et risque faible (zone A_3) qui sont forcés de s'assurer contre leur gré ;
- les ménages à revenu et risque faibles (zone I_1), qui sont incités à ne plus travailler.

Par contre, les autres ménages bénéficient de la réforme. Ce sont :

- les ménages à risque élevé et revenu élevé (zone A_1) dont la cotisation d'assurance baisse ;
- les ménages à risque élevé et revenu moyen (zone A_0) qui sont désormais incités à entrer sur le marché du travail ;
- les ménages à risque et à revenu moyen (zone A_4) qui bénéficient désormais d'une couverture maladie.

1.4 Tendances actuelles des réformes

Les évolutions récentes des systèmes de santé soulignent la pertinence de ce travail et permettent d'illustrer les choix politiques dans des systèmes aussi différents que ceux de la France et des Etats-Unis.

Aux Etats-Unis, la réforme Clinton (Cutler, 1994c) avait pour principal objectif d'instaurer l'assurance universelle. Au terme de cette réforme, de nouvelles agences, les *Health Alliances*, monopoles locaux, sont désignées comme unique intermédiaire entre les assurés et les assureurs. Seules les firmes de plus de 5000 employés peuvent s'affranchir de l'autorité des alliances. Au sein d'une HA, les assurés peuvent choisir entre différentes formes d'assurance (HMO, PPO, et assurance traditionnelle). Tous les contrats proposés doivent couvrir les mêmes biens et services mais les participations financières des usagers sont différenciées pour prendre en compte la façon dont fonctionnent traditionnellement les différents types d'organismes qui proposent la couverture. La participation financière des employeurs au financement de l'alliance est obligatoire et standardisée à environ 80%¹³ de la moyenne des primes proposées au sein de l'alliance, le reste étant financé par les ménages eux-mêmes.

Afin de garantir l'accès à l'assurance de tous, des subventions sont accordées pour les travailleurs indépendants et aux ménages dont le revenu est inférieur à 150% du seuil de pauvreté.

¹³ La participation est plafonnée à 8% des coûts salariaux, moins pour les PME.

L'échec de cette réforme s'explique probablement en partie par l'arbitrage réalisé au Congrès entre les intérêts des différentes populations que nous évoquons dans le modèle précédent et dont on se doute que la représentation politique n'est pas nécessairement assurée de façon proportionnelle.

En France, le projet d'introduction d'une couverture maladie universelle trouve aussi un écho dans l'analyse qui précède. La très grande majorité de la population bénéficie d'une couverture sociale. Quelques personnes (dont le nombre n'est pas connu avec certitude¹⁴) restent exclues non parce qu'elles n'ont pas droit à la Sécurité sociale, que parce qu'elles ne sont pas en mesure de faire valoir ce droit. Toute personne résidant en France et ne pouvant pas faire face au coût des soins a droit, en outre, à l'aide médicale du département dans lequel elle réside ou à l'aide médicale de l'Etat, pour les sans domicile fixe. L'AMG peut se traduire par la prise en charge des cotisations à l'assurance personnelle¹⁵, des tickets modérateurs et du forfait journalier. L'admission se fait sous condition de ressources (dont le plafond varie selon les départements) mais tous les allocataires du Revenu Minimum d'Insertion et de l'allocation de veuvage en bénéficient de plein droit. Malgré l'existence de ce dispositif, certaines personnes démunies ne parviennent pas à bénéficier du même accès aux soins que la population générale, notamment en ce qui concerne les soins de ville et les soins dentaires.

La nécessité d'offrir à tous l'égalité des prestations étant plus consensuelle qu'aux Etats-Unis, le gouvernement Français envisage actuellement une réforme de la protection sociale. Le projet de loi portant création d'une Couverture Maladie Universelle¹⁶ impose tout d'abord une généralisation complète de la couverture de base pour les personnes résidant de façon stable et régulière en France. Le second volet de la future loi permet aux personnes dont les revenus sont plus faibles de bénéficier d'une couverture complémentaire en matière de santé et de la dispense d'avance de frais.

Ce nouveau dispositif concerne une proportion plus grande de la population que l'aide médicale et se substitue à l'ancien dispositif. Les personnes admises bénéficieront

¹⁴ Environ 150000. Liaisons Sociales Quotidien, Jeudi 11 mars 1999, N°20/99. Cahier joint au numéro 12867.

¹⁵ Régime d'assurance facultative, dont les cotisations sont prises en charge par un tiers (Etat, Caisse d'Allocation Familiale, Fonds de solidarité vieillesse...) dans 90% des cas.

¹⁶ Conseil des ministres du 3 mars 1999.

gratuitement d'une couverture des tickets modérateurs et du forfait journalier, mais aussi d'une prise en charge des dépenses liées aux prothèses dentaires et à l'optique. Cette réforme doit cependant éviter un certain nombre d'écueils, en particulier liés aux effets de seuil. D'ores et déjà, il est prévu que les personnes qui, avant de bénéficier de la protection complémentaire gratuite en matière de santé, étaient titulaires d'une protection complémentaire de plein droit, pourront obtenir la résiliation de ces contrats ou leur transformation en garanties conformes à la nouvelle loi. Il est difficile d'estimer le nombre de personnes concernées, mais il faudra, pour évaluer pleinement l'impact de la loi, déterminer dans quelle mesure elle donnera lieu à des arbitrages, tant chez les bénéficiaires que chez leurs employeurs. Les premiers risquent à terme de perdre le bénéfice de la couverture gratuite en acceptant un emploi précaire¹⁷. Les seconds pourraient renoncer à offrir un contrat de groupe à des personnes qu'elles savent protégées par ailleurs.

En dernière analyse, on peut souligner depuis les années quatre-vingt, une tendance à l'universalisation de l'assurance maladie qui a concerné successivement les pays latins, l'Australie, Israël et la Suisse. Comme le rappelle van de Ven (1996, pp.656) *"no country is reducing the number of persons covered by mandatory health insurance"*¹⁸. Les avantages d'une couverture universelle, qui peut être articulée autour de mécanismes concurrentiels, semblent donc faire l'objet d'une adhésion croissante.

2. ACCES A L'ASSURANCE ET SELECTION DIFFERENTIELLE

Peut-on expliquer les difficultés d'accès à l'assurance dans un système concurrentiel par un seul effet revenu ? La réponse est évidemment négative et on peut plus précisément s'interroger sur le fait que les assureurs acceptent de couvrir toute personne qui en fait la demande, quand bien même les individus seraient prêts à payer des primes actuarielles élevées. Les économistes qui étudient le fonctionnement des marchés de services liés à la santé se sont penchés ces dernières années sur l'hypothèse selon laquelle certains producteurs de services, dont les assureurs peuvent avoir, dans un univers concurrentiel, intérêt à mettre en œuvre une sélection différentielle des risques. Cette sélection consiste

¹⁷ La couverture est maintenue pendant un an si les conditions de ressource sont dépassées.

¹⁸ *Aucun pays ne réduit le nombre de personnes couvertes par l'assurance obligatoire.*

en un ensemble de stratégies mises en œuvre par des agents économiques producteurs d'un service, destinées à exploiter des hétérogénéités de risque mal tarifées. Elle se traduit par le fait que les hauts risques n'obtiennent pas le service qu'ils souhaiteraient acquérir (Newhouse, 1996). Cette première définition, déclinée à l'assurance santé, décrit la sélection comme une stratégie de l'assureur qui a pour conséquence principale l'exclusion des personnes qui présentent des niveaux de risque élevés. Nous nous attachons dans un premier temps à clarifier cette définition (2.1) avant de souligner l'ensemble des conséquences préjudiciables de l'existence ou de la généralisation de telles stratégies (2.2). Nous concluons cette démarche en montrant rapidement comment la sélection est devenu un enjeu central dans la gestion des systèmes de santé (2.3).

2.1 Définition de la sélection différentielle des risques

La sélection différentielle est une réaction stratégique de l'assureur qui présente un intérêt lorsque certaines conditions sont réunies (2.1.1). Les moyens dont les assureurs disposent pour mettre en œuvre une telle stratégie sont variés (2.1.2). L'analyse des circonstances et des outils de la sélection nous conduira à clarifier la définition du concept de sélection et à préciser son articulation avec l'antisélection (2.1.3).

2.1.1 Les circonstances de la sélection

Dans quelles circonstances observe-t-on la sélection ? La sélection est une stratégie de réponse d'un prestataire de service qui reçoit un paiement prospectif imparfaitement ajusté au risque (a). En matière d'assurance, on considère aussi souvent que la sélection est induite par la réglementation des primes (b).

a) Paiement prospectif imparfaitement ajusté au risque

La sélection émerge lorsque plusieurs conditions sont réunies.

1. Tout d'abord, il faut :

- d'une part, qu'un prix d'échange contractuel soit fixé ex ante pour le service considéré ;
- d'autre part, que le coût de production du service varie en fonction des caractéristiques du bénéficiaire.

Si le coût de production pour chaque bénéficiaire ou type de bénéficiaire est parfaitement anticipé et reflété dans le prix, le marché fonctionne normalement. Il en va de même si le paiement est décidé après que le service a été produit et qu'il se fonde sur les coûts de production observés pour chaque individu. En revanche, si les caractéristiques du bénéficiaire qui expliquent les différences de coût de production ne sont pas toutes prises en compte dans le prix payé au producteur, ce dernier a rationnellement intérêt à sélectionner les risques. Dans ce cas en effet, le paiement prospectif est uniforme sur une population (ou une sous-population) hétérogène et il reflète en général le coût moyen de production pour ce groupe.

2. Si le producteur est en mesure d'anticiper le coût de production pour différents individus au sein de ce groupe, il peut générer un profit certain en fournissant le service à un groupe de personnes pour lesquelles le coût total de production est inférieur au chiffre d'affaires basé sur le prix moyen. L'existence même d'un prix moyen incite le prestataire à mettre en place une sélection différentielle qui lui permet d'améliorer son profil de risque par rapport à celui sur lequel le calcul du prix moyen est fondé.

Certaines des conditions nécessaires à l'émergence de la sélection sont réunies par construction sur un marché d'assurance concurrentiel (paiement prospectif¹⁹ et responsabilité financière). Si la tarification au risque n'est pas actuarielle, l'assureur, qui reçoit un paiement uniforme pour des risques qui ne le sont pas, peut chercher à exploiter des informations dont il dispose initialement ou qu'il accumule au fil du temps sur le niveau de risque de tel ou tel assuré.

Le terme de sélection différentielle recouvre plusieurs types de stratégies qui consistent pour le producteur à :

- éviter de couvrir les individus dont il sait qu'ils présentent des niveaux de risque élevés ;
- chercher à évincer ceux dont il apprend, après signature d'un contrat et avant un possible renouvellement, qu'ils sont dans ce cas ;

¹⁹Par prospectif, on entend essentiellement que le prix du service est déterminé avant que le service ne soit rendu, quand bien même il peut théoriquement faire l'objet d'un versement en fin de période. Dans le cas de l'assurance, le paiement prospectif est la base même du fonctionnement du marché.

- ou au contraire tenter d'attirer les individus pour lesquels il sait par avance que le montant du financement excède le coût de production, ou dans le cas précis de l'assurance, l'espérance de dommage. Cette dernière modalité de la sélection différentielle est qualifiée d'écrémage (*cream skimming, cherry picking*).

Ce terme d'écrémage des risques est parfois utilisé pour rendre compte de l'ensemble de ces pratiques. Il apparaît peu adapté puisqu'il évoque plus la discrimination positive en faveur des bas risques que la discrimination négative des hauts risques. Le terme de sélection différentielle (ou celui de sélection que nous utiliserons parfois dans la suite de cette présentation) nous semble mieux rendre compte simultanément de l'ensemble des aspects évoqués²⁰.

Il apparaît, au terme de cette première analyse que la sélection différentielle peut émerger dès lors qu'un producteur reçoit un paiement uniforme déterminé ex ante pour des bénéficiaires hétérogènes et qu'il est en mesure d'anticiper le coût de production (l'espérance individuelle de dommage dans le cas de l'assurance). S'il n'y a pas de responsabilité financière ou s'il est possible d'ajuster ex post le prix facturé au coût de production, le producteur n'a pas intérêt à adopter une telle stratégie.

b) Réglementation des primes

Certains auteurs, notamment Pauly (1984), considèrent que la sélection différentielle des risques n'apparaît en assurance que lorsque le marché est régulé et qu'elle ne peut en aucun cas émerger sur un marché concurrentiel.

Cette analyse fait référence à l'ensemble des cas dans lesquels l'Etat met en place sur les marchés d'assurance santé des mécanismes destinés à limiter la tarification au risque. L'objectif est alors d'encadrer la prise en compte du risque individuel par des assureurs afin d'en circonscrire les effets inéquitables.

On peut distinguer deux types de restrictions :

- les premières portent explicitement sur les critères de classification des risques. Le régulateur impose le type et le nombre de critères que les assureurs sont autorisés à

²⁰ L'adjectif "différentiel" signifie "relatif aux différences ou aux variations", d'après le dictionnaire alphabétique et analogique de la langue française (Robert, Tome 2, pp.220). On aurait pu utiliser sélection différenciée, mais le participe passé a un sens passif (en l'occurrence différencié signifie "marqué, séparé par quelque différence, sur lequel on opère une différenciation"). Si les assurés sont effectivement différenciés par les assureurs, la sélection elle-même, qui est une stratégie active, nous semble, en toute rigueur, plus différentielle que différenciée.

utiliser pour établir des catégories actuarielles, voire le nombre de modalités selon lesquelles un critère peut être décliné. Une restriction de ce type ne constitue pas en soi un encadrement des primes puisqu'aucune limite n'est imposée à la variation de celles-ci entre catégories. Ce type de restriction vise plus à standardiser les critères de mutualisation et donc, indirectement, la taille des pools à l'intérieur desquels le coût de l'assurance est partagé ;

- d'autres restrictions portent, pour leur part, sur les écarts de primes. L'objectif est alors, plus explicitement que dans le cas précédent, de tempérer les conséquences financières de la prise en compte par l'assureur des caractéristiques individuelles de risque. L'Etat peut ainsi limiter l'écart entre la prime maximale et minimale qu'un assureur fait payer pour un type de contrat donné. Ce mécanisme, très contraignant, entraîne de facto un encadrement des tarifs sur la population prise en charge par l'assureur. Dans une version plus souple, il est possible de limiter l'écart entre la prime minimale et maximale au sein d'une large catégorie actuarielle (âge-sexe). L'assureur est alors libre de choisir les critères sur lesquels il fonde les différences de prime à l'intérieur de chaque catégorie, mais il ne peut en répercuter tout l'effet sur l'assuré.

Le premier des cas que nous venons de décrire correspond à une situation dans laquelle l'assureur n'a le droit de faire varier ses tarifs qu'en fonction de critères précis, par tranche d'âge de 10 ans par exemple. Toute personne entre 60 et 70 ans paie la même prime, mais celle-ci peut être deux fois plus élevée que celle d'une autre tranche d'âge. Dans le second cas, l'écart entre la prime la plus basse et la plus haute est limité sur l'ensemble des assurés : la prime la plus haute représente un pourcentage fixe de la prime la plus basse. L'assureur détermine librement ses critères de classification. Le troisième cas correspondrait à une combinaison des deux critères précédents : les primes pour les personnes de 60 à 70 peuvent différer dans certaines limites (la prime la plus haute représente au maximum un pourcentage donné de la prime la plus basse, pour la tranche d'âge), mais les écarts de prime entre tranches d'âge restent à la discrétion de l'assureur.

Le terme de "*community rating*" désigne couramment ce type de réglementation. C'est à l'origine une technique tarifaire assurantielle qui consiste à faire payer une prime uniforme à toutes les personnes qui résident dans la même zone géographique. Son utilisation spontanée par les assureurs a progressivement disparu sous l'effet de la concurrence, mais le *community rating* peut être imposé sous différentes formes par le

régulateur. Ainsi, dans l'Etat de New-York, sur les marchés individuels et des petites entreprises, la prime fixée par l'assureur ne peut varier qu'avec la région, le degré de couverture et le statut familial de l'assuré. En Suisse, le prix des contrats de l'assurance maladie obligatoire ne peut varier qu'en fonction d'un critère géographique et du fait que l'assuré est un adulte ou un enfant. En France, la loi Evin²¹ stipule que le tarif payé par un assuré qui renouvelle son contrat ne peut augmenter au vu de son état de santé et que, "*si l'organisme veut majorer les tarifs d'un type de garantie ou de contrat, la hausse doit être uniforme pour l'ensemble des assurés souscrivant ce type de garantie ou de contrat*". Il s'agit aussi d'une forme de *community rating* stricte dans laquelle le tarif est uniforme pour tous les assurés choisissant le même contrat.

Dans l'ensemble de ces cas, les assureurs peuvent avoir connaissance d'une information sur le niveau de risque d'un assuré particulier, mais ils sont limités dans leurs possibilités de la répercuter dans la prime. Dans ce type d'univers concurrentiel réglementé, un assureur peut offrir des tarifs compétitifs s'il couvre une population qui présente un niveau de risque moyen plus faible que celui des autres assureurs. Si au contraire sa structure de risque est défavorable, les bas risques peuvent être tentés de contracter une assurance chez un concurrent mieux placé, voire de ne plus s'assurer. Quand bien même les assureurs en concurrence seraient collectivement obligés de couvrir une population donnée, chacun d'entre eux a intérêt à se décharger sur les autres de l'obligation d'assumer la couverture des hauts risques.

Ainsi, seul ce type de réglementation des primes peut expliquer, selon Pauly, l'apparition de sélection sur les marchés d'assurance. Cette vision, quand bien même justifiée, est notablement réductrice : même sans réglementation, la sélection est une stratégie possible pour l'assureur dans toutes les circonstances où il détient de l'information mais où il n'est pas en mesure de la répercuter intégralement dans les tarifs qu'il propose aux assurés. Concrètement, plusieurs cas peuvent se présenter :

- l'assureur peut s'interdire pour des raisons éthiques, de respect de la vie privée ou pour ne pas mettre à mal sa réputation de tarifier en utilisant certains facteurs de risque (quand bien même il en a connaissance - l'ethnie, le fait qu'une personne soit séropositive, l'histoire médicale familiale...-);

²¹ Loi 89-1009 du 31 décembre 1989 renforçant les garanties offertes aux personnes assurées contre certains risques, dite loi Evin.

- l'assureur n'a pas nécessairement les moyens de multiplier les catégories tarifaires :
 - d'une part, pour mettre en place une tarification actuariellement juste, il faut non seulement choisir les critères de tarification et estimer les primes mais aussi les réactualiser, c'est-à-dire déterminer périodiquement si les caractéristiques des individus ont changé et donc s'ils appartiennent toujours à la même catégorie tarifaire avant de réactualiser les primes. Compte tenu de la nature du risque en santé, les coûts de gestion d'un tel processus peuvent être très élevés ;
 - d'autre part, l'assureur est limité par la taille de la population qu'il couvre. Il n'a en général accès qu'aux informations disponibles dans ses propres fichiers pour évaluer les coûts par catégorie tarifaire. Plus celles-ci sont nombreuses, moins le nombre de personnes par catégorie est élevé et moins l'assureur a de chances d'estimer le risque sans commettre d'erreur. Comme la distribution des dépenses est très asymétrique, lorsque le niveau moyen des dépenses augmente, le nombre de personnes par catégorie tarifaire diminue et l'assureur court un risque encore plus élevé de mal tarifier le risque (Luft, 1995, pp. 25). Les hauts risques sont donc particulièrement vulnérables à la sélection car non seulement il est plus difficile d'établir un tarif juste pour eux mais en outre, une erreur est lourde de conséquence pour l'assureur ;
- enfin, l'assureur au fil du temps peut acquérir de l'information sur un individu (ne serait-ce qu'en observant ses consommations). Une fois qu'il détient cette information, il peut chercher à l'évincer plutôt que de réévaluer son risque en augmentant explicitement sa prime.

En pratique, les assureurs rencontrent de réelles difficultés pour remplir la condition nécessaire au fonctionnement efficace d'un marché concurrentiel de l'assurance qu'est la tarification au risque. La sélection résulte d'un arbitrage effectué par l'assureur qui peut émerger même si les primes ne sont pas réglementées : il peut en effet préférer la sélection différentielle des risques à la tarification différentielle si elle s'avère moins coûteuse. Newhouse confirme indirectement que l'arbitrage penche en faveur de la sélection lorsqu'il remarque que *"Plans typically distinguish between individuals and families.. but they often do not distinguish by age, sex, or even the number of*

children"²²(1996, pp.1258). L'effort des assureurs pour tarifer les risques semble donc dans les faits très limité, ce qui a pour conséquence inévitable le développement par les assureurs de stratégies de sélection différentielle des risques.

Cette première analyse indique que les circonstances concrètes dans lesquelles l'assureur peut avoir intérêt à mettre en œuvre une stratégie de sélection dépendent largement, mais pas exclusivement, de l'environnement réglementaire. La sélection est induite, dans tous les cas, par la non valorisation de l'ensemble des facteurs de risque dans le calcul des primes, imposée pour des raisons techniques, réglementaires ou choisie par l'assureur. Dans un tel contexte, il est essentiel de souligner qu'une asymétrie d'information en faveur de l'assureur n'est pas nécessaire à l'apparition de la sélection²³. Cette affirmation va à l'encontre de certaines analyses, sur lesquelles nous reviendrons plus longuement dans le point 2.1.3.

2.1.2 Les méthodes de sélection

L'existence de comportements de sélection différentielle sur le marché de l'assurance est très difficile, voire impossible à quantifier. A défaut d'enquêtes auprès de personnes qui auraient été victimes d'une forme ou une autre de sélection, tous les travaux scientifiques ayant pour objectif de documenter ce phénomène en sont réduits à faire état d'anecdotes illustrant la grande diversité des stratégies qui s'offrent en la matière à l'assureur.

a) Refus d'assurance

La stratégie qui consiste pour l'assureur à refuser de couvrir des personnes ou des groupes dont il soupçonne qu'ils présentent des niveaux de risque particulièrement élevés est l'exemple le plus flagrant d'un comportement de sélection.

La concentration de facteurs de risques dans certains secteurs d'activité fait que les entreprises qui en relèvent ne parviennent pas en général à couvrir les risques santé de leurs employés (Luft, 1995). Dowd et Feldman (1992) citent une enquête réalisée en 1990 auprès de 48 assureurs parmi lesquels 29 d'entre eux déclarent refuser systématiquement d'assurer certains types d'entreprises (les salons de coiffure, les compagnies de taxi). Au niveau individuel, pour offrir un contrat concurrentiel, l'assureur

²² *Les assureurs distinguent entre couverture individuelle ou familiale, mais ne distinguent pas l'âge, le sexe ou même le nombre d'enfants.*

²³ Exemple, certes caricatural, de l'origine ethnique.

doit être en mesure de déterminer le niveau de risque que présente la personne. Pour ce faire, il met en place un processus plus ou moins sophistiqué d'évaluation du risque et peut, dans ce cadre, collecter des informations générales sur l'assuré (âge, lieu de résidence...) et utiliser un questionnaire médical. Ce processus²⁴ doit théoriquement l'amener à déterminer le niveau de prime de l'individu. Dans la réalité toutefois, l'assureur peut aussi refuser de couvrir les personnes. Les études empiriques du phénomène de rejet sont très rares. Récemment Schmueli²⁵, sur la base de données d'enquête, étudie la probabilité de demander à souscrire une assurance et celle de se voir accepté. Il observe qu'en moyenne une personne âgée entre 45 et 65 ans ayant une maladie chronique a une probabilité de 50% de demander une assurance et de 30% d'être accepté. S'il déclare trois maladies chroniques, le même individu moyen, toutes choses égales par ailleurs, a une probabilité de 85% de demander une assurance et de 15% d'être accepté.

Les assureurs peuvent donc refuser de vendre de l'assurance s'ils estiment que les risques sont trop élevés (Pauly, 1988), mais ils peuvent aussi chercher à décourager certaines catégories de personnes à souscrire un contrat auprès d'eux. Luft (1995, pp.27) cite le cas de l'un des plans proposés par une université à ses employés qui avait "égaré" l'ensemble des demandes de souscription émanant d'hommes célibataires qui résidaient dans les quartiers de San Francisco privilégiés par la communauté homosexuelle. Une variante consiste à majorer la prime proposée à des personnes dont on sait qu'elles présentent des risques élevés largement au delà de ce qui est nécessaire pour couvrir le risque. Ce type de stratégie peut renforcer les difficultés d'accès précédemment évoquées pour des personnes qui ne sont pas très aisées. Elles peuvent être conduites soit à ne pas s'assurer soit à tenter de souscrire une couverture auprès d'un autre assureur. Dans le même ordre d'idée, on a vu en Suisse²⁶ (où la prime est uniforme pour tous les adultes couverts par le même assureur dans une zone géographique donnée) un assureur contacter personnellement des assurés âgés pour leur signaler que ses concurrents pouvaient leur proposer des contrats à des tarifs nettement plus avantageux.

²⁴ C'est l'*underwriting* : *screening applicants in order to determine risk class and acceptability* (Diamond et Rothschild, 1989).

²⁵ Shmueli A. (1998) *The effect of health on acute care supplemental insurance ownership: an empirical analysis*. Hebrew University School of Public Health, Working Paper.

²⁶ Anecdote rapportée dans la presse généraliste suisse.

b) Techniques de marketing

L'assureur peut mettre en œuvre des stratégies commerciales différenciées qui lui permettront d'attirer les personnes qui présentent de bons niveaux de risque. Comme le souligne Newhouse (1994, pp.137), s'il est possible de contrôler certaines pratiques flagrantes qui consistent à installer des stands de souscription dans des bals pour personnes âgées (sic), le choix d'acheter des encarts publicitaires dans tel ou tel journal dont le lectorat est susceptible de présenter un état de santé plus favorable, n'est pas non plus anodin. De même, envoyer des courriers publicitaires dans certains quartiers plutôt que d'autres (technique apparemment utilisée par certaines caisses d'assurance maladie aux Pays-Bas) ou de démarcher les entreprises plutôt que les individus (pour couvrir des personnes dont l'état de santé leur permet - à tout le moins - de travailler) permet de modifier le profil de recrutement moyen. En proposant des offres promotionnelles sur des services sans rapport direct avec l'assurance (un tarif préférentiel sur un abonnement à un club de gymnastique ou des réductions auprès d'agences de voyage), il est possible d'obtenir le même résultat (van de Ven et van Vliet, 1992). Plus généralement, le marketing, pour une entreprise quelconque, passe par la connaissance du marché et l'identification des segments de la population dont on cherche à attirer l'attention. Dans une seconde étape, on détermine l'approche la plus efficace pour atteindre ces cibles et susciter l'achat du produit que l'on propose. Les possibilités de sélection différentielle apparaissent aussi subtiles et variées que celles du marketing.

c) Calibrage des garanties

Les garanties proposées par les assureurs peuvent être plus ou moins attractives selon les niveaux de risques et les contrats moins couvrants sont plus susceptibles d'attirer des personnes présentant des niveaux de risque faible. Ceci permet à l'assureur de segmenter la population, cette fois en ne couvrant pas certains biens ou services. A titre d'illustration, un assureur qui ne couvre pas les produits injectables a peu de chances d'assurer des diabétiques. De même, les services d'aide psychologique ou psychiatriques et la chirurgie reconstructrice sont très discriminants (Newhouse, 1994). A l'inverse, en couvrant les services pédiatriques et les vaccins, l'assureur a plus de chances de répondre aux désirs de familles plus jeunes ; en couvrant généreusement les services préventifs, il attire plutôt des personnes en bonne santé ; en offrant une couverture pour les services consommés en dehors du lieu de résidence, il attire les personnes qui voyagent beaucoup

pour des raisons professionnelles et qui sont donc plutôt jeunes et actives (Chakraborty, et al., 1994).

d) Modulation de l'offre de soins

Lorsque l'assurance est couplée avec l'accès à un réseau de soins, les possibilités de sélection sont encore augmentées.

Dans le cas où l'assureur est acheteur de soins, il peut profiler son réseau pour être plus ou moins attractif aux hauts ou bas risques : en évitant d'offrir les services de spécialistes ou d'institutions réputées pour des pathologies lourdes (cancer, diabète, cardiologie), il crée une désincitation pour les personnes qui sont conscientes de présenter un haut risque (ou qui sont même déjà affectées par ces pathologies) de s'assurer auprès de lui (van de Ven et van Vliet, 1992). En sélectionnant le lieu d'installation des prestataires de son réseau , il peut aussi se rendre peu attractif à des populations plus risquées. Kronick et al. (1995) soulignent ainsi que les plans évitent généralement de développer des réseaux à proximité des quartiers défavorisés ou dans le centre des villes (*inner cities*), car les personnes plus aisées résident traditionnellement dans les banlieues aux Etats-Unis.

Enfin, l'assureur peut reporter la responsabilité de la sélection sur les prestataires de soins, en les amenant à traiter moins courtoisement les patients qui présentent des niveaux de risque élevés ou en leur laissant entendre qu'ils seraient mieux traités par des collègues (Newhouse, 1994, van de Ven et van Vliet, 1992).

Conclusion partielle

Au terme de cette présentation, on mesure à quel point les stratégies de sélection sont variées. Parmi les méthodes présentées, on peut cependant établir une gradation :

- à un bout du spectre, on trouve le refus d'assurance, qui constitue la méthode la plus brutale de sélection ;
- à l'autre bout, on trouve des stratégies beaucoup plus subtiles qui ont pour résultat de biaiser le recrutement de l'assureur et qui aboutissent à ce que le risque de sa population couverte ne soit pas représentatif du risque moyen de la population générale. Il pourra alors offrir des prix moyens plus compétitif qu'un concurrent ou réaliser un profit supérieur s'il propose le même prix moyen. Ces stratégies de sélection différentielle peuvent notamment reposer sur la tendance à la segmentation

spontanée, autrement dit, sur la propension des personnes qui ont des caractéristiques différentes à faire des choix d'assurance différents, évoquée dans le chapitre précédent.

S'il est possible d'interdire explicitement certaines formes de sélection, comme le refus d'assurance, il est très difficile de trouver des parades réglementaires à d'autres. En effet, adapter son offre au goût des consommateurs (ou de certains consommateurs) n'est pas en soi répréhensible et, quand bien même la sélection serait prohibée, apporter la preuve d'une intention délictueuse peut s'avérer difficile.

2.1.3 Sélection différentielle et antisélection : une clarification conceptuelle

Si les assureurs peuvent s'appuyer sur la segmentation spontanée pour procéder à la sélection, nous avons précédemment montré qu'ils peuvent aussi utiliser le même type de mécanisme pour surmonter l'antisélection et inciter les personnes qui détiennent des informations privées sur leur niveau de risque à les révéler indirectement.

Cette remarque souligne que les trois concepts de sélection, antisélection et segmentation spontanée entretiennent des liens complexes. On relève d'ailleurs dans la littérature économique un certain nombre d'ambiguïtés et de contradictions dans la présentation de ces différents concepts (a). Nous proposons donc une clarification (b) avant de conclure par quelques remarques de nature empirique (c).

a) Quelques ambiguïtés

En premier lieu, sélection et antisélection sont parfois présentées comme des réactions des acteurs du marché, et parfois définies en référence aux caractéristiques de l'environnement :

- Ainsi van de Ven décrit deux formes de sélection :

- *"adverse selection : the selection that occurs because high-risk consumers have an incentive to buy more coverage than low risk consumers within the same premium risk group ;*

- *cream skimming* : the selection that occurs because health plans prefer low low-risk consumers within the same premium risk group²⁷
- Pour Cutler : "adverse selection occurs both because healthy people move into less generous plans and because sick people remain in more generous plans as they age²⁸ ;
- Frank, Glazer et Mc Guire définissent l'antisélection ainsi "adverse selection : incentive (for health insurers) to distort the bundle and the quality of services to attract profitable enrollees and to deter unprofitable enrollees"²⁹ ;
- Pour Newhouse (1996, pp. 1236), la sélection est "actions of economic agents on either side of the market to exploit unpriced risk heterogeneity"³⁰.

Ces quelques exemples, que l'on pourrait multiplier, témoignent indiscutablement d'une certaine confusion, voire soulignent des contradictions entre les approches des uns et des autres. Ce sentiment est renforcé par la description des liens qu'entretiennent sélection et antisélection.

En effet, les circonstances dans lesquelles sélection et antisélection émergent sont comparables : un prix moyen est proposé pour des risques hétérogènes et cette situation, contraire au principe de tarification actuarielle, génère des réactions stratégiques de la part des acteurs du marché. Selon la définition que nous avons donnée en introduction du chapitre 4, le concept d'antisélection insiste, dans un premier temps, sur les réactions stratégiques de l'assuré qui sort du marché si on lui propose un prix supérieur à celui qu'il devrait payer, compte tenu de son niveau de risque.

²⁷ Antisélection : sélection qui émerge parce que les consommateurs "hauts risques" ont intérêt à acheter plus de couverture que les consommateurs "bas risques", à l'intérieur d'une même catégorie de prime. Ecrémage : la sélection qui émerge parce que les assureurs préfèrent couvrir, à l'intérieur d'une même catégorie de prime, les consommateurs "bas risques" que les consommateurs "hauts risques". iHEA 2d World Health Conference, Juin 1999, Rotterdam, document distribué dans la Pre-Conference Session "Risk adjustment in competitive health plan markets"

²⁸ L'antisélection émerge parce que d'une part les personnes en bonne santé choisissent moins les plans généreux et parce que les malades restent dans ces plans généreux quand ils vieillissent. iHEA 2d World Health Conference, Juin 1999, Rotterdam Abstract Book pp.44. "Combating adverse selection in multiple choice health plans".

²⁹ L'antisélection est le fait de modifier le panier et la qualité des services couverts afin de décourager les hauts risques et d'attirer des assurés profitables. iHEA 2d World Health Conference, Juin 1999, Rotterdam Abstract Book pp.67. "Measuring adverse selection in managed health care".

³⁰ Les actions d'agents économiques des deux côtés du marché pour exploiter des hétérogénéités de risque non tarifées.

Pour faire face à ce phénomène, l'assureur réagit :

- selon Rothschild et Stiglitz, en proposant des contrats différents : il ne connaît certes pas le niveau de risque des individus, mais comme il connaît leur fonction de réaction, il les incite à révéler leurs caractéristiques ;
- il peut aussi choisir de contrer le phénomène de fuite des bas risques par une stratégie de sélection active. Pour Newhouse (1996, pp.1244), ceci est notamment vrai lorsque les coûts d'introduction de contrats séparateurs sont trop élevés. L'assureur propose alors un contrat *pooling* et sélectionne les risques.

Dans l'optique de Newhouse, la sélection est donc une conséquence de l'asymétrie d'information en faveur des assurés.

Par ailleurs, nous avons noté que l'antisélection caractérise un marché sur lequel l'assureur est moins bien informé que l'assuré. Or, on considère parfois que la sélection est une stratégie fondée sur le fait que "*l'assureur est plus informé que l'assuré (ne serait-ce que parce qu'il dispose d'études épidémiologiques et statistiques) sur les risques encourus*" (Henriet et Rochet, 1999, pp.125).

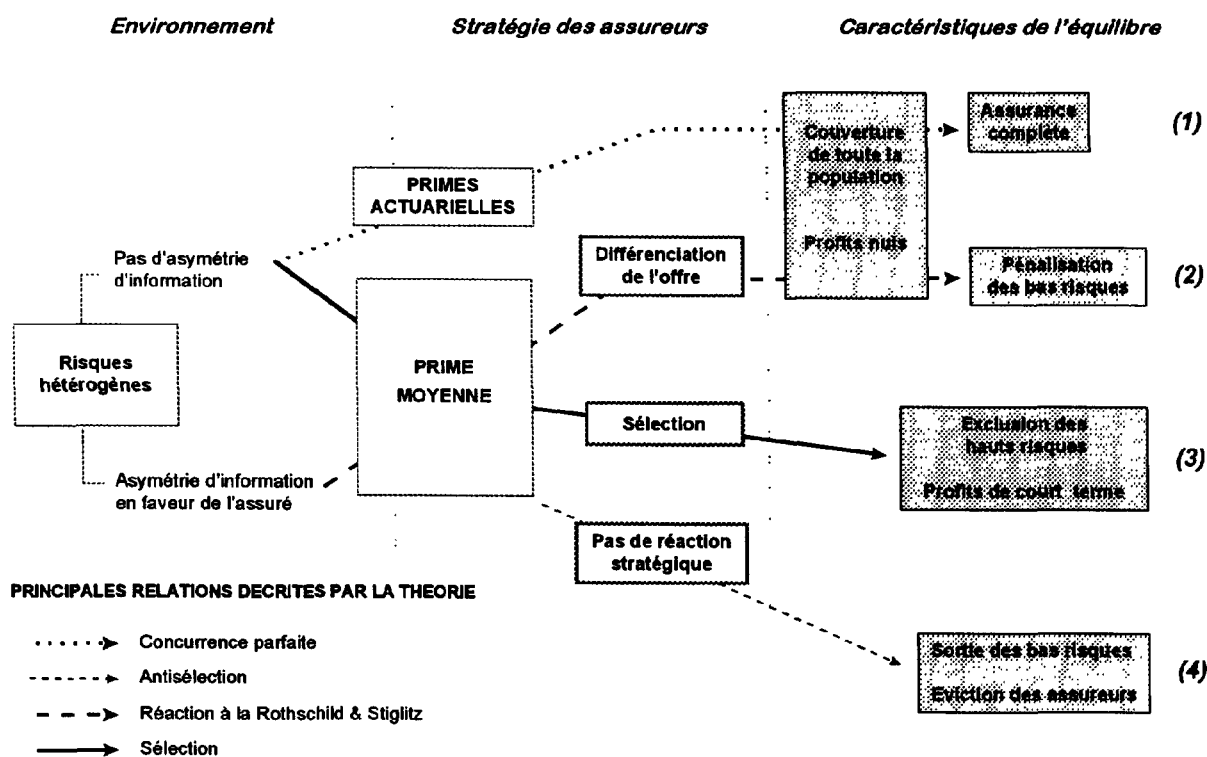
Dans cette optique, on pourrait penser que antisélection et sélection sont mutuellement exclusives puisqu'elles reposent sur des asymétries d'information de sens inverse.

Ces deux approches, à propos des liens qu'entretiennent sélection et antisélection, semblent donc, là encore, contradictoires.

b) Sélection et antisélection : une tentative de clarification

Nous allons donc proposer une analyse des différents phénomènes dans un cadre unique, sur la base du schéma de la page suivante. Il articule la description des éléments pertinents pour caractériser l'environnement dans lequel assureurs et assurés évoluent, les réactions stratégiques des assureurs et les principaux résultats en terme de dynamique et d'équilibre du marché.

ANTISÉLECTION ET SÉLECTION



Sur le marché de l'assurance maladie, les risques sont hétérogènes. Lorsqu'on leur propose un menu de contrats, les personnes qui présentent des niveaux de risque différents effectuent leur choix en fonction, d'une part, du service proposé et d'autre part, du prix du contrat, ou plus précisément du fait que la prime reflète ou non leur propre risque. La fonction de réaction des individus auxquels on propose des contrats différents, est connue des assureurs. Elle repose sur le mécanisme d'auto-sélection, appelé ici de segmentation spontanée pour refléter à la fois sa conséquence principale - la segmentation - et son caractère librement choisi par les individus.

Ceci étant donné, on peut distinguer deux types de situations :

1. Si l'information est symétrique, et si on postule que la tarification au risque est non seulement techniquement possible et autorisée, mais aussi, peu coûteuse à mettre en œuvre, l'assureur propose à chacun un contrat actuariellement juste et l'ensemble de la population est couvert³¹ (cas 1) ;

³¹ On ignore dans cette analyse les hétérogénéités de revenu.

2. Si l'une ou l'autre de ces conditions n'est pas remplie, l'assureur propose une prime moyenne, imparfaitement ajustée au risque des assurés. Plus précisément, trois raisons expliquent que l'assureur propose une prime moyenne :

- il ne connaît pas le risque de l'assuré, (Rothschild et Stigitz, 1976) ;
- la tarification au risque n'est pas autorisée (Pauly, 1984) ;
- à ces deux contraintes exogènes, nous avons proposé d'ajouter le cas où l'assureur choisit volontairement de ne pas tarifer certains facteurs de risque³².

Lorsque le prix de l'assurance est uniforme pour des risques qui ne le sont pas, certaines personnes, qui présentent un niveau de risque moins élevé que celui sur lequel se fonde le prix moyen, peuvent choisir de ne pas s'assurer. Si l'assureur ne modifie pas sa stratégie, il va couvrir des personnes qui présentent un niveau de risque trop élevé par rapport au prix proposé et il sera peu compétitif. Une telle situation est instable et peut à terme entraîner la disparition du marché (cas 4). Ce phénomène est qualifié d'antisélection dans le cas où l'assureur est moins bien informé que l'assuré (Chassagnon 1996, cf. introduction du chapitre 4). Il faut cependant souligner que les assureurs peuvent en être victimes dans tous les cas où ils n'ajustent pas les primes³³. Ce phénomène peut aussi toucher un assureur particulier qui différencierait moins les primes qu'un autre : les mutuelles en France, qui ajustent moins précisément les primes aux risques que la plupart des assureurs privés, ont ainsi plus de mal à attirer les personnes jeunes.

Afin d'éviter la fuite des bas risques, deux types de parades s'offrent à l'assureur :

- il propose une gamme de contrats, chacun d'entre eux étant susceptible d'attirer des groupes d'individus qui présentent des niveaux de risque assez proches. Il peut alors améliorer la qualité de sa tarification au risque et couvrir l'ensemble de la population. Cette stratégie de différenciation de l'offre est décrite dans les modèles "à la Rothschild et Stiglitz". Dans le cas particulier qu'ils développent, les bas risques acceptent, pour donner un signal sur leur niveau de risque, des contrats qui sont moins généreux que ceux qu'ils choisiraient en information symétrique (avec franchise

³² Le fait que la prime est actuarielle ou non dépend donc à la fois de l'environnement et des choix stratégiques de l'assureur, ce que montre le graphique.

³³ Ainsi, la flèche pointillée associée à la légende "antisélection" ne renvoie pas de façon univoque à la situation d'asymétrie d'information, mais seulement au fait que l'assureur propose une prime moyenne.

par exemple)³⁴. Quoi qu'il en soit, cette stratégie rapproche le marché d'une situation optimale dans laquelle les assurés souscrivent des contrats qui répondent à leurs besoins, paient un prix actuariel et les profits des assureurs sont minimisés (cas 2);

- l'assureur en concurrence avec d'autres peut aussi choisir de mettre en œuvre une sélection différentielle des risques (cas 3). La stratégie de l'assureur peut être brutale (refuser d'assurer une personne qui présente une caractéristique dont il sait qu'elle est liée à un niveau de risque élevé) ou elle peut s'appuyer sur la segmentation spontanée : l'assureur profilant alors son offre pour attirer de préférence les bas risques. Si un assureur met en œuvre des stratégies de sélection, il peut réaliser des profits dans la mesure où tous les autres ne la mettent pas en œuvre avec la même efficacité. La théorie des jeux pourrait permettre de caractériser les équilibres qui émergent mais l'aspect sur lequel il nous semble intéressant d'insister est le fait que la tentation de réaliser des profits de court terme peut conduire à limiter l'accès à l'assurance de certaines catégories d'individus.

Les analyses théoriques que nous avons présentées dans les chapitres 4 et 6 soulignent respectivement :

- que la stratégie de différenciation de l'offre constitue une réponse à une asymétrie d'information (flèche pointillée) ;
- que la stratégie de sélection émerge lorsque l'information disponible n'est pas reflétée fidèlement dans la prime (flèche pleine).

Nous proposons un cadre plus général qui intègre ces différents raisonnements et qui souligne que des causes différentes peuvent entraîner les mêmes effets : par exemple, le graphique permet de rendre compte de l'analyse de Newhouse selon laquelle la sélection peut être la conséquence d'une asymétrie d'information au bénéfice de l'assuré sur son propre niveau de risque. Quelle que soit la raison initiale pour laquelle l'assureur propose une prime moyenne, sélection et différenciation des contrats peuvent constituer une réponse possible de l'assureur.

³⁴ Nous nous sommes interrogés, dans le chapitre 5, sur la nécessité pour les bas risques de consentir des sacrifices sur leur niveau de couverture pour obtenir un prix actuariel.

Il reste à souligner que l'assureur peut combiner ces différentes stratégies. Il peut ainsi utiliser une stratégie de différenciation de l'offre qui amène les assurés à révéler une partie de l'information qu'ils détiennent mais qui lui échappe et, par la suite, pour chaque "segment", recourir à la sélection afin d'exploiter des informations qu'il détient mais qu'il ne peut associer à des variations de primes.

Imaginons une situation dans laquelle un assureur est en concurrence avec d'autres pour couvrir la population d'une entreprise. Afin d'augmenter la transparence de l'offre et de promouvoir une certaine forme d'équité, l'employeur peut imposer que les primes des contrats ne soient différenciés qu'en fonction d'un petit nombre de critères, comme le lieu de résidence et la structure de la famille³⁵. En n'offrant que des contrats qui couvrent généreusement les services préventifs et pas les soins palliatifs, l'assureur attirera une clientèle en meilleure santé ; en offrant un contrat avec franchise et un contrat sans franchise, il pourra segmenter cette population en fonction de caractéristiques non observables (pour caricaturer, le fait d'être hypocondriaque).

Au terme de cette analyse, il nous semble que sélection et antisélection ne sont pas mutuellement exclusives. L'information sur le risque est multidimensionnelle et les agents ne sont pas parfaitement ou symétriquement informés. Les stratégies d'un assureur peuvent être multiples, sans être nécessairement incompatibles.

c) Difficile identification empirique

Sur le plan empirique, l'analyse qui précède explique largement pourquoi le partage entre segmentation spontanée, antisélection et sélection différentielle est très délicat à établir.

Si on étudie les choix d'assurance d'une population à laquelle on propose plusieurs contrats et si on observe une corrélation entre le niveau de risque et de couverture, le fait stylisé peut résulter, comme nous l'avons vu, soit d'une segmentation spontanée en information parfaite, soit de la volonté des bas risques de consentir des sacrifices pour révéler une information (antisélection). Il peut aussi être le résultat de stratégies de sélection : si certains assureurs écrèment délibérément les bons risques en ne proposant que des contrats moins généreux sur certains aspects, les autres, qui proposeraient des contrats plus généreux peuvent être conduits à couvrir des personnes qui présentent des

³⁵ Cette situation est classique même si l'exemple précis est imaginé.

niveaux de risque plus élevés. Faire la part de la sélection et de la segmentation spontanée peut donc aussi s'avérer délicat.

Ainsi, dans la délégation de gestion par Medicare aux HMO, ces derniers reçoivent un paiement forfaitaire déterminé ex ante. Ils ont intérêt à attirer les personnes qui présentent des niveaux de risque moins élevés que ce que leur permet de couvrir la capitation et peuvent mettre en place une stratégie de sélection. Un assureur américain rencontré lors d'un voyage d'étude aux Etats-Unis³⁶ a implicitement reconnu avoir recours à ce genre de pratique. Sa stratégie commerciale était la suivante : il envoyait un document proposant de "fournir des renseignements sur les possibilités ouvertes par l'adhésion au HMO" à une liste de personnes de plus de 65 ans, dont il avait obtenu les coordonnées sur une fichier de permis de conduire (premier biais), accompagné d'une fiche de renseignements à remplir si on était intéressé (second biais). En fonction des réponses aux questions, le HMO relançait ou non le client potentiel. Il est strictement interdit de refuser de couvrir un bénéficiaire de Medicare qui en fait explicitement la demande. La dernière étape de la stratégie de l'assureur était, à cet égard, à la limite de la légalité, mais il jouait sur l'ambiguïté de l'objectif qui était simplement de "fournir des informations sur le HMO" et non pas de vendre (immédiatement) un contrat. A contrario, les HMO rendent un service qui est, de fait, plus susceptible d'attirer les personnes qui ne sont pas malades puisqu'ils insistent sur la prévention et qu'ils organisent la prise en charge dans leurs propres réseaux. Les personnes déjà malades et attachées à un prestataire particulier ont plus de difficultés à accepter de telles contraintes. Il y a dans une certaine mesure ici convergence entre les intérêt des HMO et des assurés qui présentent un niveau de risque faible. En fait, la quasi-totalité des exemples de segmentation spontanée que nous avons présentés précédemment peut résulter d'une combinaison de segmentation spontanée et de sélection.

³⁶ Entretien Avril 1996.

2.2 La sélection différentielle est-elle préjudiciable ?

Après avoir défini la sélection, il reste à souligner qu'elle représente un enjeu non seulement parce qu'elle génère de l'exclusion (2.2.1) mais aussi parce qu'elle n'est pas efficace (2.2.2).

2.2.1 Accès à l'assurance et aux soins

Les personnes qui présentent des niveaux de risque élevé sont celles qui sont le plus susceptibles de souffrir de la sélection différentielle. En effet, si tous les assureurs déploient des efforts pour avoir un profil de risque moyen meilleur que les autres, les hauts risques auront plus de difficultés à trouver un contrat qui leur convienne (Gauthier, et al., 1995, par exemple). Mais il ne s'agit pas là de leur seule vulnérabilité. En effet, nous avons signalé que la sélection des risques pouvait passer, dans le cas où l'assureur propose de fournir les soins, par une modulation de l'offre de services qui vise plutôt à décourager les personnes présentant un risque élevé, voire malades, à se soigner dans les réseaux concernés. Si une telle stratégie se généralise, la qualité des soins dont elles bénéficient peut, à terme, se dégrader (van de Ven et van Vliet, 1992).

Dans une perspective intertemporelle, la question de la sélection peut être mise en rapport avec une préoccupation que nous avons soulevée dans le troisième chapitre, à savoir l'incapacité des marchés à fournir de l'assurance de long terme. Nous avons montré qu'une parade possible pouvait être trouvée dans la constitution de groupes de risques non homogènes qui s'assurent collectivement et solidairement, par exemple au sein de l'entreprise. La tendance de nombreux employeurs aux Etats-Unis à proposer plusieurs contrats aux assurés, laisse la place à la segmentation spontanée, la sélection et l'antisélection. Ces trois phénomènes poussent les assureurs à une classification des risques toujours plus grande qui rend les solutions de type *pooling* plus précaires. Sur la base de cet argument, Newhouse affirme que "*selection can cause markets to be incomplete*"³⁷ (1996, pp.1258).

³⁷ *La sélection peut rendre les marchés incomplets.*

2.2.2 Sélection et efficacité

En admettant qu'il soit possible de protéger les personnes vulnérables et de leur garantir un accès à l'assurance, il n'en reste pas moins que la sélection est inefficace à plusieurs titres.

Sur un marché traditionnel, plus la concurrence est forte, plus elle incite les producteurs à augmenter leur l'efficacité productive. Si l'information est imparfaite, cet avantage est contrebalancé par un risque croissant de sélection. En effet, pour que le prestataire d'un service fournisse l'effort maximal de diminution de coûts, et donc pour obtenir l'efficacité productive, il faut qu'il reçoive un paiement prospectif. Le corollaire du paiement prospectif est l'appropriation par le producteur du résultat financier de son activité :

- s'il parvient à fournir le service en n'utilisant pas toutes les ressources qui lui sont allouées en début de période, il réalise un profit qui lui appartient ;
- à l'inverse, comme il est engagé contractuellement à fournir le service en contrepartie du paiement, il est responsable d'un éventuel déficit.

La rémunération prospective et forfaitaire incite donc le producteur à minimiser les coûts de production et à optimiser son effort de gestion. Or, dans le même temps, nous avons vu qu'un tel mode de rémunération, si les paiements prospectifs sont imparfaitement ajustés au risque, peuvent inciter à la sélection. A l'inverse, si les coûts de production constatés sont intégralement remboursés au producteur, il ne tire aucun bénéfice de la sélection mais il n'est pas incité à minimiser les coûts. L'existence même d'un paiement prospectif censé générer l'efficacité incite aussi le prestataire à exploiter toute hétérogénéité des risques non tarifée, en mettant en place une sélection différentielle de ces risques. Du point de vue social, il existe donc un arbitrage entre efficacité productive et sélection (Newhouse, 1996). Un tel arbitrage est valable pour un hôpital qui est payé au coût par pathologie mais aussi (entre autres) pour l'assureur qui reçoit une prime en début de période. On peut noter à cet égard que, contrairement à une idée assez largement répandue, un assureur qui pratique l'*experience rating*³⁸ qui lui permet dans les faits de reporter sur l'année suivante (dans la mesure où le contrat est renouvelé) un déficit par rapport aux primes collectées une année, n'est pas, dans un marché peu

³⁸ Cette technique consiste à ajuster systématiquement en $t+1$ la prime d'assurance en fonction des dépenses observées (l'expérience) en t .

concurrentiel, incité à diminuer les coûts puisqu'il répercute les augmentations sur les assurés. Ce système prévalait aux Etats-Unis dans les années soixante-dix et générait une situation paradoxale où les assureurs n'étaient pas mis au risque et répercutaient les augmentations de coûts sur les employeurs (Newhouse, 1996, pp.1237). Le paiement doit être véritablement prospectif et doit rester identique, quelle que s'avère être la consommation de ressources par le bénéficiaire (Rossiter, 1987).

Dans le domaine de l'assurance santé, la concurrence, dans la mesure où elle semble générer la sélection, est certes dommageable en termes d'équité, mais surtout, elle ne permet pas nécessairement d'améliorer l'efficacité globale du système de santé pour trois raisons.

1. Tout d'abord, la sélection biaise la concurrence qui ne porte pas sur les critères socialement bénéfiques. Les assureurs qui utilisent la sélection offrent des prix plus faibles et sont donc plus concurrentiels, sans pour autant augmenter leur efficacité productive par une minimisation des coûts : *"a plan can quote low price, offer high benefits, and capture a good share of employees without organizing a tough managed care system"*³⁹ (Jones, 1989, pp. 53). Ainsi, dans un contexte où la sélection existe, les différences de primes risquent de refléter plus la capacité des assureurs à constituer des groupes de risques favorables que leur efficacité productive (Cutler, 1994b, pp.32). On peut aboutir à une situation paradoxale où un assureur dont l'efficacité productive est médiocre mais qui sélectionne les bons risques évince du marché un autre, efficient mais dont la population est plus risquée (van de Ven et van Vliet, 1992). Ce mécanisme a peu de chances d'être sanctionné par les assurés qui se sont pas capables de déterminer, ni réellement intéressés par le fait de savoir, dans quelle mesure une prime basse est bien la conséquence d'une production moins coûteuse du service et non d'une stratégie de sélection.
2. Ensuite (Jones, 1989, pp. 54), le fonctionnement du marché peut décourager des assureurs de développer des innovations qui permettraient de diminuer les coûts. En effet, pour une structure de risque initiale donnée, l'assureur qui développe par exemple les soins à domicile réalise une économie. Mais si cette innovation attire plus

³⁹ Une assurance peut offrir des primes faibles, des garanties élevées et attirer une bonne partie des employés sans organiser une stricte gestion des soins.

particulièrement les personnes qui ont besoin de ce type de service, ses primes augmenteront et il apparaîtra de fait moins compétitif.

3. Enfin, on peut noter que la stratégie de sélection a un coût pour l'assureur : allouer des ressources à cette activité constitue un gaspillage net du point de vue social. Le modèle de sélection de Jack (1997) confirme cette intuition ; il postule un univers d'assurance maladie obligatoire, où les assureurs choisissent, pour offrir des prix plus compétitifs et maximiser leur profit, entre une diminution des coûts (socialement bénéfique) et une activité de sélection (qui ne l'est pas). Les assureurs sont contraints de faire payer la même prime à toute personne qu'ils couvrent. L'équilibre de Nash de cette économie est tel que, si la proportion de hauts risques n'est pas très élevée dans la population, tous les assureurs réalisent un investissement identique non nul dans la sélection. On est dans un cas typique de dilemme du prisonnier : il serait optimal qu'aucun ne fasse de sélection, mais en fait ils dépensent tous des ressources (identiques) pour obtenir un profil de risque qui est le même que s'ils n'investissaient pas dans la sélection, puisque l'ensemble de la population doit être assuré.

En conclusion de cette section, on peut souligner qu'au cours des dernières années, la gestion de la sélection différentielle a progressivement été mise au centre des débats sur les systèmes de santé.

Plus précisément, un certain nombre de pays dans lesquels la concurrence entre assureurs est une pierre angulaire du système, ont pris conscience de la nécessité de juguler la sélection pour obtenir des organismes assurantiels qu'ils se concentrent sur la recherche de l'efficacité productive. Aux Etats-Unis ainsi, des entités qui achètent l'assurance pour le compte des ménages, comme les grandes entreprises ou les Etats qui couvrent leurs employés, organisent en ce sens la concurrence entre les assureurs dont ils proposent les produits.

A l'autre bout du spectre, des pays dans lesquels la couverture des risques liés à la santé est largement socialisée ont introduit une forme de concurrence entre assureurs ou caisses d'assurance maladie afin de promouvoir l'efficacité. Dans cette démarche, les régulateurs ont tenté de mettre en place une responsabilisation financière, tout en cherchant à éviter que les organismes gérant le risque recourent à la sélection.

CONCLUSION

Au terme de ce chapitre, il nous semble intéressant de souligner que les deux hypothèses évoquées pour expliquer le phénomène de non assurance se combinent probablement dans un système où l'assurance est volontaire. En effet, pour un niveau de revenu donné, lorsque le risque d'une personne augmente :

- elle doit s'acquitter d'une prime plus élevée et peut se trouver dans une situation dans laquelle il lui est difficile de souscrire un contrat ;
- dans le même temps, elle est plus vulnérable à la sélection des risques.

A l'inverse, lorsque son risque est faible, elle peut trouver plus facilement à souscrire une assurance, non seulement parce que sa prime est faible, mais aussi parce que l'assureur peut avoir intérêt à la couvrir.

Un tel système est, certes inéquitable, mais aussi générateur d'inefficacité à plusieurs titres : les personnes qui ne sont pas couvertes ont accès à des soins de façon plus tardive, ce qui peut augmenter leur niveau de risque et donc la tentation d'avoir recours aux "filets de sécurité" de type hôpital public. En outre, quand bien même les personnes qui n'ont pas les moyens de souscrire une assurance, bénéficieraient de subventions pour acheter de l'assurance, elles resteraient vulnérables à la sélection et la concurrence ne se déroulerait pas nécessairement sur les critères souhaitables.

CONCLUSION DE LA PARTIE II

La maîtrise de la sélection au sens large, sélection par le revenu et sélection différentielle par le risque, est une motivation commune aux grandes réformes de l'assurance maladie entreprises ces dernières années. L'objectif est d'éviter que la concurrence ne se traduise par l'exclusion de l'assurance de certains individus mais aussi d'obtenir que les entités, dont la responsabilité financière est en jeu et qui couvrent le risque maladie des individus, gèrent au mieux ce risque et maîtrisent les coûts sans avoir la possibilité ou la tentation de recourir à la sélection différentielle.

Pour répondre à cet objectif, les grandes orientations des réformes convergent. Elles passent par une limitation de l'efficacité allocative, qui augmente ou maintient l'équité des systèmes et cherchent à en améliorer l'efficacité productive.

Ces évolutions se traduisent de façon différente selon que le système de santé est à l'origine plutôt libéral (Suisse ou Etats-Unis) ou socialisé (Belgique ou Pays-Bas).

Dans les systèmes traditionnellement socialisés, on observe que l'introduction de mécanismes concurrentiels dans les systèmes d'assurance maladie a pour principal objectif d'en augmenter l'efficacité productive. Ces réformes ne conduisent jamais à diminuer le nombre de personnes qui bénéficient d'une couverture sociale. Elles n'ont pas d'impact sur l'efficacité allocative, du point de vue des assurés sociaux, dans le sens où elles ne leur confèrent ni le choix de souscrire ou non un contrat, ni le droit de choisir le niveau de couverture. En revanche, ils peuvent choisir l'opérateur auquel ils confient la gestion de cette assurance.

Pour les pays de tradition libérale, dans la mesure où l'on facilite l'accès à l'assurance santé pour les hauts risques et les personnes peu aisées, on augmente de fait l'équité du système car cette généralisation de la couverture s'accompagne d'une redistribution des ressources. Comme nous l'avons montré dans le chapitre six, l'universalisation de la couverture est la conséquence d'un arbitrage politique qui implique de dépasser l'obstacle constitué par le fait que dans tout passage de ce type, il y a des gagnants et des perdants, et que les perdants sont plutôt les personnes aisées et en bonne santé.

Un tel choix n'est cependant pas seulement la conséquence de l'émergence d'un sentiment de solidarité : on peut aussi considérer qu'il est motivé par la recherche de l'efficacité.

a) Tout d'abord, si l'on cherche à améliorer l'efficacité des prises en charge dans le système et si l'on considère que les assureurs peuvent se doter des moyens de le faire, en laissant se développer la sélection par le risque et par le revenu, on limite le champ sur lequel les assureurs opèrent cette gestion des risques.

b) Indépendamment de cela, il apparaîtra, dans la suite de ce travail, que, pour obtenir des opérateurs gestionnaires du risque en concurrence qu'ils concentrent leurs efforts sur maîtrise des coûts, il s'avère préférable :

- de standardiser les contrats ;
- d'encourager, voire d'imposer, une généralisation de la couverture ;
- de définir explicitement et de limiter les critères de tarification de l'assurance.

Pour obtenir l'efficacité productive, on limite donc l'efficacité allocative et on réintroduit de fait, a minima, une forme d'équité entre hauts et bas risques.

c) Ensuite, l'organisation de la concurrence selon les modalités que nous venons de décrire est susceptible de résoudre les difficultés créées par l'antisélection. En effet, elle passe par une harmonisation de la couverture de base à laquelle on associe un prix moyen. Si toutes les personnes sont obligées de souscrire ce contrat, elles peuvent ensuite compléter la couverture par des contrats "complémentaires" (configuration proche de celle proposée par Pannequin 1992 décrite dans le chapitre 4).

d) Enfin, le caractère obligatoire ou collectif de l'assurance permet d'apporter une solution à un certain nombre de difficultés évoquées dans la première partie de ce travail, en imposant une coopération mutuellement avantageuse entre les individus. L'assurance collective ou obligatoire, qui généralise la péréquation entre niveaux de risques et en garantit la pérennité, constitue plus précisément un moyen de résoudre le problème de l'assurance de long terme qui n'est pas disponible sur les marchés. Atkinson souligne en outre que l'assurance universelle permet de couvrir ex post des risques difficilement assurables par le marché (1995) : "*providing for contingencies*

*which are not foreseen, or to allay fears about events which we cannot forecast...*⁴⁰
(pp.210). Ces éléments contribuent à alimenter le sentiment qu'il peut y avoir optimalité sociale à se couvrir collectivement et de manière coercitive : "*cette forme d'entraide, au niveau de la société comme de la famille, conduirait ainsi à un partage plus efficace (avantageux pour tous)... des risques difficilement assurables ou diversifiables par le marché parce que collectifs ou corrélés, imprévisibles ou encore évolutifs,*" (Masson, 1995).

Il est donc difficile de distinguer, dans ces réformes, ce qui relève exclusivement de la recherche d'une plus grande efficacité globale de ce qui procède d'un arbitrage explicite en faveur de l'équité, mais il semble que ces deux objectifs ne soient pas inconciliables.

L'analyse de ces réformes sera détaillée dans la partie suivante.

⁴⁰ *Fournir une assurance pour des risques inconcevables ou alléger la crainte qu'un événement qu'on ne peut prévoir se réalise.*

PARTIE III

CONCURRENCE ENTRE ASSUREURS DANS LES SYSTEMES DE SANTE

PARTIE III

CONCURRENCE ENTRE ASSUREURS DANS LES SYSTEMES DE SANTE

Les grandes réformes des systèmes de santé sont souvent inspirées par les expériences des autres pays. Malheureusement, comme le note ironiquement Maynard (1995, pp.49), l'élaboration des politiques se fonde plus souvent sur des effets de mode et des croyances que l'étude rigoureuse du fonctionnement des systèmes.

Une des missions du chercheur consiste à analyser de façon aussi rigoureuse que possible ce fonctionnement sans se contenter de comparer les agrégats macro-économiques qui reflètent des réalités profondément différentes. Il doit analyser, au delà des principes affichés, les implications concrètes des réformes et déterminer les paramètres institutionnels et techniques qui en garantissent ou non le succès.

L'objectif de cette troisième partie est d'approfondir l'analyse qui précède en tentant d'appliquer ces principes. L'idée couramment avancée selon laquelle un certain nombre de réformes entreprises ces dernières années, notamment en Europe, visent à introduire la concurrence entre assureurs dans le domaine de la santé, nous semble un peu simplificatrice.

Pour le démontrer nous commencerons par dresser, dans le chapitre sept, un bilan du rôle que remplit l'assurance privée dans les systèmes de santé des pays développés. Nous montrerons notamment qu'à une exception près, les Etats-Unis, l'assurance privée a une place marginale et que les assureurs rendent rarement le service qui nous semblait leur au terme de la première partie, pouvoir relever de leur responsabilité.

Dans le chapitre huit, nous décrivons un certain nombre de réformes et de ré-organisations des marchés à l'initiative d'acteurs qui ont un pouvoir de régulation. Ces évolutions procèdent de la problématique que nous avons développée dans la seconde partie. Nous nous attacherons à souligner en quoi elles répondent à la nécessité de gérer la sélection afin d'améliorer l'efficacité productive. Nous montrerons qu'elles se traduisent par une forme de concurrence entre des opérateurs chargés de gérer les risques, qui peuvent être, entre autres, des compagnies d'assurance. Nous démontrerons en quoi cette concurrence, en toute rigueur, ne nous semble pas pouvoir être qualifiée de concurrence entre assureurs. Dans le même temps, nous détaillerons les conditions qu'il semble nécessaire de remplir pour atteindre une concurrence porteuse d'efficacité productive et nous présenterons la diversité des solutions techniques adoptées.

CHAPITRE 7

FONCTIONS DE L'ASSURANCE PRIVEE

DANS LES SYSTEMES MIXTES

A la question de savoir si la couverture du risque maladie peut relever exclusivement d'une assurance privée, la majorité des pays développés a répondu par la négative. Plus encore, l'assurance privée ne constitue la principale source de financement des dépenses de santé dans aucun de ces systèmes, pas même celui des Etats-Unis. La nature du partage entre les couvertures privées et socialisées des dépenses de santé est pourtant très variable d'un pays à l'autre, ce qui s'explique notamment par des critères historiques, sociologiques, et politiques. On oppose ainsi traditionnellement, en Europe, les systèmes d'assurance sociale Bismarckiens¹ (Allemagne, France...) et les systèmes nationaux de santé Beveridgiens² (Royaume-Uni).

Le principe Bismarckien consiste à l'origine à encourager des systèmes déjà existants de solidarité entre groupes professionnels. Les prestations (pensions diverses) ne sont pas des minima, mais sont proportionnelles aux cotisations (principe contributif) : la retraite est par exemple d'autant plus élevée que le salaire l'était. En matière d'accès aux soins, le principe contributif conduit à une couverture segmentée et proche de l'assurance : les prestataires de soins sont des producteurs indépendants, la couverture maladie rembourse les frais induits par le recours aux prestataires.

Le principe de Beveridge est de fournir à tout citoyen d'une nation un socle de droits sociaux minimaux, indépendants de son activité. En matière de prestations sociales (retraites, chômage), l'Etat social garantit donc un minimum à tous, étendu aux prestations en nature de l'assurance maladie et financé par l'impôt (principe universel). Il est assez naturel, dans une telle perspective, que l'Etat social salarie les prestataires de soins, ou au moins une partie d'entre eux, puisqu'ils fournissent ce service social.

¹ Bismarck fut chancelier de la Prusse, puis de l'Empire allemand dans les années 1865-80.

² De Lord Beveridge, rédacteur d'un rapport sur l'Etat Social au Royaume-Uni en 1942.

La ligne de démarcation entre les deux types de systèmes s'est toutefois estompée au fil du temps. Pour prendre l'exemple de la France, le basculement progressif des cotisations sur l'ensemble des revenus (par la Contribution Sociale Généralisée) et l'extension des droits à la Sécurité sociale sur un critère de résidence, prévu dans le projet de loi sur la lutte contre les exclusions³, font que le système s'éloigne progressivement du modèle originel.

Il nous semble donc intéressant d'offrir une ligne alternative de démarcation pour établir des comparaisons entre les systèmes de santé. Nous choisissons de mener une analyse transversale en fonction de la responsabilité que l'Etat détient dans la couverture des risques liés à la santé. Cette démarche nous conduira à étudier le rôle de l'assurance privée, car la place qui lui est dévolue est conditionnée par la nature et l'étendue de la couverture publique choisie par chaque pays.

Nous avons souligné dans la première partie de ce travail que l'assurance était théoriquement le moyen le plus efficace de se couvrir contre un risque individuel, même si les caractéristiques des biens considérés amènent les assureurs à moduler le service qu'ils proposent. L'objectif de ce chapitre est d'étudier le rôle effectif de l'assurance dans un grand nombre de systèmes de santé développés et de montrer qu'en pratique, elle joue un rôle assez marginal, même si elle répond aux attentes des personnes qui souscrivent les contrats et qu'elle remplit en cela une fonction importante.

Nous montrerons tout d'abord en quoi il nous semble intéressant de développer une typologie fonctionnelle pour analyser le marché de l'assurance (Section 1). Puis, nous en testerons la validité en l'appliquant à différents systèmes (Section 2).

³ Projet de loi portant création d'une couverture maladie universelle - Conseil des ministres du 3 mars 1999.

1. UNE APPROCHE METHODIQUE DES ARTICULATIONS PUBLIC-PRIVE

Un cadrage macro-économique simple montre que l'assurance privée a, de fait, un rôle de complément à la couverture publique dans l'ensemble des pays développés (1.1). Ceci dit, les systèmes de santé et de protection sociale ayant été élaborés selon des modèles différents, cette première analyse mérite d'être affinée pour déterminer en quoi l'assurance privée complète la couverture publique (1.2). Pour comparer, par la suite, les systèmes étudiés, nous utiliserons une démarche précise que nous présentons en dernier lieu (1.3).

1.1 Cadrage macro-économique

L'assurance se traduit, à quelques rares exceptions près, par le financement direct ou indirect des frais auxquels les individus peuvent devoir faire face pour se soigner. Pour estimer son importance en première approche, on peut étudier la part de l'assurance privée dans le financement des dépenses de santé (Tableau 1).

Tableau 1 : Part de l'assurance privée dans le financement des dépenses de santé (1995)

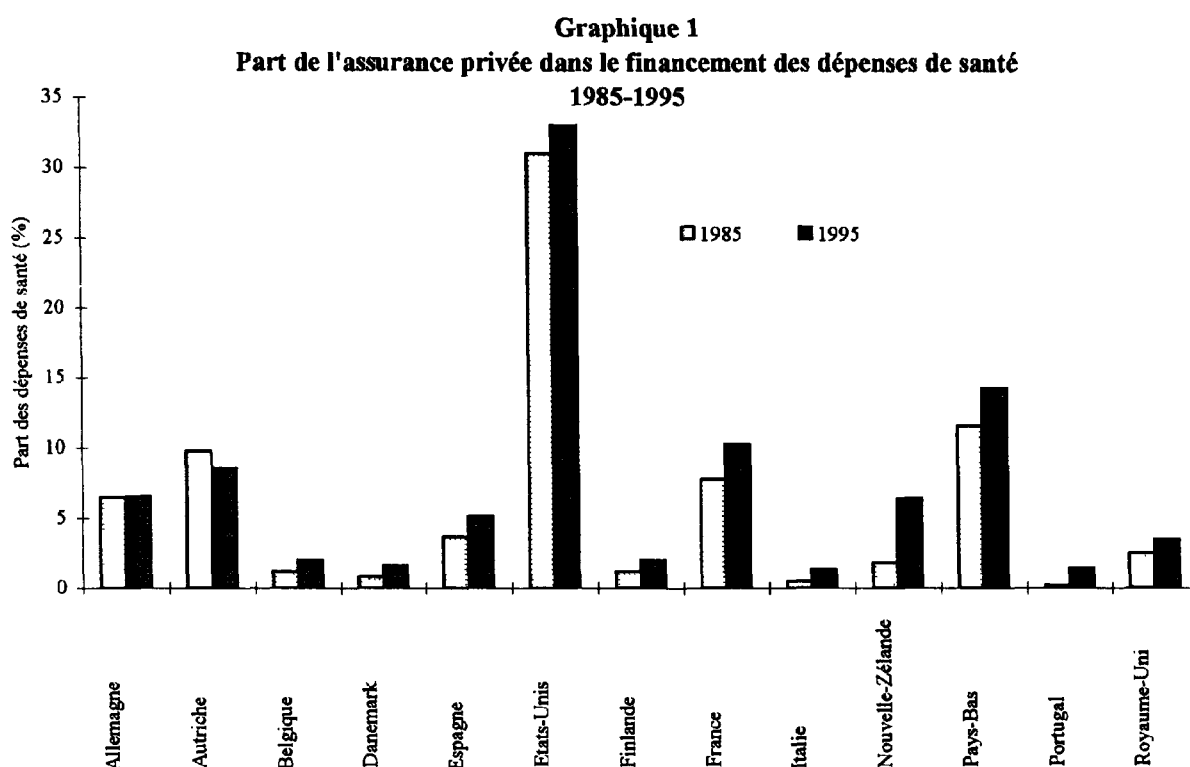
	Source OCDE* (%)
Etats-Unis	33
Pays-Bas	14,3
Suisse	11,6
France	10,3
Irlande	8,9**
Canada	8,8
Autriche	8,6
Allemagne	6,6
Espagne	5,2
Royaume-Uni	3,5
Belgique	2
Finlande	2
Danemark	1,6
Portugal	1,4
Italie	1,3
Grèce	nd.
Norvège	nd.
Suède	nd.

*OECD Health Data 98 : financement par l'assurance à but lucratif et non lucratif

**1996

La part couverte par l'assurance privée, qu'elle soit ou non à but lucratif, n'atteint dans les pays développés au maximum que 33% de la dépense totale de santé (Tableau 1). En outre, les Etats-Unis, qui sont au premier rang, se détachent nettement des autres pays. Aux Pays-Bas, qui occupent la seconde place, la part des dépenses de santé financée par l'assurance privée n'est que de 14%. La Suisse, qui était au troisième rang en 1995, a depuis réformé son système de santé et la part de l'assurance privée y a diminué.

Néanmoins, dans la quasi-totalité des pays pour lesquels l'information est disponible, le Graphique 1 montre que la part de l'assurance privée dans le financement des dépenses de santé a augmenté pendant la dernière décennie.



Source : OECD Health Data 98

La place dévolue à l'assurance privée, en augmentation régulière depuis plusieurs années, reste tout de même largement minoritaire, exception faite des Etats-Unis. La part de l'assurance sociale et des financement directs de l'Etat reste donc prépondérante dans tous les pays (50% au minimum aux Etats-Unis), même si elle a tendu à se stabiliser, voire à décroître depuis plusieurs années. De fait, l'intervention de l'Etat conditionne la marge de manœuvre et la taille du marché potentiel de l'assurance privée. A partir du

moment où des individus financent par des cotisations obligatoires ou par l'impôt l'accès à des soins ou à une assurance sociale, leur besoin, leur désir et leur capacité financière à recourir à une assurance privée sont largement réduits. L'assurance privée est donc complémentaire de l'assurance publique.

Cette notion de complémentarité, traditionnellement utilisée en France pour qualifier la couverture privée, doit être entendue plus au sens mathématique qu'au sens économique du terme. Si l'on considère assurance privée et publique comme deux biens, la notion de complémentarité économique renverrait à l'idée qu'à chaque niveau de couverture publique, il faut associer un niveau de couverture privée minimal, sans lequel l'individu ne tire aucune utilité de la "consommation" d'assurance publique. En retenant l'acception mathématique du terme, on considère que l'ensemble des dépenses de santé peut être a priori couvert par une assurance qu'elle soit de type public ou privé. Si un sous-ensemble est couvert par l'assurance publique, le champ *potentiel* de l'assurance privée est défini comme le complément de ce sous-ensemble. Cette seconde notion de complémentarité peut décrire plusieurs situations :

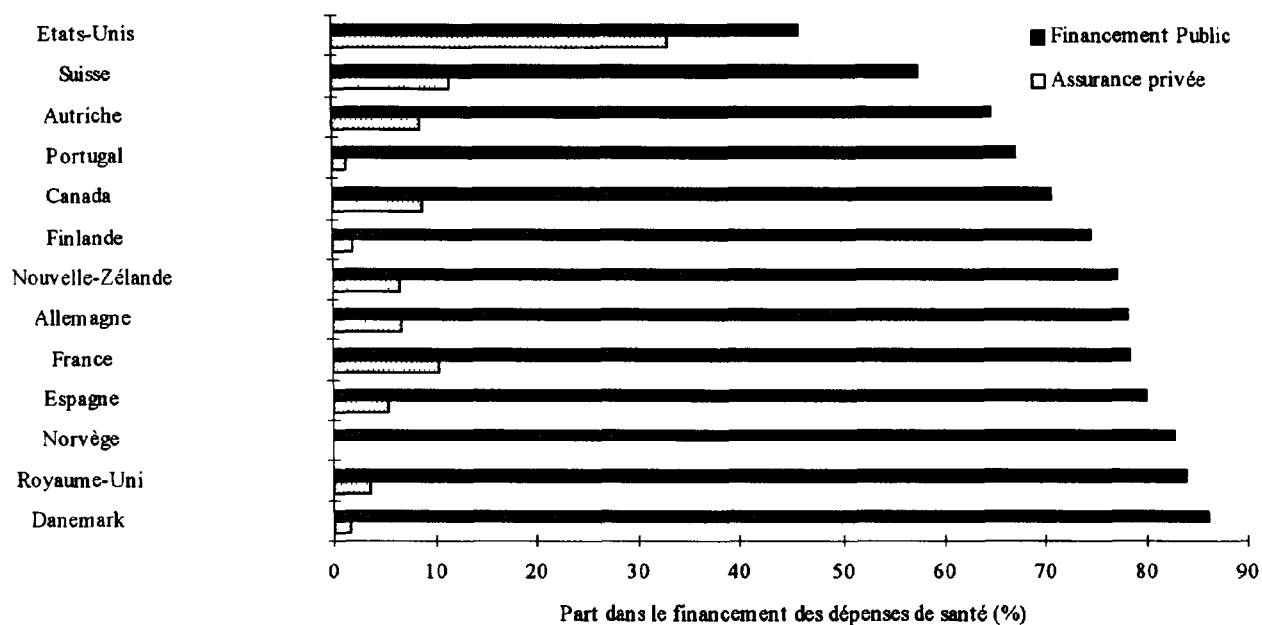
- si l'assurance publique couvre l'ensemble de la population mais pas l'ensemble des dépenses, les deux types d'assurance sont plutôt des substituts, au sens où chaque individu pourrait compenser la diminution (l'augmentation) de l'assurance publique par un recours accru (moindre) à l'assurance privée ;
- si l'assurance publique couvre une partie de la population, l'assurance privée est la seule source de couverture possible pour le reste de la population. Dans ce cas, il n'y a pas réellement de substitution possible.

Dans les deux cas cependant, le marché privé est largement défini par défaut et en référence à la couverture publique préexistante.

Le graphique 2 fournit une illustration de ce constat. Les pays⁴ sont classés en fonction de la part des dépenses de santé financées par l'assurance sociale et les financements publics directs.

⁴ Pays pour lesquels l'information est disponible dans ECO-SANTE OCDE 98.

Graphique 2
Parts de l'assurance privée et de l'ensemble des organismes publics dans le financement des dépenses de santé 1995



Source : OECD Health Data 98

On constate à nouveau que l'assurance privée ne finance qu'une part relativement minoritaire des dépenses de santé. Toutefois, ce constat masque des différences profondes qu'une analyse plus fine du rôle de l'assurance privée dans les différents pays va permettre de mettre en lumière.

1.2 Pour une typologie fonctionnelle

Quelques travaux récents étudient, pour différentes raisons, le rôle de l'assurance privée et leurs auteurs se sont attachés à construire des typologies pour le décrire. Ces études illustrent en général chaque fonction par un exemple connu mais ne cherchent ni à nuancer l'analyse pour chaque pays ni à appliquer de façon systématique les classifications proposées.

La typologie proposée par van Doorslaer et Wagstaff (1993) et complétée par Wasem (1995) distingue trois rôles possibles pour l'assurance santé volontaire, comme celle proposée par Rovira et al. (1998). Seul Schneider (1995) distingue explicitement quatre fonctions.

L'assurance volontaire peut, en premier lieu, selon van Doorslaer et Wagstaff, permettre de couvrir les copaiements imposés par le secteur public. Wasem élargit cette catégorie à l'ensemble des cas où l'assurance complète la couverture publique d'un individu, soit en couvrant des franchises ou copaiements laissés à sa charge, soit en finançant l'accès à des biens qui ne sont pas couverts par l'assurance publique. Il qualifie cette couverture de *complémentaire*. Rovira *et al.* nomment pour leur part cette assurance *supplémentaire* et distinguent trois cas : la couverture volontaire peut porter sur de nouveaux risques (les soins dentaires dans leur ensemble), ou pour un même risque, donner accès à d'autres services que ceux couverts (certains types de prothèses qui ne seraient pas couvertes par l'assurance publique), ou enfin limiter le risque financier que supportent les assurés associé à des copaiements du régime général. La typologie de Schneider est proche de celle de Wasem, mais il distingue explicitement l'assurance qui couvre les copaiements publics, qualifiée de *résiduelle*, et celle qui couvre des biens non pris en charge par le système public, dite *supplémentaire*.

Tous les auteurs distinguent ensuite deux autres fonctions.

D'une part, l'assurance privée peut dupliquer l'assurance publique. Elle est qualifiée alors de *supplémentaire* (van Doorslaer et Wagstaff), de *double couverture* (Wasem), d'*alternative* (Rovira *et al.* et Schneider), ou de *parallèle* (Schneider). Les trois premières études insistent surtout sur l'aspect institutionnel et sur le fait que les personnes qui souscrivent un tel contrat privé, ont les mêmes droits d'accès aux services publics que les autres. Schneider explicite le rôle de l'assurance volontaire parallèle : elle permet d'accéder à des services privés parce qu'il y a des listes d'attente dans le système public ou plus généralement parce que celui-ci est perçu comme offrant une qualité moindre.

D'autre part, l'assurance privée peut constituer la principale protection de certaines catégories de personnes. Elle est alors, pour Wasem, *alternative* à la couverture publique. Rovira *et al.* soulignent que cette fonction recouvre deux situations : les assurés volontaires n'ont pas d'assurance obligatoire parce qu'ils ne sont pas éligibles ou parce que la possibilité de sortir de l'assurance obligatoire était ouverte et qu'ils ont fait ce choix. Pour Rovira *et al.* l'assurance est alors dite *substitutive*, terminologie adoptée par Schneider.

L'ensemble de ces propositions est synthétisé dans le Tableau 2.

Tableau 2 : Terminologie utilisée dans les différentes typologies

Fonction de l'assurance	van Doorslaer Wagstaff	Wasem	Rovira et al.	Schneider
Couverture unique pour certaines catégories de personnes	n.d.	alternative	substitutive	substitutive
Couverture de biens hors panier public	n.d.	complémentaire	supplémentaire	supplémentaire
Couverture de copaiements publics				résiduelle
Duplication de la couverture publique	supplémentaire	double couverture	alternative	parallèle / alternative

On observe donc, malgré la difficulté d'aboutir à une terminologie commune et univoque, une convergence dans l'analyse des fonctions que l'assurance privée volontaire peut remplir. A différents types de systèmes publics, on peut associer différents "besoins" de complément privé. D'où l'idée d'établir une typologie explicitement fonctionnelle de l'assurance privée qui rende compte de la nature de ce besoin et indirectement du rôle de l'assurance privée.

Nous distinguerons quatre catégories, selon que l'assurance :

1. constitue la seule source de couverture contre le risque maladie. En France, on parle alors de couverture au premier franc (fonction 1) ;
2. finance des risques ou des biens non couverts par l'assurance publique (fonction 2) ;
3. re-finance des copaiements publics (fonction 3) ;
4. finance l'accès à une qualité supérieure (fonction 4).

La première fonction est celle qui semble, à la lumière du modèle théorique que nous avons présenté dans la première partie de cette thèse, le plus naturellement échoir à l'assurance privée puisqu'il s'agit de protéger les individus contre la réalisation de l'ensemble des risques liés à la santé. Toutefois, nous verrons que l'assurance ne joue que rarement ce rôle dans les pays développés et qu'en tout état de cause, cette fonction ne concerne jamais la totalité de la population.

On peut ensuite considérer que les fonctions 2 et 3 constituent des applications de la fonction standard de l'assurance (fonction 1) à des champs plus restreints. L'Etat garantit une couverture de base à la population concernée. Si cette couverture publique exclut explicitement certains biens et services, l'assurance privée peut remplir la fonction 2. Si l'assurance publique couvre une partie des dépenses liées à chaque consommation, l'utilisateur qui doit, à chaque contact avec le système de santé, couvrir une partie des frais, peut souscrire une assurance qui remplit alors la fonction 3.

De la même manière, la fonction 3 est un cas particulier de la fonction 2, dans lequel un bien est couvert partiellement. Si la couverture publique est faible, on se rapproche du cas de figure 2 (pas de couverture publique de certains services) voire de la configuration 1 (aucune couverture publique) si l'ensemble des biens et services est concerné.

Au total, la différence entre ces trois fonctions peut être décrite en termes de degré plus que de nature.

Le dernier cas est différent à cet égard. Dans les systèmes étudiés, l'Etat garantit un accès à la quasi-totalité des soins dans un système public mais les individus souscrivent une assurance pour, en cas de nécessité, avoir accès à des biens et services de meilleure qualité. Il s'agit toujours d'assurance, mais on ne peut pas considérer que la motivation première de l'individu est de se protéger contre le risque de devoir s'acquitter du prix des soins. L'assurance, dans ce cas, joue plutôt le rôle d'une option de sortie du système public.

1.3 Méthodologie

Dans une première étape, nous proposons une définition opérationnelle de l'assurance privée (1.3.2), avant de présenter les choix méthodologiques et leurs conséquences (1.3.2).

1.3.1 Définition opérationnelle de l'assurance privée

Dans la suite de ce travail, l'assurance est dite "privée" lorsque les trois conditions suivantes sont réunies.

a) L'assureur peut accepter ou non de couvrir la personne ou l'entité qui en fait la demande et l'entité peut choisir de se couvrir ou non

L'entité qui se couvre n'est pas nécessairement l'assuré lui-même. L'assurance privée peut, par exemple, être souscrite par un tiers au bénéfice de l'assuré, même si celui-ci ne choisit pas directement de se couvrir. En France, un employeur peut ainsi souscrire un contrat d'assurance maladie auquel ses employés sont obligés de contribuer. Pour autant, on considère que l'assurance est privée car l'employeur n'est pas obligé de souscrire et l'assuré peut, au moins théoriquement, échapper à cette obligation en changeant d'employeur.

b) Les assureurs sont en concurrence

Dans certains pays comme l'Irlande ou le Danemark, le marché est, pour des raisons historiques, dominé par une entreprise qui couvre la quasi-totalité du marché privé. Cette notion de concurrence correspond donc plutôt à l'idée que l'entrée sur le marché est possible.

c) L'assureur et l'assuré ont une grande marge de manoeuvre pour déterminer les différentes dimensions du service rendu

Ce choix porte spécifiquement sur :

- le champ des prestations et le niveau de couverture (*benefit package*). Des participations financières des assurés peuvent être prévues dans le contrat, ainsi que des limites au remboursement et des exclusions ;
- les primes. Elles sont liées au risque individuel ou au risque des cohortes couvertes (âge sexe, état de santé, profession..).

Tous les marchés d'assurance sont peu ou prou réglementés par l'Etat, en particulier pour obliger les assureurs à respecter un certain nombre de règles prudentielles afin qu'ils soient en mesure de faire face à leurs engagements. Pour des raisons qui tiennent d'une part, à la spécificité du risque maladie et aux imperfections de marché qu'elle entraîne et, d'autre part, à la reconnaissance par la société que la couverture des risques maladie représente un enjeu particulier, l'activité des assureurs intervenant dans le secteur de la santé fait l'objet de réglementations spécifiques. Il y a donc une part d'arbitraire dans cet élément de définition du champ de l'assurance privée qui

dépend du degré de réglementation des primes et des prestations. Si les forces de marché déterminent largement les caractéristiques de l'interaction entre les parties au contrat, l'assurance peut être privée (mais ce n'est pas une condition suffisante).

Cette définition appelle deux remarques :

- le statut privé, en particulier à but lucratif, de l'entreprise qui produit le service n'est pas intégré dans la définition. Ceci permet de prendre en compte l'ensemble des prestataires possibles du service d'assurance (Rovira *et al.*, 1998). Par exemple, des caisses publiques qui, au delà de leurs engagements obligatoires, offrent des services complémentaires optionnels, peuvent participer à l'offre d'assurance privée au sens où nous l'entendons ;
- l'élément le plus central de la définition est le fait de savoir si la personne a bien le choix de s'assurer et si l'assureur peut refuser de l'assurer. De fait, nous étudions l'assurance maladie en ce qu'elle est volontaire et que la décision d'assurance est strictement privée.

Illustrons cette définition par deux exemples.

1. Un certain nombre de pays ont instillé des éléments de concurrence dans une assurance maladie publique sans pour autant la "privatiser". Aux Pays-Bas, par exemple, une partie de la population est couverte par une assurance obligatoire dont les conditions sont définies par l'Etat et qui est financée par une cotisation proportionnelle au revenu. Depuis plusieurs années, les assurés peuvent choisir leur caisse d'assurance maladie ; pour autant, il ne s'agit pas d'une assurance privée. En effet, les assureurs ne peuvent refuser de couvrir une personne, tous les assurés doivent se couvrir, le niveau de couverture est défini par l'Etat et les primes ne sont pas déterminées ou prélevées par les caisses (sauf à la marge).
2. A l'inverse, un marché sur lequel opèrent des assureurs privés peut être réglementé à un point tel que l'on ne puisse plus qualifier, selon les termes choisis, l'assurance de privée. Ce cas de figure se présente en Suisse où la loi LAMal de 1994 introduit une obligation d'assurance pour l'ensemble de la population, interdit la sélection par les assureurs, définit précisément le champ des prestations de cette assurance et réglemente strictement la tarification. Les assureurs restent en concurrence et ils définissent globalement les primes de façon à pouvoir faire face à leurs engagements

(et donc en fonction du risque total assumé). Pour autant, la décision d'assurance n'est pas privée.

Dans tous les cas où l'individu a le choix entre différents opérateurs en concurrence mais où il n'a pas la possibilité de ne pas s'assurer ou de définir les prestations pour lesquelles il se couvre, l'assurance, selon notre terminologie, n'est pas privée.

1.3.2 Choix méthodologiques

Comme tout exercice de taxinomie, le travail entrepris ici nécessite de faire des choix et comporte des limites qu'il convient d'explicitier, par rapport au contexte dans lequel il est mené.

1. Tout d'abord, la classification étant articulée autour du rôle de l'assurance privée, elle conduira à rapprocher des systèmes qui diffèrent fondamentalement sur d'autres critères couramment utilisés pour les classer (mode de financement, caractéristiques de l'offre...). Ces rapprochements nous permettront de mettre en lumière des problématiques communes à des pays que l'on compare rarement.
2. Ensuite, la nature de la couverture publique, qui conditionne le champ de l'assurance privée, peut lui conférer plusieurs fonctions possibles parmi les quatre que nous distinguons. Si, dans deux systèmes, l'assurance privée remplit les mêmes fonctions mais dans des proportions différentes, chacun de ces systèmes pourra être classé dans un groupe différent. Cet exercice comporte une part d'arbitraire. Dans plusieurs cas, il sera ainsi difficile de séparer ce qui relève du refinancement des copaiements publics de ce qui procède du financement de services non couverts par le public. Dès lors, il n'est pas surprenant que les fonctions 2 et 3 soient confondues dans deux des quatre typologies existantes alors que, sur le plan conceptuel, la séparation est claire. Dans la mesure où il est possible de rattacher certains systèmes sans ambiguïté à une seule de ces fonctions, il nous semble intéressant de conserver cette distinction.

3. Dans le même ordre d'idée⁵, les soins dentaires sont largement exclus de l'analyse qui suit pour plusieurs raisons :

- d'une part, l'Etat considère souvent que son intervention n'est pas prioritaire dans le domaine dentaire. Ces soins sont donc l'objet de mesures particulières. Ils sont parfois exclus de la couverture publique, notamment pour certaines catégories de la population (adultes) ou font l'objet de mesures de déremboursement dans le cadre de plans de maîtrise des dépenses ;
- d'autre part, il est quasiment impossible, en ce qui concerne les prothèses qui constituent le gros des dépenses sur le risque dentaire, de distinguer ce qui relève de l'achat d'une prothèse "standard", qui répondrait à un besoin de santé "objectif" (fonction 2), de ce qui relève d'un achat d'un produit de qualité supérieure (fonction 4).

En fait, il conviendrait d'étudier séparément le dentaire, par type de bien et de service et dans certains cas, par type de population. Ainsi, en France, les soins dentaires relèvent d'un régime de copaiements comparable à celui qui s'applique pour les autres soins de ville (fonction 3). Par contre, les prothèses sont remboursées selon un barème fixé par la Sécurité sociale. Comme les prix pratiqués sont beaucoup plus élevés que ces tarifs, on peut quasiment considérer que les prothèses sont exclues de la couverture publique (fonction 2). Le prix d'une prothèse, pour un dentiste donné, varie fortement en fonction du matériau qu'il utilise (indépendamment du prix du matériau lui-même). Cette variation reflète des niveaux de qualité supposés ou réellement différents. Selon le taux de remboursement prévu au contrat, on pourrait estimer que l'assurance remplit la fonction 2 ou la fonction 4. Au total, si l'on prend en compte le rôle de l'assurance privée pour le dentaire dans le cadre d'une typologie générale, on risque non seulement de commettre des erreurs et mais aussi d'articuler la classification des pays autour d'un critère peu discriminant.

⁵ Je remercie les participants à l'ECuity II Workshop (Londres, 22-23 avril 1999) pour leurs nombreux commentaires sur ce travail et particulièrement Adam Wagstaff, qui m'a suggéré cette remarque, ainsi que Carol Propper, Eddy van Doorslaer et Brian Nolan.

4. Pour finir, l'exercice qui consiste à rapprocher le système de protection sociale d'un pays d'une fonction spécifique de l'assurance privée est compliqué par le fait que les régimes de protection sociale n'accordent pas les mêmes droits à l'ensemble des citoyens. Au sein d'un pays peuvent donc coexister plusieurs régimes qui laissent une place différente à l'assurance privée. On pourrait choisir d'étudier le régime "principal" d'un pays. Mais il est probable que selon le critère de choix du régime principal⁶, on serait amené à privilégier des sous-systèmes différents. Nous préférons donc étudier des sous-systèmes afin de rendre compte de la diversité du rôle de l'assurance au sein d'un pays. Dans cette optique, afin de conserver la cohérence de l'analyse, nous rendrons compte de l'ensemble des sous-systèmes d'un pays. Cette approche comporte une limite : dès lors que plusieurs sous-systèmes coexistent dans un pays, les chiffres disponibles au niveau macro-économique (notamment la part des dépenses de santé financée par l'assurance privée) doivent être désagrégés au niveau de chaque sous-système. Les sources d'informations comme l'OCDE ou les synthèses publiées permettent rarement d'atteindre ce niveau de détail. Nous avons cherché à obtenir des informations complémentaires auprès d'acteurs susceptibles de les détenir (fédérations professionnelles des assureurs, Ministères, chercheurs). Nous rendrons compte des résultats de cette recherche. Ce travail constitue une première étape qui permet déjà de souligner les limites d'analyses moins détaillées.

2. DECLINAISON DE LA TYPOLOGIE FONCTIONNELLE DE L'ASSURANCE PRIVEE

Nous nous proposons ici d'appliquer la typologie proposée pour classer les systèmes et, si nécessaire, les sous-systèmes d'une quinzaine de pays. Ce premier travail nous conduira tout d'abord à illustrer et valider la typologie.

⁶ On pourrait par exemple choisir de présenter pour un pays, le régime public qui couvre la proportion la plus élevée de la population, celui qui couvre la part la plus élevée des dépenses de santé, celui qui représente la part la plus élevée des dépenses publiques ou encore le régime public pour lequel le rôle de l'assurance privée est le plus important.

Dans le même temps, nous soulignerons les différences ou les convergences entre :

- d'une part, les systèmes dans lesquels l'assurance volontaire joue un rôle semblable et la dynamique propre à chaque fonction ;
- d'autre part, les systèmes dans lesquels l'assurance privée joue un rôle différent.

Nous étudions tout d'abord le cas où l'assurance couvre l'ensemble des risques liés à la santé, fonction qui semble, en référence à la théorie économique, pouvoir naturellement lui échoir (2.1). Puis, nous présentons les cas où l'assurance privée couvre des biens exclus de la couverture publique (2.2) et des copaiements laissés à la charge des assurés sociaux (2.3). Pour finir, nous nous pencherons sur les systèmes dans lesquels l'assurance a un rôle de nature différente et porte sur la qualité (2.4).

2.1 Couverture "au premier franc": fonction 1

Le premier cas que nous étudions ici, dans lequel l'assurance volontaire assume l'intégralité de la couverture des risques liés à la santé (fonction 1), s'avère être le moins répandu. L'assurance volontaire n'assume cette fonction que dans trois pays : les Etats-Unis, les Pays-Bas et l'Allemagne. Elle ne constitue la principale source de couverture de la population qu'aux Etats-Unis. Dans les deux autres pays, l'assurance volontaire couvre des proportions résiduelles de la population et concerne seulement les personnes dont le revenu dépasse un certain seuil.

2.1.1 États-Unis

Nous avons, dans un précédent chapitre, esquissé les contours du système américain : la quasi-totalité de la population âgée de moins de 65 ans (88% de la population) a pour seule source de couverture l'assurance volontaire. Passé cette limite d'âge, toute personne est couverte par Medicare, programme public sur les caractéristiques duquel nous reviendrons par la suite.

Sur le marché de l'assurance volontaire, la couverture du risque maladie s'effectue le plus souvent au sein de l'entreprise. Dans ce cas, la prime payée par l'assuré reflète le risque moyen de la main d'œuvre de son entreprise. La souscription n'est jamais obligatoire. Ce mode d'accès à l'assurance n'est possible que dans la mesure où l'individu est un actif employé et où son entreprise lui propose une assurance. En dehors du cadre de

l'entreprise, l'assurance à titre individuel est très coûteuse et ne concerne qu'une faible partie de la population âgée de moins de 65 ans.

L'assurance sociale pour les moins de 65 ans est conçue comme un filet de sécurité et couvre des populations jugées particulièrement vulnérables :

1. Medicare couvre quelques personnes qui présentent des niveaux de risque très aggravé ;
2. Medicaid, est un programme d'assurance géré par les Etats destiné à quelques catégories de personnes très défavorisées (notamment les femmes enceintes ou les enfants sous un certain seuil de revenu).

Ces filets de sécurité de nature assurantielle sont donc réservés à des catégories précises de la population et près de 20% des personnes de moins de 65 ans ne bénéficient d'aucune couverture, et peuvent parfois être prises en charge par des institutions publiques ou charitables.

Tableau 3 : Couverture maladie aux Etats-Unis, premier semestre 1996

	Couverture privée collective	Couverture privée individuelle	Couverture Publique	Pas de couverture
Personnes âgées de moins de 65 ans (88% de la population)	64%	5%	12%	19%
Total de la population	61%	7%	15%	17%

Source : Medical Expenditure Panel Survey Household Component (Round 1)

2.1.2 Allemagne et Pays-Bas

Dans ces deux pays européens, tout ménage dont le revenu est inférieur à un plafond légal, est couvert de façon obligatoire par un système public d'assurance. Le reste de la population est potentiellement concerné par l'assurance privée. Signalons que, dans ces deux pays, les assurés privés ne sont, en règle générale, pas soignés par des prestataires différents de ceux auxquels ont recours les assurés publics. La dichotomie entre le traitement des personnes aisées et des autres ne s'étend pas aux modes de prise en charge.

Quelles sont les spécificités des deux pays ?

En **Allemagne**, les employés dont le revenu excède un plafond légal et les indépendants ont le choix entre une adhésion à l'assurance publique (aux mêmes conditions que les assurés obligatoires) et une assurance privée.

L'assurance privée est volontaire à double titre : les personnes concernées choisissent de relever d'un régime privé, puis décident ou non de s'assurer. Ce dernier cas reste rare, puisqu'on estime que plus de 99,9% de la population bénéficie d'une assurance maladie⁷. La possibilité d'opter pour le régime privé est ouverte à environ 20% de la population. Une majorité reste dans le système public et l'assurance privée au premier franc ne concerne que 9% de la population totale (OECD, 1997). Plusieurs raisons peuvent expliquer cette relative désaffection :

- d'une part, sur le marché privé, les primes sont ajustées au risque. Une personne qui présente un niveau de risque élevé ou qui estime que son niveau de risque pourrait augmenter réalise donc un arbitrage et peut préférer s'acquitter d'une contribution proportionnelle au revenu (régime public) que d'une prime élevée ou destinée à augmenter ;
- d'autre part, l'assurance privée ne connaît pas le système des ayants droit. Si la personne dont le revenu dépasse le plafond choisit le régime privé, elle doit aussi acheter un contrat pour chaque autre membre du ménage alors que, dans le régime public, l'assurance couvre les ayants droit sans cotisation supplémentaire.

Les personnes qui présentent un niveau de risque élevé et / ou qui ne sont pas célibataires choisissent donc majoritairement le système obligatoire. Cette décision est encouragée par une réglementation au terme de laquelle, si une personne opte pour le privé, elle ne peut plus, par la suite, adhérer à l'assurance publique, à moins que son revenu ne passe sous le seuil légal. Cette règle vise à limiter les arbitrages opportunistes qui aboutiraient à ce que les personnes jeunes tirent avantage de primes faibles dans le privé et intègrent le public lorsqu'elles fondent une famille ou lorsque leur niveau de risque s'aggrave.

Pour atténuer l'effet de cette réglementation, depuis 1994, le marché privé est réglementé pour garantir l'accès des personnes âgées à l'assurance. Les assureurs privés doivent

⁷ Enquête auprès des ménages, 1995.

offrir un contrat standard aux personnes de plus de 65 ans qu'elles couvrent depuis au moins dix ans. Les garanties de ce contrat sont équivalentes à celles de l'assurance publique et la prime ne peut dépasser la moyenne des primes prélevées par des caisses publiques. Toutefois, ce contrat standard est "*de façon surprenante*" rarement souscrit (OCDE, 1997, pp.150). La fédération des assurances déclare que seulement mille personnes en bénéficiaient au 31 décembre 1997.

Aux **Pays-Bas**, l'assurance privée volontaire concerne uniquement les personnes dont le revenu dépasse un certain seuil, soit un peu plus de 30% de la population. Contrairement à l'Allemagne, elles n'ont pas la possibilité d'opter pour le système public d'assurance. La quasi-totalité de la population concernée est cependant couverte. Une analyse plus fine de la situation de ce pays, qui est un de ceux dans lesquels l'assurance privée semble jouer le rôle le plus important après les Etats-Unis, conduit à nuancer cette position pour deux raisons.

1. Tout d'abord, par rapport à l'Allemagne, le champ couvert par l'assurance volontaire est restreint. En effet, les résidents des Pays-Bas bénéficient d'une couverture socialisée et obligatoire des risques médicaux graves en vertu de la loi générale sur les frais exceptionnels de maladie (AWBZ). L'AWBZ couvre l'ensemble de la population et se limite théoriquement à une assurance des risques exceptionnels, qui, selon les termes de la loi, "*ne peuvent pas être assurés par le secteur privé à des taux acceptables*". Elle couvre les soins pour les pathologies chroniques lourdes (hospitalisation de plus d'un an), les handicapés, les malades mentaux (soins de jour et en institution) et les soins de long terme pour les personnes âgées : soins médicaux et aide à domicile ou en institution. Le principe de financement qui prévaut est celui d'une solidarité entre des individus qui ont un risque avéré et l'ensemble de la population. L'assurance privée des personnes aisées ne porte donc que sur le risque courant et aigu. Malgré tout, le champ de cette assurance privée est comparativement beaucoup plus vaste que celui de l'assurance volontaire du "risque mineur" des travailleurs indépendants en Belgique que nous étudierons par la suite. Nous avons donc choisi de classer les néerlandais aisés parmi les personnes qui ont pour seule source de couverture l'assurance volontaire et non parmi celles pour lesquelles une partie des biens et service est couvert par l'assurance publique (fonction 2). Ce choix est conforté par le fait que les prestations versées au titre de l'AWBZ constituent une grande partie des dépenses des personnes qui en bénéficient. Autrement dit, les

personnes qui utilisent les prestations de l'AWBZ basculent vers ce régime et sortent, de fait, du marché privé.

2. Mais surtout, si les assurances privées aux Pays Bas couvrent 30% de la population et financent environ 15% des dépenses de santé, il faut rappeler qu'il ne s'agit pas uniquement d'assurance volontaire au sens où nous l'avons définie. Le contrat standard WTZ (cf. Chapitre 5 §3.2), dont les conditions et le prix sont fixés par l'Etat sort du champ de l'assurance volontaire. Il représente 45% de la dépense couverte par le privé (cf. Chapitre 5).

L'analyse plus détaillée de l'exemple néerlandais montre donc que l'assurance privée concerne les personnes aisées, ou plus précisément celles d'entre elles qui ne présentent pas de risque particulièrement élevé.

2.1.3 Synthèse sur la fonction 1

Le tableau 4 présente les trois systèmes dans lesquels l'assurance volontaire constitue théoriquement la seule source de couverture de certaines catégories de personnes.

Tableau 4 : Assurance privée et couverture au premier franc

	Population concernée		Couverture de la population concernée			Proportion de la population totale couverte par l'assurance volontaire	
			Nature du filet de sécurité assurantiel	Répartition**			
		Proportion		Filet	Assurance privée		Non assurance
Etats-Unis	Population de moins de 65 ans	88%	Personnes démunies et très vulnérables (Medicaid) + quelques malades chroniques (Medicare)	12%	69%	19%	61%
Allemagne	Personnes aisées	20%	Assurance publique (pour tout volontaire) + contrat standard	55%*	45%*	0 ⁺	7%
Pays-Bas	Personnes aisées	30%	Contrat standard	15%	85%	0 ⁺	25%*

Sources : sources propres détaillées dans le texte
 * estimation sur la base des informations disponibles
 ** la somme des trois colonnes représente 100%

Compte tenu de l'imbrication des différents régimes d'assurance qui coexistent dans ces trois pays, on ne peut proposer qu'une estimation de la part des dépenses de santé couverte par une assurance volontaire dans le cadre de la fonction 1. Elle est naturellement la plus élevée aux Etats-Unis puisqu'au total 33% des dépenses de santé sont couvertes par l'assurance privée. Ce chiffre contient cependant la part des dépenses couvertes par les assurances privées au titre des couvertures qui complètent Medicare. En Allemagne, les 7% de la dépense totale couverts par l'assurance privée comprennent les dépenses des personnes qui souscrivent une assurance complémentaire au régime obligatoire (régime que nous étudions plus loin). Dans ces deux cas, on peut cependant estimer que la quasi-totalité de la part de la dépense totale couverte par l'assurance privée l'est au titre de la fonction 1. En revanche, pour les Pays-Bas, la part couverte par l'assurance volontaire, hors contrat standard, est probablement plus proche de 8% que de 15%.

Il serait intéressant de connaître, pour les personnes dont la principale source de couverture est l'assurance volontaire, la répartition des dépenses de santé entre différentes sources de financement (assurance privée / participation financière directe). Les données d'enquêtes nécessaires ne semblent toutefois pas disponibles.

Quoi qu'il en soit, le contraste entre les pays européens dans lesquels les individus obtiennent l'intégralité de leur couverture maladie sur une base volontaire et les Etats-Unis est frappant. Dans ce pays, l'assurance volontaire est la norme et l'assurance sociale, filet de sécurité, l'exception. Une proportion élevée de la population de moins de 65 ans ne bénéficie d'aucune couverture maladie. En Europe, l'assurance volontaire est réservée aux personnes aisées, ce qui tempère largement l'inéquité d'un système de tarification au risque. En outre, le marché privé et les droits à l'assurance sont structurés de façon à garantir à tous l'accès à une couverture, à des conditions financières jugées acceptables.

L'ensemble de cette analyse confirme que l'accès à l'assurance est bien une question centrale dans les systèmes où l'assurance est volontaire⁸.

⁸ Comme nous l'avons souligné dans le chapitre 6.

2.2 Couverture de biens exclus de la couverture sociale : fonction 2

Dans un grand nombre de pays, l'assurance sociale limite sa responsabilité à la couverture de certains soins ou de certains risques. On considère alors que le financement et l'accès aux biens et services médicaux exclus relève de l'initiative et de la responsabilité individuelle. L'assurance volontaire peut alors permettre de faire face à ces dépenses et son rôle exact dépend de celui de l'assurance sociale.

Parmi les exemples d'assurance volontaire remplissant cette mission, on peut distinguer deux groupes et établir une gradation selon la part du risque que l'assurance sociale assume. Dans le premier groupe, l'assurance volontaire porte sur un champ assez restreint et comprend des soins non prioritaires ou de confort. Dans le second, l'assurance sociale assume la couverture de dépenses exceptionnelles et l'assurance volontaire couvre une large proportion du petit risque.

2.2.1 Assurance volontaire marginale

Parmi les systèmes dans lesquels on peut considérer que l'assurance volontaire est marginale, trois d'entre eux : les Pays Bas, la Belgique et l'Allemagne sont plus pérennes que les autres (Israël et la Suisse). Ces deux derniers systèmes ont été réformés dans les trois dernières années et le rôle de l'assurance privée s'en est trouvé modifié.

a) Pays-Bas, Belgique, Allemagne

Aux **Pays Bas**, la majorité (63%) de la population est couverte par une assurance publique obligatoire qui porte sur un panier de biens assez large. Il est toutefois possible de souscrire une assurance facultative pour financer l'accès à des biens non couverts, comme par exemple les soins de kinésithérapie au delà d'une certaine limite et la fécondation in vitro. Le champ de l'assurance volontaire semble assez réduit : le ministère qui avait tenté d'exclure les prothèses dentaires de la couverture de base a été contraint de les réintégrer en 1997. L'assurance volontaire est en général servie par les caisses publiques qui fournissent l'assurance obligatoire du risque courant et elle est souscrite massivement par les personnes concernées. Schneider (1995, pp.65) donne un taux de souscription de 90%.

En Belgique, 85% de la population bénéficie d'une assurance maladie obligatoire qui est servie par des mutuelles. Chacun peut souscrire une assurance complémentaire pour des services additionnels, en général auprès de ces mêmes mutuelles. L'assurance libre complète les remboursements dans certains secteurs (prothèses et implants) mais offre surtout des services complémentaires en cas de maladie : services de convalescence, prêt de matériel médical... Ces services sortent du cadre strictement médical et constituent un argument commercial important dans la concurrence que se livrent les mutualités belges (Maarse et Paulus, 1998). Il est donc assez délicat de discerner la part véritable de cette assurance dans les dépenses de santé, mais elle est probablement très marginale. Schneider (1995) avance que 92% des personnes concernées souscrivent une assurance de ce type.

Statuer sur l'Allemagne s'avère un exercice difficile. Néanmoins comme la configuration générale du système est très proche de celle des Pays-Bas, nous choisissons de la conserver dans cette catégorie. L'assurance publique couvre environ 90% de la population et l'assurance privée peut proposer une couverture optionnelle. En fait, les assurances privées proposent un menu de services :

- couverture, comme en France, du forfait hospitalier journalier laissé à la charge de chaque assuré "public" ;
- couverture de soins auprès de praticiens non conventionnés par les caisses publiques et couverture de services hospitalier exclus de l'assurance de base (qui peuvent être des services de confort). En ce sens, l'assurance permet d'accéder à une meilleure qualité. Cependant, les documents décrivant l'assurance privée⁹ insistent plus sur le fait que ces services ne sont pas couverts par l'assurance publique que sur l'aspect qualitatif ;
- enfin, elle peut couvrir explicitement des biens ou des services comme les prothèses qui sont assez largement exclues de la couverture publique depuis 1997.

Il n'est pas possible de déterminer avec exactitude le nombre de personnes couvertes par le régime public qui souscrivent une assurance privée, car chacun peut souscrire différentes options séparément, parfois auprès de compagnies différentes, à titre

⁹ *Private Health Insurance : Facts and Figures 1997/1998*. Verband der privaten Krankenversicherung.

individuel ou au sein de contrats groupes. En 1997, la fédération des sociétés d'assurance allemande estime qu'environ 10% de la population couverte par le régime public bénéficie (au moins) d'une assurance volontaire.

b) Deux systèmes en transition : Israël et Suisse

En Israël et en Suisse, une assurance sociale universelle a été instaurée respectivement en 1995 et 1996 qui garantit à l'ensemble de la population l'accès à un panier standard. La mise en place du panier standard a cependant eu un effet inverse dans les deux pays sur l'assurance volontaire.

En Suisse, avant la réforme, la souscription d'une assurance maladie était facultative. Les caisses sociales d'assurance maladie étaient toutefois tenues d'offrir un panier standard minimal et de respecter des règles de tarification et de fonctionnement précises. Les personnes pouvaient ensuite compléter ce panier par une assurance volontaire (souscrite auprès des mêmes caisses ou des assurances privées). Ces assurances volontaires servaient avant tout à compléter le panier de base (fonction 2), mais finançaient aussi des copaiements assez élevés en matière hospitalière (fonction 3) et pouvaient couvrir des soins aigus dans des cliniques privées, remplissant en cela la fonction 4.

Depuis 1996, l'assurance de base est obligatoire et le panier standard a été étendu par rapport à la situation antérieure. Il comprend les soins prodigués par des médecins, l'hospitalisation (sans limite et y compris les frais d'hébergement). Ont été ajoutés, entre autres, les soins à domicile et en maison de retraite, les soins de psychothérapie et d'orthophonie prescrits.

La transition vers le nouveau système n'est pas complète, mais il apparaît d'ores et déjà que l'assurance volontaire jouera dans le futur un rôle différent. A l'heure actuelle, 35% de la population bénéficie d'une assurance volontaire et on estime que cette proportion devrait diminuer à 25% en l'an 2005 (Mosca, 1997), tendance que les observateurs attribuent à l'augmentation de la couverture de base. En réaction, il semble que l'assurance volontaire se recentre sur la fonction 4 et couvre plus souvent qu'auparavant des soins aigus dans des cliniques privées.

En Israël, jusqu'en 1995, l'assurance maladie est volontaire. Quatre-vingt-quinze pourcent de la population souscrit une assurance auprès d'une des quatre caisses semi-publiques, qui organisent la délivrance de soins sur un modèle proche du HMO

américain. Chaque caisse est libre de décider des garanties qu'elle offre mais les primes sont liées au niveau de revenu. La réforme de 1995¹⁰ impose le passage à l'assurance universelle : tous les Israéliens doivent maintenant souscrire une assurance de base auprès de la caisse de leur choix et le panier de biens et services couvert est défini légalement. Comme ce panier standard s'est avéré moins couvrant que certains contrats proposés dans la période antérieure, les caisses ont été incitées à basculer les services non compris dans le panier standard sur l'assurance volontaire. Elles peuvent offrir ces contrats, tout comme les assureurs privés. Ces derniers peuvent par ailleurs offrir une couverture de biens déjà compris dans la couverture de base.

Les éléments d'information disponibles sur l'assurance volontaire sont incomplets mais confirment que la tendance est à l'expansion de ce marché :

- les personnes couvertes par la principale caisse d'assurance (80% de la population) sont 16% de plus à avoir souscrit une assurance volontaire (Chinitz et Israeli, 1997) qu'avant la réforme ;
- les personnes inscrites auprès des autres caisses sont plus de 50% à avoir souscrit ou maintenu une couverture volontaire (ibid.).

Ce développement est encouragé par l'Etat car, depuis la réforme, les caisses font face à des problèmes de financement. La *Budget Reconciliation Bill*¹¹ de 1998 :

- souligne la volonté de limiter l'assurance de base en reportant explicitement sur l'assurance volontaire le financement des soins gériatriques ;
- encourage les caisses à développer de nouveaux produits d'assurance "pour rééquilibrer leurs budgets". Il se pourrait alors que les caisses développent des produits qui ne serviraient plus seulement à compléter la couverture de base mais financeraient l'accès à des structures privées (pour contourner des listes d'attentes, avoir accès à des cliniques privées...). L'assurance pourrait être amenée à remplir alors la fonction 4.

¹⁰ Que nous étudierons en détail dans le prochain chapitre.

¹¹ The Privatization of Health Care, Health Care in Israel. *The Israel Equality Monitor*, nov. 1998.

Ces deux expériences confirment la nature complémentaire de l'assurance volontaire dont le rôle évolue lorsque l'étendue de la couverture publique change. De plus, comme le système a récemment été remis à plat, il est plus facile d'y comprendre le rôle de l'assurance privée. Dans les systèmes plus anciens, l'Etat a pu, au fil du temps, et dans un objectif de maîtrise de ses dépenses, parfois diminuer les taux de remboursement des biens, parfois exclure certains services du panier. L'assurance privée pouvant se substituer à la couverture publique dans les deux cas, au titre de la fonction 2 ou 3, il est plus difficile d'établir à un instant donné quelle fonction l'emporte.

2.2.2 Assurance volontaire d'une gamme étendue de biens

Cette distinction entre la couverture des copayments, d'une part, ou des services non pris en charge par la couverture publique, d'autre part, est probablement plus pertinente dans le cas où l'assurance sociale couvre une gamme limitée de bien et services. Cette configuration est toutefois assez peu répandue et elle correspond à des sous-systèmes qui concernent de faibles proportions de la population :

- En **Belgique**, les travailleurs indépendants, qui représentent 15% de la population, n'ont pas de couverture publique en dehors des frais hospitaliers et des prestations techniques exigeant la compétence d'un médecin spécialiste telles que la radiologie, la biologie clinique... Entre 65 et 70% d'entre eux souscrivent une assurance volontaire qui couvre ce qu'on appelle - en Belgique - le "risque mineur" (soins ambulatoires, dentaires...).
- Les fonctionnaires **allemands** ¹² dont la moitié des dépenses de santé est couverte directement par l'Etat, peuvent compléter cette couverture par une assurance volontaire. Selon la fédération de l'assurance maladie privée (1999, pp.12) la quasi-totalité des fonctionnaires ont souscrit une assurance privée.

Peu d'information est disponible sur ces régimes particuliers et marginaux. Notons toutefois que ce type de fonctionnement semble susceptible de générer des incitations peu favorables à l'économie générale du système. Si seuls les soins très coûteux (par exemple l'hospitalisation) sont couverts par le régime public, des personnes qui ne

¹² Environ 2% de la population (estimation sur la base de Schneider 1995).

seraient pas couvertes par une assurance volontaire peuvent être confrontées à des barrières financières dans l'accès aux soins de ville. Or, lorsque la prise en charge d'une maladie par le système de soins est tardive, le coût total (qui sera en grande partie supporté par l'assurance sociale) peut être plus élevé.

2.2.3 Synthèse sur la fonction 2

Le tableau 5 synthétise les informations sur l'assurance volontaire lorsqu'elle porte sur les biens et services exclus de l'assurance publique.

Tableau 5 : Assurance privée et couverture de biens et services exclus de la couverture publique

	Population concernée	Proportion de la population concernée par la souscription de l'assurance volontaire	Taux de souscription
<i>assurance volontaire marginale</i>			
Pays-Bas	Bénéficiaires de l'assurance publique	63%	90%
Belgique	Bénéficiaires du régime général	85%	92%
Allemagne	Bénéficiaires de l'assurance publique	90%	10%
Suisse	Population entière	100%	35% (en baisse)
Israël	Population entière	100%	n.d. (en augmentation)
<i>assurance volontaire d'une large gamme de biens et services</i>			
Belgique	Travailleurs indépendants	15%	65%
Allemagne	Fonctionnaires	2%	100%

Sources propres détaillées dans le corps de texte

Si l'on observe les trois premiers pays dans lesquels le rôle de l'assurance privée volontaire n'a pas été modifié ces dernières années, on voit que le taux de souscription est beaucoup plus faible en Allemagne qu'aux Pays-Bas et en Belgique, alors que les paniers couverts par l'assurance publique dans ces trois pays sont comparables tout comme les primes de l'assurance volontaire, puisqu'elles correspondent à des services proches. Cette différence s'explique probablement par le fait que les caisses d'assurance publique allemandes n'ont pas le droit d'offrir l'assurance volontaire. A contrario, dans les deux autres pays, l'assurance est souscrite massivement, parce que l'assurance volontaire est utilisée comme argument commercial et comme moyen de fidélisation par les caisses. L'impression générale qui se dégage de l'ensemble de ces éléments est que l'assurance

volontaire, pour ces trois sous-systèmes, ne joue pas un rôle crucial dans l'accès aux soins et dans le financement des dépenses de santé.

La situation est plus contrastée dans les deux pays suivants, même si les réformes sont trop récentes pour qu'on puisse tirer des conclusions définitives. La Suisse a augmenté la couverture publique obligatoire. En passant à une couverture universelle, Israël a au contraire accru la part laissée à l'assurance volontaire. Les observateurs du système craignent que les mesures récentes n'entraînent une dérive rapide vers un système à deux vitesses, selon que les gens auront accès ou non à une assurance volontaire (Chinitz et Israeli, 1997).

Dans les deux derniers systèmes, on perçoit une nouvelle fois l'importance du revenu dans l'accès à l'assurance : le taux de souscription parmi les travailleurs indépendants belges est plus faible que parmi les fonctionnaires allemands. Ceci reflète le fait que les revenus de ces derniers sont globalement plus homogènes et plus élevés que ceux des travailleurs indépendants belges, dont la frange modeste n'a pas accès à l'assurance.

2.3 Refinancement des copaiements publics : fonction 3

Dans les systèmes que nous étudions ici, l'assurance privée re-finance la participation financière des usagers (fonction 3). Une large gamme de soins est prise en charge ou remboursée par la couverture sociale. Mais contrairement aux systèmes où l'accès est gratuit, des copaiements plus ou moins élevés peuvent être laissés à la charge des patients. L'objectif affiché de ce type de mesure est d'inciter financièrement les patients à consommer avec discernement, autrement dit, de limiter le risque moral. Par ailleurs, le copaiement diminue la part socialisée de la dépense, et constitue, en dehors de tout effet incitatif, un outil simple de limitation des dépenses publiques.

2.3.1 La France

En France, la Sécurité sociale couvre une palette très large de biens et services. Pour chaque consommation, une part plus ou moins élevée du coût reste à la charge des assurés sous forme de ticket modérateur. Ce ticket modérateur représente la différence entre le tarif du bien ou du service fixé par la Sécurité sociale, dit tarif de responsabilité, et le remboursement. Des exonérations de paiement du ticket modérateur sont prévues pour certaines pathologies lourdes ou sous condition de ressource. Elles concernent

environ 10% de la population. Les tickets modérateurs sont donc très régressifs : à niveau de consommation égal, toute personne supporte une charge identique, quelles que soient ses ressources. Les tickets modérateurs pèsent donc relativement plus sur les personnes dont le revenu est faible. Ils ne constituent en outre qu'une partie des dépenses que l'assuré social doit assumer. S'y ajoutent :

- le forfait hospitalier payé chaque jour par un patient hospitalisé,
- la différence entre le tarif fixé par la sécurité sociale et le prix effectivement payé. Ceci concerne :
 - les soins dispensés par certains praticiens qui appartiennent au secteur 2¹³ et qui sont autorisés à facturer des dépassements aux patients ;
 - certains biens, comme les lunettes et les prothèses dentaires, pour lesquels le tarif est très inférieur au prix effectif.

L'assurance volontaire sert principalement à refinancer cette participation financière. Selon le type de contrat, elle peut rembourser, outre les tickets modérateurs et le forfait hospitalier, les dépassements, la part des dépenses de lunetterie et de prothèse non prise en charge par la Sécurité sociale et éventuellement des suppléments de chambre individuelle en cas d'hospitalisation. En revanche, l'assurance volontaire ne couvre pratiquement jamais des biens et services qui ne sont pas pris en charge par la Sécurité sociale¹⁴. En France, 85% de la population souscrit une assurance volontaire, ce qui explique la part élevée de la dépense de santé financée par ce moyen (10,6%). L'assurance est souscrite et payée par les individus et leurs employeurs (deux tiers des personnes sont couvertes dans le cadre de contrats de groupe).

A l'origine, le ticket modérateur était conçu comme une participation financière destinée à inciter les individus à consommer avec discernement. Or, le remboursement du ticket modérateur par les assurances gomme littéralement cette incitation. En revanche, les augmentations successives du ticket modérateur, destinées à contenir les dépenses de la Sécurité sociale, se sont traduites par une diminution des taux de remboursement des

¹³ Le secteur 2 est ouvert sous certaines conditions à des médecins qui ont la liberté de déterminer "avec tact et mesure" leurs honoraires et doivent théoriquement délivrer un service de meilleure qualité.

¹⁴ Les lentilles jetables constituent une exception courante dans les contrats offrant des garanties étendues.

soins, qui a touché plus durement les personnes non couvertes par l'assurance volontaire, en moyenne beaucoup plus défavorisées que le reste de la population. Au total, la participation financière des assurés sociaux au paiement de chaque soin est assez élevée, malgré un accès universel à l'assurance sociale et rend le système globalement inéquitable. L'introduction de la Couverture Maladie Universelle (chapitre 6) vise à pallier cette inéquité et à garantir l'accès aux soins des personnes les plus démunies.

2.3.2 Medicare

Au Etats-Unis, le programme d'assurance sociale Medicare, financé par l'Etat fédéral, s'adresse essentiellement aux personnes de plus de 65 ans, sans conditions de revenu, mais aussi aux sujets handicapés ou souffrant de maladies chroniques âgés de moins de 65 ans. Medicare représente, en 1996, 34% des dépenses de santé¹⁵, et couvre 38,1 millions de personnes (13,2% de la population), dont plus de 86% sont des personnes âgées (HCFA Statistics, 1997). Medicare recouvre en fait deux programmes distincts :

- la partie A garantit le remboursement des dépenses hospitalières, moyennant paiement d'une franchise et acquittement de copaiements croissants au delà du 60^{ième} jour d'hospitalisation et dans la limite de 150 jours d'hospitalisation par période de couverture¹⁶.
- la partie B assure un remboursement partiel des dépenses autres qu'hospitalières à hauteur de 40%, en raison d'un système de ticket modérateur, de franchises, et surtout parce que les médicaments ne sont pas pris en charge. Pour accéder à la partie B, les assurés doivent verser des cotisations mensuelles, mais l'essentiel du programme est financé par l'Etat fédéral (75%).

Au total, 94% de la population concernée par Medicare bénéficie des parties A et B. La participation financière des bénéficiaires est assez élevée et les personnes couvertes par Medicare souscrivent des assurances volontaires. Elles leur permettent, avant tout, de refinancer les copaiements et les franchises imposées par le système public et elles couvrent, dans certains cas, des biens et services non couverts par l'assurance publique. En 1996, selon l'enquête MEPS (*Medical Expenditure Pannel Survey*), environ 62% des

¹⁵ National Health Expenditures, 1996.

personnes de plus de 65 ans, qui représentent la majorité des personnes couvertes par Medicare, sont aussi couvertes par une assurance privée.

2.3.3 Les pays scandinaves

Les pays scandinaves, pour ce qui concerne le rôle de l'assurance volontaire, pourraient a priori être assez proches de la France. Les systèmes de santé de ces pays sont majoritairement financés par l'impôt et offrent une couverture universelle. Ils sont marqués par une composante locale forte dans l'organisation comme le financement des services. Les services de santé sont majoritairement publics et il n'existe pratiquement pas d'exercice libéral pur ou de structures sanitaires privées¹⁷.

Au **Danemark**, l'assurance volontaire joue un rôle similaire à celui de la France. L'enjeu y est cependant beaucoup moins important, car l'accès aux soins est largement gratuit. L'assurance volontaire complète le remboursement de biens et services partiellement pris en charge par le système public, en particulier les médicaments, les prothèses dentaires et l'optique. Au milieu des années quatre-vingt-dix, 26% environ de la population avait souscrit une assurance privée.

En **Norvège** et en **Suède**, les patients paient une partie des soins dont ils bénéficient, sous forme de paiement forfaitaire par type de soin. Dans ces deux pays, l'assurance privée joue pourtant un rôle très marginal. La raison fondamentale de ce phénomène réside probablement dans le fait que ces deux pays ont instauré une limite aux dépenses auxquelles les individus doivent faire face au cours d'une année (franchise sociale). La **Finlande**, qui applique un tel type de franchise aux médicaments et aux consultations médicales, envisage de généraliser ce système afin de limiter la participation financière, jugée inéquitable, des usagers. Les paiements résiduels par les individus représentent en effet une part assez élevée des dépenses (21.5%). Dans ce dernier pays, l'assurance volontaire se concentre cependant sur la fonction 4, que nous étudions par la suite.

¹⁶ Une période de couverture commence le jour où une personne entre à l'hôpital et prend fin lorsqu'elle est sortie de l'hôpital depuis plus de 60 jours.

¹⁷ Espace Social Européen (1995).

2.3.4 Synthèse sur la fonction 3

Le tableau 6 synthétise les informations concernant l'assurance privée dans les pays où elle sert principalement à financer les copaiements pris en charge par les assurés sociaux.

Tableau 6 : Assurance privée et copaiements publics

	Pourcentage souscrivants*	Pourcentage de la dépense couverte**
France	85%	10,3%
Medicare (~13% de la population)	62%***	n.d.
Danemark	26%	1,6%
Suède	0+	0+
Norvège	0	0

Sources : * sources propres détaillées dans le texte
** identiques au tableau 1
*** taux de souscription pour les plus de 65 ans

Lorsque l'assurance privée complète le financement de l'accès à une large gamme de soins, elle est souscrite assez massivement, comme en France et aux Etats-Unis pour Medicare. Dans ces deux cas, le fait de ne pas pouvoir souscrire une assurance volontaire peut poser de réels problèmes d'accès aux soins. En France, on sait que le montant des copaiements constitue, avec le fait que les assurés sociaux doivent payer les prestataires de soins avant de se faire rembourser (avance de frais), une des limites à l'accès aux soins de ville des plus défavorisés. A l'inverse, dans les pays qui ont mis en place des systèmes de franchise sociale limitant l'impact des copaiements sur le budget d'un ménage, l'assurance privée est quasiment inexistante.

Pour les systèmes publics dans lesquels les copaiements sont généralisés sans être modulés en fonction du revenu, l'accès à l'assurance volontaire joue un rôle majeur dans la capacité qu'ont les personnes d'avoir recours au système de soins (Rochaix, 1995). En France, l'idée d'instaurer un plafonnement des tickets modérateurs en fonction du revenu, avancée dans le rapport *Santé 2010* (1993, pp.437) n'a jamais vraiment fait son chemin et la réforme actuelle qui vise à introduire une couverture maladie universelle cherche à corriger l'inéquité générée par les copaiements pour les plus démunis. Aux Etats-Unis, de façon comparable, le programme Medicaid, destiné aux personnes démunies apporte un complément de couverture à 15% de la population couverte par Medicare.

Par ailleurs, la possibilité de souscrire une assurance volontaire pour rembourser les copaiements illustre la difficulté de mettre en œuvre une participation financière efficace : si l'instauration des copaiements a pour objectif d'inciter les assurés à consommer avec circonspection, et si cet objectif est atteint pour les personnes qui ne sont pas couvertes, il est gommé lorsque la couverture volontaire finance les copaiements.

2.4 Accès à une qualité supérieure : fonction 4

Dans le cas où l'assurance privée permet d'avoir accès à une plus grande qualité de service (fonction 4), on peut distinguer deux groupes de systèmes. Dans le premier, l'assurance privée duplique complètement l'assurance publique. Dans le second, l'architecture globale du système prend plus explicitement en compte l'existence de l'assurance privée (Irlande et Autriche).

2.4.1 Effet doublon complet et système de santé parallèle

Dans les systèmes de santé de type N.H.S.¹⁸, l'offre de soins est essentiellement publique. Le système de santé est financé par l'impôt et l'ensemble de la population bénéficie d'un accès gratuit à la quasi-totalité des soins. La régulation du système passe par celle de l'offre, ce qui se traduit en général par une forme de rationnement et des listes d'attente pour certains services. Dans les systèmes de ce type, l'assurance privée permet de financer l'accès à des prestataires de soins privés. Elle est donc réservée à une faible proportion aisée de la population. En effet, comme la souscription d'une assurance privée ne limite pas les droits d'accès au système public, elle ne permet pas de se dispenser des contributions à son financement. La couverture privée duplique donc la couverture publique, sur le plan du financement et elle contribue au financement d'une duplication de l'offre dans un système privé parallèle. La souscription d'une assurance privée permet à l'assuré d'avancer le moment d'un traitement, d'avoir un plus large choix de prestataires, de bénéficier d'un plus grand confort hospitalier. Autrement dit, elle permet aux assurés d'avoir accès à une prise en charge de meilleure qualité que celle dont ils pensent pouvoir bénéficier dans le service public.

¹⁸ National Health System, Système National de Santé s'inspirant du modèle Beveridgien.

a) Grande-Bretagne

L'exemple emblématique est le système de santé de la Grande-Bretagne. L'assurance privée finançait au début des années quatre-vingt-dix essentiellement des soins aigus dispensés au sein d'hôpitaux privés vers lesquels les personnes aisées choisissaient de s'orienter (Propper, 1992). Il semble qu'elle se soit depuis étendue vers les soins de long terme et à domicile pour les personnes âgées et les malades mentaux. Selon une estimation récente (Evans, 1998), environ 12% de la population a souscrit une assurance privée santé¹⁹. Cette proportion est très variable selon les régions : le taux de souscription atteint 25% dans le sud de l'Angleterre, où les revenus sont plus élevés et l'activité économique plus florissante que dans le reste du pays. La séparation entre système public et privé n'est toutefois pas complète car les acteurs locaux, qui, dans le système britannique, achètent les soins pour le compte de leurs patients peuvent, depuis le début des années quatre-vingt-dix, passer des contrats avec des hôpitaux ou des structures privées²⁰. Dans des domaines comme les soins de long terme où l'offre publique est nettement insuffisante, des patients du NHS peuvent, sans participation financière supplémentaire, bénéficier de services privés. Il existe une certaine dichotomie entre les prises en charge publiques et privées et la dérive vers un système à deux vitesses est une préoccupation constante des régulateurs selon Propper²¹.

b) Europe du sud

A la fin des années soixante-dix, plusieurs pays de l'Europe du sud ont choisi de mettre en place des systèmes nationaux de santé inspirés du modèle britannique parmi lesquels la Grèce, l'Espagne et l'Italie. Pour ces trois pays, ce passage à un système national traduit en partie la volonté de rupture avec les régimes politiques de la période précédente dans lesquels "*corporatisme et mutualisme s'avéraient être les éléments fondamentaux du système de protection sociale*" (Duriez et Lequet, 1996, pp. 62). La réforme procède donc d'une recherche d'harmonisation républicaine et de reconstruction de la cohésion

¹⁹ Si l'on prend en compte les personnes qui ont choisi une assurance qui leur verse directement des compensations monétaires forfaitaires en cas de maladie mais qui ne rembourse pas les soins, la proportion d'assurés privés est de 17%.

²⁰ Ces acteurs sont les *District Authorities* (autorités locales), les généralistes gestionnaires de budgets. La réforme mise en œuvre depuis le premier avril 1999 les a fait disparaître au profit des *Primary Care Groups*. Les PCG devront à terme acheter les soins pour les usagers.

²¹ Entretien, Londres 1999.

sociale. En Italie, l'objectif politique cherché dans le passage au NHS, toujours selon Duriez et Lequet, est plutôt "*la relégitimisation de la classe politique dans la gestion des affaires publiques à une époque où l'évolution économique génère un fort risque d'explosion sociale (extrémisme politique, développement de mouvements terroristes)*" pp.63.

L'ensemble de ces pays était marqué par une grande diversité inter-régionale de l'offre de soins, qu'un système de type NHS semblait a priori bien armé pour réduire. Mais les conditions économiques difficiles ont freiné les réformes et le développement du secteur privé d'offre de soins et d'assurance s'explique en partie par cet échec relatif. Il se traduit, beaucoup plus qu'au Royaume-Uni, par une dérive vers un système à deux vitesses.

Ainsi en est-il du système de protection sociale en Grèce qui est resté très désorganisé : la qualité des services publics est inégale ; un secteur privé parallèle s'est développé et les paiements occultes sont très élevés, en particulier dans le secteur public hospitalier. Il est donc impossible de déterminer avec exactitude la part des dépenses de santé financée par les particuliers. Le ministère de la santé estime en 1994 qu'elle est environ de 30% de la dépense. Sur la base d'une enquête, Liapopoulos (1995) avance que 10% de la population a souscrit une assurance volontaire. Cette assurance finance l'accès aux services privés ainsi qu'à des services dont la couverture n'est pas assurée par le système public.

En Espagne, la population bénéficie de l'accès à un système national de santé mais à l'instar des Grecs ou des Britanniques, une partie de la frange aisée de la population souscrit une assurance privée. Elle donne accès à un système privé parallèle et permet de contourner les listes d'attentes considérables. En 1995, on estime (Costes, 1995) qu'entre 15 et 20% des Espagnols ont souscrit une telle assurance. Elle donne accès à des réseaux de soins privés (dans 90% des cas) ou elle rembourse (10% des cas) les frais engagés pour bénéficier des soins dispensés par des prestataires privés (spécialistes en particulier). L'importance que les Espagnols attribuent à la différence de qualité entre les secteurs public et privé est illustrée par le comportement des fonctionnaires qui représentent, avec leurs ayants droit 4% de la population. Ils ont la possibilité de choisir un opérateur privé pour mettre en œuvre leur couverture (définie et financée par l'Etat) dans le cadre d'une délégation de gestion. En 1995, 85% des fonctionnaires (Pellisé, 1995) ont choisi de bénéficier des services mis en place par les assureurs qui ont constitué des réseaux de soins privés, plutôt que de rester dans le NHS. Contrairement aux autres Espagnols

cependant, les fonctionnaires n'ont pas à payer deux fois pour avoir accès aux services privés.

L'Italie est dans une situation proche de celle de l'Espagne. Les chiffres de l'OCDE, qui évaluent la part du financement par les assureurs privés à 1,3% en 1995, donnent à penser que l'accès à des soins privés est plus souvent financé directement par les patients que par l'assurance. Les informations disponibles sur le secteur privé sont assez peu fiables et Mapelli (1995) estime qu'environ 20% de la population a souscrit une assurance volontaire et que l'assurance privée couvre en réalité, en 1991, 10% des dépenses de santé. Mapelli souligne toutefois, comme l'OCDE, que les Italiens semblent préférer avoir recours à l'épargne qu'à l'assurance, dans laquelle ils ont une confiance limitée, pour financer leur accès à des soins privés.

c) La Finlande

En Finlande²², nous avons signalé que l'assurance privée permet aux assurés de bénéficier de services privés. Selon l'enquête santé de 1995/1996, 12% de la population souscrit une telle assurance. L'assurance est utilisée pour couvrir les dépenses liées à des consultations auprès de médecins privés. Elle est majoritairement souscrite au bénéfice des enfants (la proportion d'enfants de moins de 14 ans couverts est proche de 31%). La souscription d'assurance volontaire est concentrée dans les régions plus riches (taux de souscription standardisé par âge-sexe à Helsinki : 16%) et parmi les personnes aisées (7% dans le quintile de revenu le plus faible, 16% pour le plus élevé).

2.4.2 Effet doublon atténué et système de santé intégré

L'Autriche et l'Irlande sont deux pays dans lesquels l'assurance privée a aussi pour fonction première de financer l'accès à une meilleure qualité. Cependant, à la différence des pays mentionnés précédemment, cette possibilité est plus explicitement prise en compte dans l'architecture globale du système et le sentiment qu'il existe un secteur privé, entièrement disjoint du secteur public, est beaucoup moins prégnant.

²² Merci à Unto Hakkinen et Jan Klavus de m'avoir aidée à comprendre le fonctionnement de l'assurance dans leur pays.

a) Autriche

Le système autrichien est fondé sur un modèle d'assurance sociale et son organisation tranche radicalement avec les systèmes de type NHS. Néanmoins, l'assurance volontaire y joue un rôle similaire : elle permet de financer l'accès à une meilleure qualité de service. L'assuré social peut ainsi choisir de consulter un prestataire de soins qui n'est pas conventionné par sa propre caisse. Cette possibilité est utilisée pour réduire le temps d'attente pour des examens. Quand il opte pour un prestataire non conventionné par sa caisse, l'assuré est moins bien remboursé par son assurance obligatoire. De même, il existe en Autriche deux classes d'hospitalisation. Si l'individu choisit la classe spéciale (confort supérieur), sa caisse verse un prix de journée à l'hôpital qui varie selon l'établissement et le patient doit compléter le paiement. L'assurance privée permet de financer ces dépenses et aussi accessoirement de compléter les remboursements pour des biens et services partiellement couverts par l'assurance publique (soins dentaires de luxe et quelques produits médicaux). Environ 35% de la population souscrit un tel type d'assurance (Ministère fédéral du Travail, 1998).

Le parallèle avec la situation des pays de type NHS, précédemment cités, repose sur l'idée que l'assurance volontaire permet d'obtenir une meilleure qualité de service. Il existe cependant deux différences fondamentales :

- dans le système autrichien, la "qualité" se mesure avant tout en terme de confort pour l'individu. Le coût de ce confort est explicité et les assurés sociaux peuvent en bénéficier, moyennant une participation financière plus élevée. Pour autant, ils n'ont pas accès à un système parallèle comme c'est le cas dans les pays de type NHS : s'ils choisissent de consulter un praticien qui n'est pas conventionné par leur caisse, celui-ci est en général conventionné par au moins une autre caisse. La rémunération du médecin ne provient pas seulement de son activité privée. De même, lorsque les patients choisissent l'hospitalisation de classe spéciale, ils sont soignés dans les mêmes établissements que le reste de la population ;
- par ailleurs, l'assurance privée ne duplique pas la couverture publique de façon aussi flagrante que dans les pays précédemment étudiés :
 - d'une part, un patient qui choisit le service de meilleure qualité ne paie pas le service de base auquel il aurait droit, puisque les prestataires ne sont rémunérés que s'ils prennent en charge un patient. A l'inverse, dans un système

de type NHS, on ne peut se dispenser de sa part du financement obligatoire même si on renonce à bénéficier des soins gratuits offerts à tous ;

- d'autre part, l'assurance sociale participe au financement des soins de meilleure qualité que l'individu choisit : pour les soins ambulatoires, la caisse rembourse à l'individu 80% de la somme qu'il recevrait s'il consultait un prestataire conventionné, l'assurance privée remboursant la différence. Par contre, la participation de la caisse publique au financement de l'hospitalisation de classe spéciale semble être nettement plus faible que pour l'hospitalisation de classe normale. Pour ce volet, on peut donc considérer que le renoncement au "droit à la classe normale" a un coût d'opportunité qui s'ajoute au coût direct que la prime d'assurance représente, mais là encore il n'y a pas duplication du paiement du service.

b) Irlande

En Irlande, l'ensemble des soins est fourni gratuitement à 35% de la population dont le revenu est inférieur à un seuil légal. Le reste de la population a gratuitement accès à l'hôpital public, mais fait face à l'intégralité de la dépense pour les soins ambulatoires²³.

Quarante pour-cent de la population irlandaise souscrit une assurance volontaire.

Paradoxalement, cette assurance couvre peu les soins ambulatoire. Elle joue plutôt un rôle de filet de sécurité : les assurés peuvent se faire rembourser - partiellement - lorsque leurs dépenses dépassent une franchise d'environ 7000 FF²⁴ par an et par ménage (et dans la limite d'un plafond). Le rôle principal de l'assurance privée est de prendre en charge les frais d'hospitalisation en lits privés, installés au sein des hôpitaux publics. Elle permet dans ce cas au patient de contourner les listes d'attentes, de choisir son hôpital et de s'assurer qu'il sera suivi par le même spécialiste tout au long de son traitement. Enfin, elle offre un meilleur confort aux patients hospitalisés.

²³ Quelques rares exceptions : prise en charge des maladies infectieuses, certains soins pédiatriques et quelques dépistages.

²⁴ Source : entretien avec Aisling Kennedy, MERCER déc. 1998. La franchise inscrite au contrat est moins élevée que celle donnée ici (environ 4000F pour un ménage). Cependant, le tarif maximal que l'on peut comptabiliser pour chaque soin afin de déterminer si on atteint la franchise est inférieur au prix payé. Le ménage ne pourra être remboursé par l'assurance qu'après avoir déboursé directement 7000FF.

Pour la partie hospitalière, l'assurance privée duplique la couverture offerte par le public. En effet, les soins hospitaliers dans le service public sont ouverts à tous. Les patients privés, qui n'exercent pas ce droit, paient en réalité deux fois l'accès à l'hôpital (une fois par l'impôt qui finance le service public et une fois par le biais de l'assurance pour accéder au secteur privé). Cet effet de doublon est cependant limité par des avantages fiscaux liés au paiement des primes et par le fait que les hôpitaux facturent aux patients privés et aux assureurs un montant inférieur au coût réel des séjours hospitaliers (Smyth et O'Faherty, 1995). L'Etat trouve un avantage à cette situation et cherche à maintenir cet équilibre pour éviter une dérive trop forte vers un système de prise en charge à deux vitesses. Ainsi, à l'heure actuelle la quasi-totalité des praticiens hospitaliers partagent leur temps entre activité privée et publique. Il n'y a donc pas de réelle différence dans la qualité des soins prodigués aux patients privés et publics. En revanche, les patients privés sont soignés plus tôt que les patients publics.

2.4.3 Synthèse sur la fonction 4

Le tableau 7 synthétise les informations disponibles sur les systèmes dans lesquels l'assurance volontaire permet d'accéder à une meilleure qualité.

Tableau 7 : Assurance privée et qualité

	Pourcentage souscrivant*	Pourcentage de la dépense couverte**
UK	12%	3.5%
Espagne	15-20%	5.2%
Italie	20%	1.3% ou 10% (Mapelli)
Grèce	10%	n.d.
Finlande	12%	2%
Autriche	35%	8.6%
Irlande	40%	8.9%

Sources : * sources propres détaillées dans le texte
 ** identiques au tableau 1, sauf mention contraire.

Dans l'ensemble de ces pays, la totalité de la population est théoriquement concernée par l'assurance privée au titre de la qualité, mais les taux de souscription restent modestes.

Une fois encore, le revenu joue un rôle central dans l'accès à l'assurance volontaire : seules les personnes relativement aisées peuvent payer deux fois l'accès aux soins. On remarque d'ailleurs que, dans les pays où l'effet de doublon sur le financement est atténué, la souscription est beaucoup plus massive et la part des dépenses de santé financée par l'assurance privée est plus élevée. L'équité dans les systèmes de ce type n'est donc pas une préoccupation de court terme, puisque l'accès aux soins est théoriquement garanti à l'ensemble de la population. En revanche, si l'écart de qualité entre les systèmes privés et public est très élevé, une déstabilisation du système qui se veut solidaire est à craindre : les personnes aisées se soignant dans des systèmes parallèles n'ont pas intérêt à pousser à l'amélioration de la qualité d'un système public dont elles se détournent. A terme, elles pourraient même être tentées de remettre en cause leur participation au financement du système public. Sans être aussi alarmiste, on peut toutefois noter que le développement de l'assurance volontaire qui remplit la fonction 4 constitue un marqueur de la qualité et des limites d'un système de santé public. Il est intéressant d'observer que la médecine de première ligne au Royaume-Uni reste dans le domaine public, alors que l'offre privée concerne les soins aigus que le système hospitalier public, confronté à de grandes difficultés financières et organisationnelles, ne prend pas en charge de façon satisfaisante. En Finlande, l'assurance privée concerne à l'inverse l'ambulatoire.

CONCLUSION

Au terme de ce chapitre, un certain nombre de conclusions se dégagent.

Tout d'abord, malgré l'idée couramment admise selon laquelle les réformes récentes ont conduit à une "privatisation" des systèmes d'assurance publique, l'assurance volontaire reste fondamentalement complémentaire de la couverture publique. Sa part dans le financement des dépenses de santé et le rôle qu'elle va jouer dépendent de l'architecture du secteur public et du service qu'il propose, dans ses différentes dimensions (comme l'indique le tableau 8).

Dans tous les cas, l'assurance volontaire reste marginale dans le financement des dépenses de santé, même si elle répond à des besoins spécifiques. Il apparaît en outre clairement que l'assurance, au sens où nous l'entendions dans la première partie, n'est que rarement mise en œuvre dans les faits et que les Etats-Unis constituent la seule véritable exception. Chacun sait qu'ils ont cherché, dans les années quatre-vingt-dix, à s'éloigner de ce modèle.

Ce travail nous a amené à vérifier que le revenu est une décision centrale dans la décision d'assurance. Dans le même temps, nous avons affiné l'analyse du chapitre six en soulignant que l'accès à l'assurance volontaire ne conditionne pas seulement l'accès aux soins quand elle constitue la seule source de couverture possible. L'accès à l'assurance volontaire peut avoir un impact sur l'accès aux soins et la structure de la consommation lorsque les copaiements sont élevés ou lorsque certains biens ne sont pas couverts (les services ambulatoires). En conséquence, le rôle de l'assurance volontaire a un impact sur l'équité globale du système de santé. Ceci avait été qui avait été mis en lumière par l'étude européenne ECuity (van Doorslaer et Wagstaff, 1993) pour les quelques pays sur lesquelles on disposait d'informations suffisamment complètes pour mesurer la redistributivité de l'assurance volontaire. La présente analyse complète ces résultats d'un point de vue plus qualitatif.

Tableau 8 : Synthèse générale : le rôle de l'assurance volontaire

		Fonction 1	Fonction 2	Fonction 3	Fonction 4
Etats-Unis	Population concernée	Moins de 65 ans		Plus de 65 ans	
	Proportion	88%		12%	
	Taux de souscription	<ul style="list-style-type: none"> • 69% assurance privée • 12% Medicaid/Medicare • 19% non couvert 		62%	
Pays-Bas	Population concernée	Personnes au dessus d'un niveau de revenu	Personnes au dessous d'un niveau de revenu		
	Proportion	30%	63%		
	Taux de souscription	<ul style="list-style-type: none"> • 85% assurance privée • 15% contrat standard 	90%		
Allemagne	Population concernée	Personnes au dessus d'un seuil de revenu choisissant le privé	Bénéficiaires du régime public	Fonctionnaires	
	Proportion	9% (sur 20% potentiels)	89%	2%	
	Taux de souscription	<ul style="list-style-type: none"> • 99,99% assurance privée • 0⁺% contrat standard • 0⁺% ne se couvre pas 	10%	100%	
Belgique	Population concernée		Régime général	Travailleurs indépendants	
	Proportion		85%	15%	
	Taux de souscription		90%	65%	
Suisse	Population entière		35%		
Israël	Population entière		n.d.		
France	Population entière			85%	
Danemark	Population entière			26%	
UK	Population entière				12%
Espagne	Population entière				15-20%
Italie	Population entière				20%
Grèce	Population entière				10%
Finlande	Population entière				12%
Autriche	Population entière				35%
Irlande	Population entière				40%

CHAPITRE 8

UNE CONCURRENCE ENTRE ASSUREURS POUR L'EFFICACITE PRODUCTIVE

Dans un très grand nombre de pays, l'assurance privée volontaire joue donc un rôle relativement marginal. Cette affirmation semble paradoxale quand on sait que de grandes réformes des systèmes de santé des années quatre-vingt dix ont porté sur l'assurance maladie et ont visé à introduire des mécanismes de marché dans ce secteur.

Comme nous l'avons souligné au terme de la seconde partie, ces réformes avaient pour principal objectif d'obtenir que les organismes concernés concentrent leurs efforts sur la gestion des risques et la recherche de l'efficacité productive en limitant la sélection différentielle. Autour de cette problématique commune, nous avons établi une distinction entre les systèmes de tradition libérale ou sociale.

Quels sont les principes généraux qui caractérisent les évolutions récentes dans ces deux types de système ?

1. Lorsque le marché de l'assurance est traditionnellement concurrentiel, les entités qui couvrent les risques prélèvent directement des primes d'assurance, plus ou moins ajustées au risque, auprès des assurés. L'organisme régulateur estime le niveau de risque des personnes qui sont effectivement couvertes par chaque assureur. Il met alors en place un mécanisme de refinancement croisé entre les assureurs qui couvrent des personnes présentant des risques relativement élevés et les autres. Que la structure de risque de l'assureur résulte du hasard ou de stratégies plus ou moins efficaces de sélection, il peut bénéficier de transferts compensatoires ou être au contraire obligé de financer les malchanceux. Les primes qu'il propose, dès lors, prennent en compte cet ajustement et ne reflète plus sa structure de risque mais son efficacité productive relative.
2. Lorsque le régulateur introduit une forme de concurrence ou de responsabilité financière des organismes gérant le risque maladie, on part d'une situation initiale dans laquelle des entités publiques sont financées au prorata de leurs dépenses. Les

ressources destinées à financer le système sont prélevées par l'Etat auprès des ménages bénéficiaires et les contributions sont généralement liées au revenu. Afin d'inciter les organismes à être plus efficaces, l'Etat met progressivement en œuvre un financement prospectif normatif, tout en maintenant la solidarité du financement. Il collecte les fonds auprès des ménages et alloue aux entités assurant la gestion de la couverture des capitations ajustées au risque.

Si les problématiques dans ces deux types de systèmes convergent, on constate aussi que les voies empruntées pour obtenir le résultat souhaité, mis à part le mode de prélèvement auprès des ménages, sont comparables : dans les deux cas, l'objectif est d'obtenir que les entités qui assurent la gestion des risques, reçoivent un paiement normatif ajusté au risque. Dans le premier cas, le régulateur met en place un ajustement des risques ex post ; dans le second, il verse aux assureurs une capitation et procède à un ajustement de cette capitation en fonction des risques ex ante¹. Parallèlement, on maintient (ou on introduit) la possibilité pour les assurés de choisir un opérateur parmi ceux qui lui sont proposés. Ceci renforce théoriquement les incitations des opérateurs à rendre un service coût-efficace et à améliorer la qualité des services rendus². L'objet de ce chapitre est d'approfondir l'étude technique de ce mode de gestion de la concurrence.

Au préalable notons que les entités qui prennent part à cette concurrence fortement encadrée peuvent être des assureurs privés, des caisses publiques, voire des groupes de prestataires de soins. Lorsqu'il n'est pas utile de les distinguer, nous les qualifions de façon générique d'Organismes Gestionnaires du Risque (OGR), terminologie qui témoigne du fait que l'efficacité productive est au cœur du questionnement.

Nous avons précédemment signalé que la sélection différentielle en assurance n'a pas été étudiée de façon formalisée dans la littérature économique. De même, les analyses théoriques des outils nécessaires pour limiter la sélection en stimulant la recherche de l'efficacité productive dans le cadre de l'assurance maladie sont rares. Pour autant, une littérature assez riche s'est développée depuis plusieurs années sur ce sujet. Elle s'inspire généralement de l'observation d'un système précis.

¹ Afin de lever une ambiguïté possible, rappelons que la notion de "paiement prospectif" fait référence à une situation dans laquelle un paiement normatif est déterminé en début de période. On oppose à cette notion celle de paiement rétrospectif qui qualifie une situation dans laquelle on rembourse les coûts de production lorsque le service a été produit. Dans les deux cas présentés ici, le paiement est prospectif normatif, mais les ajustements peuvent être réalisés ex ante ou ex post.

² Le concept d'efficacité productive rend compte simultanément des dimensions coûts et qualité.

Dans la première partie de ce travail, nous dégagons les éléments de convergence de ces travaux sur les conditions nécessaires à l'établissement d'une concurrence qui porte véritablement sur l'efficacité productive. Cette approche vise dans le même temps à rendre compte des contraintes techniques que le régulateur (qui n'est pas nécessairement l'Etat) doit surmonter, des arbitrages qu'il doit réaliser et des questions qui restent ouvertes (Section 1). Puis, dans un second temps, nous étudions plus précisément les expériences des différents pays en déterminant si les conditions nécessaires que nous avons identifiées sont remplies. Ce travail de recensement, qui se veut exhaustif, permet dans le même temps de tirer un premier bilan des réformes en cours et de montrer que l'organisation du financement et de l'activité des OGR selon les principes édictés n'est pas une condition suffisante pour atteindre l'objectif souhaité (Section 2).

1. REGLEMENTATION DE LA CONCURRENCE : ENJEUX ET PRINCIPES

La responsabilisation financière incite théoriquement les OGR en concurrence à améliorer leur efficacité productive. Nous avons toutefois vu que la sélection est la conséquence directe du versement d'un paiement prospectif mal ajusté au risque. Il n'est dès lors pas étonnant que l'organisation du mode de financement des OGR soit au cœur des solutions proposées. L'essentiel de la régulation passe par l'élaboration d'un mécanisme tel que les entités en concurrence ne soient pas tentées de sélectionner les assurés (1.1). Toutefois, pour parvenir à une solution satisfaisante, la concurrence doit se dérouler dans un cadre particulier et nous dégagons les conditions supplémentaires auxquelles les OGR et les assurés doivent se plier (1.2).

1.1 L'ajustement au risque

La solution naturelle pour réduire la sélection, du point de vue économique, consiste à mettre en place un mécanisme de financement tel que l'OGR ne puisse réaliser, à efficacité productive constante, un profit plus élevé s'il couvre une catégorie de risque plutôt qu'une autre. Cette condition est remplie si le financement obtenu pour chaque personne est actuariellement juste. Dans ce cas, des OGR en concurrence, qui sont incités à augmenter leur efficacité productive, pourraient même avoir intérêt à se

spécialiser dans l'offre de service pour les hauts risques, notamment pour bénéficier d'économies d'échelle³(Bowen, 1995, Kronick *et al.*, 1995).

Nous nous sommes attachés à détailler, dans la première partie de ce travail, les raisons pour lesquelles les niveaux de risque que présentent les individus diffèrent. Nous avons souligné que, pour pouvoir produire le service de couverture des risques, un assureur doit proposer à chaque assuré une prime actuarielle. Plus précisément, il différencie théoriquement les primes selon les facteurs de risque qui contribuent à expliquer la variance explicable. Par la suite, nous avons montré, d'une part, que l'assureur n'est pas toujours en mesure d'avoir accès à l'information dont l'individu dispose sur son propre niveau de risque, et d'autre part, qu'il pouvait choisir de pratiquer une tarification grossière des risques. Nous nous interrogeons désormais explicitement sur la faisabilité technique, pour le régulateur, d'un ajustement au risque, qu'il soit réalisé ex ante ou ex post. Au préalable, il convient de choisir un critère de jugement qui permette de déterminer la qualité d'un tel ajustement par rapport à une situation théorique idéale (1.1.1). Puis, nous nous demandons si, et par quels moyens, il est possible de mettre en œuvre un tel ajustement (1.1.2). Nous montrons enfin que l'on peut mettre en place des mécanismes qui complètent le dispositif et pallient ses défauts (1.1.3).

1.1.1 Critère d'évaluation de la qualité d'un ajustement au risque

Dans un premier temps, nous présentons le critère à l'aune duquel la qualité d'un ajustement au risque quelconque se mesure. Ceci nécessite d'une part, de caractériser d'une situation d'ajustement parfait et d'autre part, de trouver un critère de comparaison d'un mode d'ajustement particulier avec cette situation idéale.

Dans un premier temps, il faut déterminer quelle proportion de la variance est explicable, autrement dit quel est le R^2 maximal qui peut être prédit⁴. De nombreuses études proposent des estimations de la part de la variance explicable, mais les résultats diffèrent selon les échantillons, même lorsque les auteurs utilisent des spécifications économétriques similaires. Pour van Vliet et van de Ven (1992), elle est environ de 14%. Newhouse estime pour sa part qu'elle représente entre 20 et 25% de la variance totale (Newhouse, 1996, Newhouse *et al.*, 1989). Wolfe et Goddeeris (1991), dont l'objectif

³ Rentabilisation d'investissements spécialisés, d'un investissement dans la formation de personnels soignants...

⁴ cf Chapitre 3, § 3.1.1.

n'est pas de déterminer le R^2 maximum, obtiennent un R^2 de 22% (tableau 4 pp.448). On peut noter de plus que le R^2 maximal dépend du type de dépense que l'on observe : il est beaucoup plus élevé pour les dépenses ambulatoires (environ 50%, Newhouse *et al.* 1989) que pour les dépenses hospitalières (8%, *ibid.*).

Il n'existe donc pas d'étalon unique et les études qui visent à estimer la qualité d'un ajustement au risque choisissent en général, de façon assez arbitraire, un seuil d'environ 20%.

Dans un second temps, on estime la variance expliquée par les facteurs de risques utilisés par l'assureur ou le régulateur (le R^2 empirique) et on la compare au R^2 maximal choisi. Plus la part de la variance expliquée dans la variance explicable augmente, meilleur est l'ajustement au risque.

La pertinence d'un tel critère de jugement reste partielle, même s'il est couramment utilisé. Newhouse (1996) souligne ainsi qu'en ajoutant un facteur de risque, on peut augmenter la part totale de la variance expliquée mais diminuer la part de la variance expliquée dans certains sous-groupes. A titre d'illustration, on peut citer l'étude de Howland *et al.* (1987) : les mêmes facteurs de risques (mesures physiologiques et utilisation passée de ressources) expliquent 10% de la variance des hospitalisations observées dans les deux années suivantes pour les hommes et 3% pour les femmes.

L'existence d'un tel phénomène soulève plusieurs réflexions. D'une part, l'introduction d'un facteur de risque qui explique mieux la variance des dépenses d'un sous-groupe que d'un autre peut générer une incitation à une sélection différentielle, en faveur du groupe dont la variance est mieux rendue au détriment de l'autre. En prolongeant le raisonnement, on peut se demander s'il ne serait pas intéressant d'utiliser des facteurs de risque différents pour des populations disjointes. Luft (1995) estime que le programme de recherche ne devrait pas porter sur l'étude de la variance des dépenses individuelles dans un échantillon représentatif de la population mais sur les dépenses de groupes vulnérables caractérisés par des facteurs de risques particuliers⁵. De telles approches compliqueraient les mécanismes d'estimation et d'allocation, mais il serait intéressant de mener une analyse comparative en utilisant les différentes méthodes sur un même

⁵ "Concern about the poor explanatory power of models focusing on individual expenditures is misplaced if the real question is how well one can predict for groups, particularly groups reflecting non random risk selection" (Luft, 1995, pp.28).

échantillon, afin de déterminer celle qui rend le mieux compte de la variation des dépenses entre individus.

A défaut, pour juger de la véritable qualité d'un ajustement au risque, il ne faudrait pas se fonder sur la seule observation du R^2 . Il est préférable de tester la robustesse de l'ajustement en comparant les dépenses prédites aux valeurs observées sur des sous-populations spécifiques ou, comme le suggèrent van de Ven et van Vliet (1993), sur plusieurs intervalles de dépenses.

1.1.2 La tarification du risque en pratique et ses limites

Muni d'un critère de jugement, le régulateur doit choisir les critères qu'il utilise effectivement pour procéder à l'ajustement des risques. Cet exercice nécessite de prendre en compte un certain nombre de contraintes techniques (a). Une recherche très prolifique s'attache à déterminer les facteurs de risque qui devraient être pris en compte (b). Compte tenu de ces éléments, nous proposons un bilan provisoire de l'état de la recherche et des perspectives (c).

a) Les contraintes techniques de l'ajustement au risque

Que la tarification idéale du risque soit mise en œuvre par l'assureur directement, par un organisme chargé de la compensation des risques ex post, ou par une institution chargée de déterminer et de verser aux OGR une capitation normative, le choix des critères d'ajustement doit respecter un certain nombre de contraintes techniques (van de Ven et van Vliet, 1992). Ces contraintes sont plus ou moins fortes selon le contexte.

1. La valeur du critère sur laquelle est fondé le calcul doit pouvoir être estimée sans erreur de mesure.
2. Chaque critère doit rendre compte d'informations suffisamment structurelles pour que le paiement reflète une tendance valable sur l'ensemble de la période et ne caractérise pas seulement l'individu au moment où la mesure est effectuée.
3. Le coût d'extraction de l'information ne doit pas être trop élevé et, en règle générale, l'information doit être relativement facile à obtenir.
4. Les critères choisis ou les modes de collecte de l'information doivent respecter la vie privée des individus. Cette contrainte est plus ou moins forte selon le contexte et la qualité des personnes qui collectent l'information. Si la personne remplit un

questionnaire que l'employeur transmettra à l'assureur, il sera plus réticent à révéler certaines informations que s'il s'adresse à un médecin tenu de respecter le secret médical qui transmettra une information statistique agrégée décrivant les risques couverts par chaque OGR.

5. L'information ne doit pas être manipulable. Dans le cas où l'assuré paie une prime qui est fonction de son risque, il n'a pas intérêt à révéler à l'assureur des informations qui prouvent qu'il est un haut risque. Mais la manipulation peut aussi être le fait de l'OGR : dans la mesure où sa dotation est fonction des risques qu'il couvre, il a intérêt à surévaluer ces risques afin de recevoir un montant plus élevé. Les circuits d'information doivent être élaborés pour minimiser les possibilités de manipulation et l'institution qui détermine les paiements ou les ajustements, doit se ménager la possibilité de vérifier l'information.
6. Les critères ne doivent pas créer d'incitations perverses, notamment d'incitations à une moindre efficacité pour les OGR.
7. Enfin, il faut que chaque "cellule" définie par le croisement des critères comprenne un nombre suffisant d'individus pour pouvoir évaluer le paiement, sans commettre d'erreur de mesure. Nous avons souligné que c'était là une des raisons pour lesquelles l'assureur, qui n'a accès qu'à l'information contenue dans ses propres fichiers, pouvait chercher à éviter de couvrir certaines personnes. Un organisme qui centralise les informations émanant d'un grand nombre d'OGR peut avoir un avantage à cet égard. Mais un tel système centralisé comporte aussi des limites : il faut que chaque OGR soit en mesure de restituer périodiquement une grande masse d'information détaillée et standardisée sur la base de laquelle l'organisme central estimera la capitation ou l'ajustement. En outre, dans l'idéal, l'information restituée doit aussi permettre de distinguer les OGR qui produisent de façon peu efficace de ceux qui sont victimes d'une sélection défavorable (Swartz, 1995). Or, les systèmes d'information existant ne permettent pas toujours de respecter ces conditions. Aux Etats-Unis, ainsi il n'est pas rare que les plans passent eux même des contrats avec des groupes médicaux sur une base de capitation (ce qui est, après tout, un moyen de les inciter à réaliser un effort de gestion optimal) ou qu'ils salarient les prestataires de soins. Dans ce cas, les plans

n'ont pas nécessairement accès à une information détaillée de nature médicale (sur l'état de santé, les pathologies...) et ne sont pas en mesure de la restituer⁶.

b) Quelles variables utiliser dans l'ajustement au risque ?

Compte tenu de ces contraintes, nous présentons les principaux résultats de la recherche empirique, qui vise à déterminer les facteurs de risque les plus pertinents pour l'ajustement au risque.

Variables socio-démographiques

Le recours à la consommation médicale varie tout d'abord en fonction de l'âge et du sexe. Ces facteurs sont quasiment toujours pris en compte pour estimer les risques : les dépenses de santé sont plus élevées au début et à la fin de la vie et les dépenses des femmes sont plus élevées que celles des hommes pendant leur période de fécondité. Ces deux données sont donc des facteurs de base dans la détermination des dépenses de santé.

On sait par ailleurs que le niveau et le mode de recours aux soins varient en fonction de l'environnement économique et social des individus. On peut donc théoriquement envisager de prendre en compte un certain nombre de variables comme le niveau de revenu, le niveau d'étude, la profession (qui peut aussi refléter des facteurs de risque particuliers), le fait d'être employé, de vivre seul... En général, ces variables sont prises en compte sous forme d'indicateurs, comme le fait d'être au chômage ou de bénéficier de certaines prestations sociales réservées aux personnes défavorisées. Les Britanniques, dans le cadre de la réflexion sur l'allocation macro-économique des ressources aux régions⁷ ont tenté d'élaborer un indice synthétique de caractéristiques socio-économiques, mais cette réflexion n'a, semble-t-il, jamais été étendue au calcul de l'ajustement des risques dans le cadre plus micro-économique de l'allocation à des OGR. L'inconvénient majeur de certaines de ces variables est leur caractère transitoire (revenu, chômage) et endogène (l'individu peut les influencer).

Des variables caractérisant l'environnement peuvent aussi permettre de rendre compte de différences dans le coût de la prise en charge des patients. Ainsi, le degré d'urbanisation

⁶ Ce manque d'information des plans sur les patients et les soins effectivement administrés soulève des difficultés dans le cadre de l'estimation d'indicateurs de qualité.

⁷ Procédure RAWP : *Risk Allocation Working Party*

peut refléter des variations de prix des infrastructures ou de mode de recours liées à la densité et à la nature de l'offre. Toutefois, on hésite parfois à prendre en compte ce type de variable par crainte de figer des situations peu efficaces. Par exemple, si la densité plus élevée de praticiens en ville encourage des comportements d'induction de la demande, les différences de dépenses observées entre milieu rural et urbain ne sont pas justifiées du point de vue de l'efficacité. La question plus générale posée ici est de savoir si l'ajustement au risque ne doit prendre en compte que des variations d'état de santé, ou s'il doit prendre en compte des facteurs déterminants des coûts comme les prix et l'offre de soins (van de Ven *et al.*, 1995). Ceci dit, les variables géographiques (code postal) peuvent aussi permettre de rendre compte indirectement de différences de type socio-économique. Par exemple, Kronick *et al.* (1995,51) montrent que le nombre d'admissions et de journées d'hôpital est beaucoup plus élevé quand la proportion de personnes vivant sous le seuil de pauvreté dans une zone géographique augmente. Le choix d'inclure ou non des variables de type géographique est donc assez délicat.

L'ensemble de ces caractéristiques socio-démographiques (âge, sexe, région, degré d'urbanisation, revenu, chômage, niveau d'éducation, taille de la famille...) dont on peut imaginer qu'elles sont disponibles de manière routinière ou qu'elles peuvent être aisément collectées en respectant la vie privée des individus, ne permettent pourtant d'expliquer qu'une faible proportion de la variance totale (autour de 3%), soit entre 10 et 20% de la variance explicable (Newhouse, 1996, van de Ven et van Vliet, 1992).

Indicateurs de l'état de santé

L'ajustement cherche avant tout à rendre compte du niveau de risque des individus. Or, ce risque est largement lié à l'état de santé général. Une recherche foisonnante s'est développée ces dernières années, surtout aux Pays-Bas et aux Etats-Unis, pour élaborer des critères d'ajustement qui reflètent l'état de santé. Nous ne détaillons par ici les résultats⁸ mais nous rendons compte des principales pistes de recherche et des enjeux qui président au choix des différentes variables.

- L'état de santé perçu (physique ou psychologique), ainsi que la mobilité, sont des informations en général obtenues dans des enquêtes, sur une base déclarative. Elles ont une bonne capacité prédictive mais elles restent en partie subjectives et peuvent

⁸ Lamers (1997) van de Ven et al. (1992) van de Ven et al. (1995) proposent des revues de littérature empirique détaillées.

aisément être manipulées. Au delà des problèmes techniques de collecte de l'information que leur utilisation poserait, il semble difficilement envisageable de les inclure dans une procédure d'ajustement au risque.

- Les mesures physiologiques (pression sanguine, cholestérol...) permettent d'expliquer une partie non négligeable de la variance explicable mais elles ont un coût de collecte très élevé et leur prise en compte nécessite d'avoir accès à des informations complémentaires qui permettraient de les interpréter : un niveau de cholestérol faible peut en effet résulter d'un traitement approprié (Newhouse, 1996). Or, une personne qui présente un taux de cholestérol faible parce qu'elle suit un traitement régulier consomme plus de ressources qu'une personne qui a "naturellement" un taux identique.
- Les variables de style de vie (obésité, consommation de tabac et d'alcool) semblent par contre améliorer faiblement le pouvoir prédictif des modèles (à court terme et au niveau individuel).
- La mortalité est corrélée avec les dépenses de santé qui sont souvent très élevées dans les mois qui précèdent un décès (Beck et Zweifel, 1995). Outre le fait que l'ajustement en fonction d'une telle variable est nécessairement réalisé ex post, la logique selon laquelle on devrait augmenter la rémunération d'un OGR quand le nombre de personnes décédées augmente semble contestable. De fait, les taux de mortalité sont parfois utilisés dans les processus d'allocation macro-économique (au Royaume-Uni) mais jamais dans des calculs d'allocations individuelles.
- L'utilisation passée de ressources médicales est un bon estimateur des dépenses futures. Il est même le meilleur lorsqu'il est utilisé seul. Van Vliet et van de Ven (1992) montrent par exemple que les 10% de la population qui ont les dépenses annuelles les plus élevées une année, auront pendant (au moins) les 4 années suivantes des dépenses en moyenne deux fois plus élevées que les personnes qui appartiennent aux mêmes groupes âge-sexe⁹. Toutefois, ce critère d'ajustement est critiquable à plusieurs égards. Tout d'abord, l'utilisation passée reflète l'ensemble des déterminants de la demande (y compris éventuellement l'offre). En lui accordant une importance trop grande, on peut pérenniser ou encourager des comportements de recours et des

⁹ On trouve un résultat inverse pour les personnes qui ont les dépenses les plus faibles. L'observation des dépenses passées est un moyen peu coûteux de sélection différentielle pour les OGR.

pratiques médicales inefficaces ou inflationnistes. En fait, on décourage les efforts de maîtrise des coûts puisque, plus les dépenses augmentent, plus la rémunération prospective des OGR est élevée (Jones, 1990). Ensuite, la capacité prédictive de l'utilisation passée dépend de la nature du besoin de santé. En effet, une consommation élevée une année peut s'expliquer par un problème aigu ou chronique. Dans le premier cas, la consommation une année n'est pas un aussi bon estimateur des dépenses subséquentes que dans le second.

- Les informations de nature diagnostique semblent pouvoir échapper à ces critiques et les travaux récents explorent cette piste de façon systématique. Des variables indicatrices de la présence de certaines pathologies diagnostiquées lors d'hospitalisations une année ou détectées au travers de la consommation de certains médicaments permettent d'améliorer la valeur prédictive des modèles pour les années qui suivent. On peut comptabiliser le nombre de pathologies chroniques et les pondérer par une variable de coût relatif (Kronick *et al.*, 1995, van Vliet et van de Ven, 1992). La recherche s'oriente vers l'élaboration de groupes de diagnostics qui donnent lieu à des prises en charge d'intensité comparable. Ash *et al.* (1989) ont développé le modèle de référence pour l'hôpital¹⁰ et des modèles ont aussi été construits pour des pathologies détectées en ambulatoire¹¹ (Lamers, 1997, pp. 29 et s., Newhouse et Chairman, 1998, pp. 75).

La principale critique adressée à ces modèles est qu'ils reposent sur des informations manipulables par les OGR ou les prestataires de soins. Les chercheurs tentent donc de classer les diagnostics en fonction du pouvoir discrétionnaire dont les médecins disposent pour les diagnostiquer et d'ajuster la pondération en fonction de ce critère.

Il reste cependant une difficulté à surmonter (Newhouse, 1994, pp. 141) : si les pratiques de traitement diffèrent radicalement entre par exemple les HMOs et des assureurs traditionnels, sur quelle référence doit-on se baser pour estimer le coût moyen de la prise en charge d'une pathologie ? Le moins élevé, au risque d'ériger une sous-estimation en norme ? La moyenne ? Ou doit-on estimer le coût moyen de la méthode de traitement qui fait l'objet du consensus scientifique le plus large ? Cette dernière méthode semble objectivement la plus juste, mais elle nécessite de faire appel

¹⁰ *Diagnostic Cost Groups DCGs.*

¹¹ *Ambulatory Cost Groups ACGs.*

à des compétences très spécifiques et complique plus encore le calcul des facteurs d'ajustement.

- L'invalidité, l'incapacité ou le handicap fonctionnel expliquent des variations interindividuelles significatives dans les dépenses de santé, y compris quand d'autres variables au pouvoir explicatif élevé, comme l'utilisation passée de ressources sont prises en compte. Ces informations sont, dans certains cas, recueillies pour des raisons administratives (versement de pensions) et leur utilisation semble être consensuelle (van de Ven et van Vliet, 1992).

En conclusion, signalons qu'à l'heure actuelle, l'utilisation des informations individuelles sur l'état de santé permet, au mieux, d'expliquer environ les deux tiers de la variance explicable sur des échantillons représentatifs de la population totale (van de Ven et van Vliet, 1992). L'utilisation de facteurs de risque multiples communs à des sous-populations d'assurés ne permet en effet jamais de tarifer la composante strictement individuelle du risque et donc d'intégrer la totalité de la variance explicable dans l'évaluation prospective du risque.

c) Bilan et perspectives

Le moyen théorique idéal d'éradication de la sélection, qui consisterait à garantir que les OGR soient parfaitement rémunérés en fonction des risques qu'ils assument et que cette rémunération les incite à maîtriser les coûts, semble donc un objectif irréaliste dans l'état actuel de la recherche. L'opinion de Luft qui qualifie la locution "actuarialement juste" d'oxymore (1995, pp.24) nous semble cependant par trop pessimiste. On peut, en premier lieu, raisonnablement estimer que les progrès de la science permettront à moyen terme de mieux prévoir le risque. Ensuite, de façon très pragmatique, si on instaure une capitation imparfaitement ajustée au risque, les OGR qui sont desservies par ce système ont intérêt à collaborer à l'effort de recherche pour apporter la preuve du fait qu'elles sont moins bien placées que d'autres au regard des critères utilisés, voire pour en proposer de nouveaux.

Il faut en réalité se demander si l'amélioration du processus d'ajustement au risque sera, à terme, suffisante pour éliminer les profits espérés d'une stratégie d'écrémage. L'OGR réalise en effet un arbitrage : il faut que le profit de la stratégie de sélection soit supérieur à son coût, coût de mise en œuvre de la sélection elle-même et coût de leur réputation.

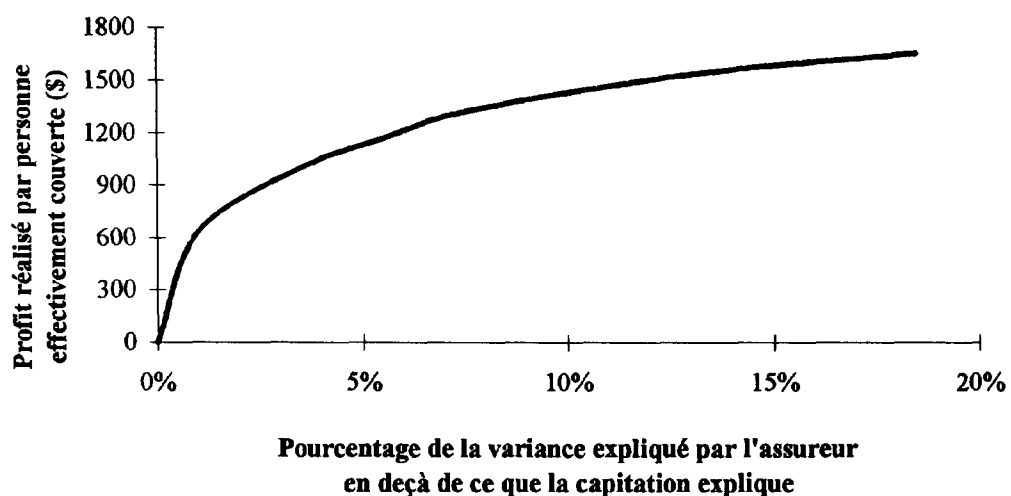
Plusieurs auteurs ont cherché à estimer le profit que les assureurs, ou les OGR pouvaient tirer de la sélection, mais leurs résultats sont divergents.

Newhouse *et al.* (1989) montrent que le profit de l'écrémage n'est pas une fonction linéaire de l'information privée que l'OGR détient sur le risque. Les auteurs considèrent une situation dans laquelle l'Etat verse à l'OGR une capitation ajustée au risque. Ils calculent le profit de la sélection en fonction du différentiel d'information dont bénéficie l'OGR. Ils se fondent sur les hypothèses comportementales suivantes :

1. l'OGR calcule les dépenses anticipées par assuré sur la base de l'information qu'il détient ;
2. si l'information qu'il détient n'est pas prise en compte dans la capitation ajustée au risque qu'il reçoit de l'Etat, il est donc en mesure de déterminer quelles sont les personnes pour lesquelles la capitation est sur ou sous-estimée¹² ;
3. il évince les personnes pour lesquelles il sait que la capitation est sous-estimée (cette stratégie est supposée de coût nul). Le profit qu'il réalise par personne effectivement couverte est alors calculé.

Comme le montre la figure 1, le profit de l'OGR croît avec la part de la variance des dépenses qu'il explique, au delà de la part prise en compte par la capitation ajustée au risque (fonction croissante). De plus, ce taux de croissance est décroissant (fonction concave).

Figure 1 : Profit de la sélection
D'après Newhouse et al. 1989, tableau 4, pp. 50



¹² On se place ici dans le cas où l'OGR reçoit une capitation ajustée au risque déterminée par une entité publique. La situation serait la même si l'assureur déterminait lui-même une prime fondée sur des critères simples (âge sexe) et utilisait l'information supplémentaire dont il dispose pour la sélection. En particulier, il pourrait atteindre le même niveau de profit.

Si l'OGR parvient à expliquer 1% de plus de la variance totale que la capitation ajustée au risque, il réalise un profit substantiel (alors que si la courbe de profit était convexe, il lui faudrait réaliser une estimation bien meilleure pour atteindre ce résultat).

Newhouse extrapole le résultat en affirmant qu'il sera toujours possible à l'OGR de réaliser un profit en améliorant de façon très marginale la capitation.

Cette analyse a été reprise par van de Ven, van Vliet, van Barnevelt et Lamers (1994b, 1995) qui démontrent tout d'abord (sur la base des mêmes chiffres) que Newhouse surestime le profit réalisé sur le premier pour-cent de variance expliquée "en plus" d'environ 60% lorsque la variance expliquée par la capitation est plus élevée¹³. Autrement dit, si la capitation ajustée au risque versée à l'OGR prend un plus grand nombre de facteurs de risque en compte, le profit que l'OGR réalise en l'améliorant de 1% est beaucoup plus faible. En outre, l'analyse de Newhouse ne tient pas compte de plusieurs éléments :

- d'une part, lorsque la variance expliquée par la capitation augmente :
 - l'information dont doit disposer l'OGR pour faire "mieux" est de plus en plus sophistiquée et coûteuse à acquérir,
 - de plus, van de Ven *et al.* (1994b) démontrent que la variance du profit estimé augmente. Autrement dit, lorsque la capitation ajustée au risque est améliorée, le profit réalisé en faisant 1% de mieux est de plus en plus aléatoire et de moins en moins élevé ;
- d'autre part, Newhouse fait implicitement l'hypothèse que seul un OGR sélectionne (parfaitement et sans coût) et que les autres ne réagissent pas, ce qui le conduit à surestimer le profit de la stratégie. En réalité, une stratégie de sélection ne fonctionne que dans la mesure où elle est mise en œuvre par un nombre limité d'OGR, voire par un seul OGR. Si plusieurs OGR mettent en œuvre des stratégies similaires pour sélectionner les risques, ils devront en partager le bénéfice alors que chacun devra assumer le coût de sa stratégie.

¹³ La démonstration est détaillée dans la note technique à laquelle van de Ven fait référence (1994).

En conclusion, il semble que l'ajustement au risque reste à l'heure actuelle nécessairement imparfait et laisse une marge de manoeuvre à l'OGR pour sélectionner les risques. Toutefois, il est permis d'espérer qu'à terme, une meilleure connaissance des déterminants de l'état de santé et de la consommation médicale, permettront une tarification des risques plus fine. A cet égard, il nous semble utile de réfuter l'idée selon laquelle une tarification au risque parfaite gomme toute incitation à l'efficacité. Certains auteurs, notamment Newhouse *et al.* (1989) laissent entendre que si l'intégralité de la variance est expliquée, on retombe dans une situation identique au remboursement des coûts (paiement rétrospectif). Même si on fait abstraction du fait qu'il est impossible de prévoir la variance pure, cet argument reste discutable. En effet, tout paiement déterminé de façon prospective est incitatif dans la mesure où l'OGR peut bénéficier d'une diminution des coûts et en tirer un profit. Si au contraire il ne maîtrise pas les dépenses, il devra assumer un éventuel déficit. Dès lors que le paiement est normatif, il est intrinsèquement incitatif.

Quoi qu'il en soit, à l'heure actuelle, si les techniques d'ajustement au risque ne sont pas parfaites, celles de la sélection ne le sont probablement pas non plus et le profit de la sélection n'est pas nécessairement élevé et certain. Enfin, et c'est l'aspect que nous étudions maintenant, il est possible d'élaborer des modes de financement mixtes qui contribuent à mitiger les effets de la sélection.

1.1.3 Les systèmes de paiement mixtes

Dans le début de cette analyse, nous avons implicitement étudié le cas où l'ajustement au risque, qu'il soit réalisé ex post ou ex ante, porte sur l'intégralité de la dépense. Le calcul de la somme que doit finalement recevoir l'OGR se fonde sur une estimation qui ne prend pas explicitement et directement en compte les dépenses auxquelles il fait face.

Or, l'analyse théorique oppose deux cas extrêmes :

- le paiement rétrospectif : en remboursant en fin de période les dépenses de l'OGR, on ne génère aucune incitation à l'efficacité ;
- le paiement prospectif ou normatif : en versant à l'OGR un paiement prospectif normatif, on crée une incitation maximale à l'efficacité productive, mais aussi à la sélection dans la mesure où la dotation n'est pas parfaitement ajustée aux risques.

Si on sait qu'il est impossible d'ajuster parfaitement le paiement au risque, on peut envisager, pour limiter la sélection, de mettre en place un mode de paiement en partie prospectif et en partie rétrospectif. Newhouse (1996) souligne qu'une solution intermédiaire, à savoir un système de paiement mixte, est probablement optimale du point de vue théorique.

Un certain nombre de raisons pragmatiques plaident aussi (selon les contextes) pour un mode de paiement mixte :

- Newhouse (1994) estime qu'un système mixte est toujours préférable à un système pur. En effet, lorsqu'on rembourse les dépenses observées, on fonde le paiement sur une estimation biaisée des coûts qui ne prend pas en compte l'efficacité ; lorsqu'on calcule un paiement prospectif, on risque de commettre des erreurs dans l'estimation du coût des traitements. Il en tire la conclusion que "*Pricing errors are inevitable...a mixed system is better than a pure system because it averages these errors*"¹⁴ (pp. 143) ;
- ensuite, selon le contexte institutionnel dans lequel évoluent les OGR, ils peuvent être amenés à supporter des coûts plus ou moins fixes et déterminés de façon exogènes. Dans la mesure où ils n'ont pas les moyens d'exercer un réel contrôle sur ces coûts, il ne semble pas logique de les responsabiliser financièrement sur la totalité de la dépense. Aux Pays-Bas, par exemple, toutes les caisses sont obligées de passer des contrats avec tous les hôpitaux et n'ont aucune marge de manoeuvre pour négocier les tarifs. Les dépenses hospitalières des caisses sont donc remboursées intégralement et ne font pas partie de l'ensemble des dépenses sur lesquelles elles sont responsabilisées ;
- enfin, dans les systèmes où le paiement prospectif est destiné à augmenter l'efficacité des organismes qui servent l'assurance maladie (comme les Pays-Bas ou la Belgique), on part d'une situation antérieure où les caisses n'avaient aucune incitation financière à gérer efficacement et recevaient des paiements entièrement rétrospectifs. Pour éviter de les déstabiliser et quand bien même on chercherait à tendre vers un paiement largement prospectif, il faut introduire la responsabilité financière progressivement.

¹⁴ Les erreurs de tarification sont inévitables, un système de paiement mixte est meilleur qu'un système pur parce qu'il permet de répartir, en moyenne, ces erreurs.

S'il on admet qu'un mode de rémunération mixte est optimal, il faut s'interroger sur la façon dont il doit être calibré. Newhouse (1996) propose une revue de littérature complète des modèles dont les conclusions pourraient éclairer ce choix¹⁵, mais il conclut que les solutions théoriques proposées ne sont pas adaptées au cas de l'assurance et que la détermination du partage optimal nécessitera de nombreuses expérimentations. Parmi les pistes de recherches envisagées :

- Newhouse (1994, pp. 142) propose d'augmenter la part du financement rétrospectif (remboursement) quand les dépenses sont très élevées afin de protéger les personnes plus vulnérables à la sélection ;
- Bowen (1995) suggère qu'on pourrait exclure certaines pathologies du paiement prospectif ;
- Luft (1995, pp.30) propose de rembourser les OGR pour la part des dépenses qui dépasse un certain plafond ou de constituer un pool de hauts risques vers lequel les OGR transféreraient ceux qu'ils estiment être des hauts risques et qui serait financé collectivement pas les OGR et / ou l'Etat.

Ce type de solution (notamment les deux dernières) a pour inconvénient majeur de déresponsabiliser assez largement les OGR sur la maîtrise des dépenses lorsque celles-ci sont très élevées. Ce faisant, ils ne sont pas incités à collecter l'information médico-économique sur les personnes concernées. Or, cette information permet d'étudier la façon dont les patients sont pris en charge, de contrôler les pratiques afin de les rendre plus efficaces et, à terme, elle est essentielle à la recherche de l'efficacité.

Ceci dit, aucun de ces auteurs n'explore en détail le fonctionnement de tels systèmes. Van Barneveld *et al.* (1996) constituent une exception à cet égard. Ils proposent un mécanisme original et en développent une analyse systématique sur des données empiriques. Le principe du mécanisme est le suivant : les auteurs partent d'un système où la capitation ajustée au risque (imparfaite) concerne l'ensemble de la population et proposent, pour en pallier les défauts, d'instaurer une mutualisation obligatoire des hauts risques qui fonctionne selon les principes exposés ci-après.

¹⁵ Entre autres, Laffont et Tirole (1993), Ma (1994) qui étudie le cas de l'hôpital, Ellis et Mc Guire (1990) pour les médecins.

1. L'OGR désigne périodiquement un groupe d'individus dont les dépenses vont être prises en charge collectivement par l'ensemble des OGR.
2. La proportion d'assurés dont les coûts seront mutualisés est fixée par les autorités pour chaque OGR.
3. Chacun d'entre eux est libre de déterminer ses critères de sélection et peut, en particulier, exploiter les informations dont il dispose sur ses propres assurés. A titre individuel, l'OGR a donc intérêt à déterminer qui va engendrer une dépense au dessus de la capitation avec autant de précision que possible. Ainsi, si les dépenses passées ne sont pas prises en compte dans la capitation ajustée au risque, l'OGR peut les utiliser pour déterminer qui doit être dans le pool. Les personnes dont l'OGR se doute au départ qu'elles vont entraîner des dépenses supérieures à la capitation ajustée au risque, sont donc intégrées dans ce pool.
4. La stratégie de sélection de l'OGR est détournée et n'aura pas de conséquences néfastes pour les assurés : puisque l'OGR ne supporte le coût que dans la mesure où il participe au financement du pool, il a un intérêt moindre à évincer les hauts risques. Un autre avantage de ce système est qu'il n'oblige pas le régulateur à déterminer des critères d'admission au pool qui pourraient être contestés ou devraient être négociés (comme c'est le cas si on choisit d'exclure certaines pathologies du système de paiement prospectif).
5. En fin d'année, les OGR partagent les dépenses du groupe poolé. La contribution d'un OGR ne dépend que des dépenses de pool et du nombre total de personnes qu'il couvre. Autrement dit, elle ne dépend pas directement des dépenses que ses propres assurés ont générées.

Ce mécanisme ne concerne pas spécifiquement les personnes qui ont des dépenses élevées. En effet, même si une personne génère ex post un déficit que l'OGR ne pouvait pas rationnellement anticiper, le risque d'écrémage disparaît puisque l'OGR n'était pas capable de détecter ce "haut risque". La fonction d'assurance est maintenue, au sens où l'OGR est responsable des variations imprévisibles des coûts autour de la moyenne, mais la tentation de recours à l'éviction des hauts risques a été limitée.

Concrètement, pour les personnes poolées, la mutualisation peut s'appliquer au coût total, au surcoût par rapport à la capitation reçue ou aux dépenses supportées par l'OGR au delà d'un plafond arbitraire. En outre, elle peut être partielle : l'OGR ne mutualise alors pas l'ensemble de la dépense mais une fraction plus ou moins élevée. Cette solution semble préférable pour ne pas générer de déresponsabilisation collective sur la maîtrise des coûts engendrés par les personnes concernées.

Enfin, on peut noter que ce mécanisme est neutre vis-à-vis des assurés eux-mêmes qui, typiquement, ne savent pas que leurs dépenses ont été poolées.

Une simulation de ce mécanisme est proposée dans l'article cité ci-dessus sur la base de données d'assurance collectées pour deux années consécutives (1992 et 1993).

Dans un premier temps, une capitation ajustée au risque (grossière) est calculée en fonction de l'âge et du sexe. Les auteurs font l'hypothèse que chaque OGR va intégrer, en 1993, dans le pool, les personnes qui, en 1992, ont eu les dépenses les plus élevées et que l'intégralité des dépenses qu'elles engendrent sera prise en charge par le pool.

On classe en ordre décroissant les différences entre les dépenses effectives de 1993 et les capitations, puis on regroupe les observations en classes qui représentent chacune 1% de la population totale.

Selon que le pool concerne 1 à 4% des assurés, les auteurs calculent :

- le pourcentage de la dépense totale poolée. Plus celui-ci est élevé, moins l'incitation de l'OGR à maîtriser les coûts est importante. Ce pourcentage est donc un des critères qui permet de déterminer (de façon pragmatique) le pourcentage de la population qui devrait être poolé ;
- la prime à payer par personne couverte pour financer le pool.

On étudie dans quelle mesure le profit / la perte prévisible par sous groupe varie, autrement dit, dans quelle mesure les incitations à écrémer diminuent.

Si 4% des personnes sont poolées, la dépense qui reste à la charge des OGR représente 75% de la dépense totale. Sur les 25% des dépenses poolées, la capitation âge sexe représente 6% de la dépense totale. On voit donc que le critère de choix utilisé par les OGR est simple et efficace (de leur point de vue). Cette impression est confirmée par le test suivant : lorsqu'on partitionne la population que l'OGR conserve à sa charge selon

différents critères - ce qu'il ferait s'il cherchait à trouver de nouveaux critères lui permettant de déterminer quelle catégorie non poolée il faudrait, soit chercher à détourner vers d'autres OGR, soit faire entrer dans le pool dans la période suivante - aucun profil de risque particulier ne se dégage.

Autrement dit, lorsque l'OGR s'est déchargé des 4% des personnes qui ont coûté le plus cher l'année précédente, les informations disponibles ne lui permettent plus de distinguer précisément ex ante des catégories de personnes qui génèrent une perte très élevée. Ceci démontre que le critère simple choisi est le bon.

Ce mécanisme incite-t-il réellement l'OGR à ne plus maîtriser des coûts pour les personnes poolées ? Notons d'abord que la proportion de personnes concernées par le mécanisme est faible : il est, dès lors, vraisemblable que les mesures mises en œuvre par l'OGR pour gérer les soins et augmenter son efficacité productive sur les 96% des personnes non poolées pourront aussi s'appliquer aux 4% poolées. L'argument selon lequel les soins des personnes dont les dépenses sont les plus élevées ne sont pas gérés doit donc être relativisé, en particulier si le pool ne prend en charge qu'une partie de leur dépense et non l'intégralité comme c'est le cas dans cette simulation.

Une dernière série de tests permet de vérifier la robustesse du mécanisme proposé. Sur la base des données disponibles, on génère 11 OGR fictifs auxquels on assigne par simulation des structures de risque délibérément différentes. Initialement, la différence entre le résultat moyen par assuré le plus élevé (profit) et le plus faible (perte) est de 500 florins (pour un coût moyen sur la population de 1700 florins). Si chacun des OGR poolé 4% de sa population, la différence est réduite à 340. Dans une dernière étape, on module le pourcentage poolé par plan en fonction de la perte nette par assuré enregistrée l'année précédente¹⁶. Ceci se justifie parce que la capitation ajustée au risque est très grossière et qu'une perte moyenne très élevée ne peut être entièrement imputée à une mauvaise gestion de l'OGR. En utilisant ce critère, la différence des profits tombe à 265 environ. Au total, la perte effectivement réalisée sur les 1% de personnes dont on avait anticipé, sur la base d'un critère simple, qu'elles engendreraient des coûts largement supérieurs à la capitation, diminue de 80%. Autrement dit, ce mécanisme n'exonère pas l'OGR de toute sa responsabilité dans la maîtrise des coûts, ce d'autant moins que des personnes non

¹⁶ Pour la détermination du pourcentage poolé par plan, le régulateur utilise une information assez simple (cf. van Barneveld, et al., 1996, pp. 141.)

poolées peuvent générer des pertes aussi élevées, sans que cela ait été prévisible : toute mesure destinée à mieux gérer les soins s'appliquera aussi à elles.

Cette analyse détaillée confirme que, même si l'ajustement au risque est imparfait, il est possible d'élaborer des mécanismes originaux qui permettent de contrer les stratégies d'écrémage des OGR. Celui qui est proposé ici repose sur une approche pragmatique et empirique. Les hypothèses sur l'information disponible et sur les comportements des différents agents semblent aussi réalistes.

Au terme de cette analyse, nous sommes en accord avec la position de van de Ven et al. : *"Although we are optimistic about the technical possibilities to prevent cream skimming, the realization in practice requires considerable effort"*¹⁷ (1995, pp.18).

Nous décrirons, dans la seconde partie de ce chapitre, les expériences en cours d'ajustement des risques qui témoignent de la diversité des mécanismes possibles, et qui confirment que cette recherche est prometteuse.

Toutefois, si l'ajustement au risque est une condition nécessaire, il faut que les assureurs et les assurés acceptent de se soumettre à d'autres contraintes.

1.2 Les contraintes supplémentaires

La détermination d'un mode de financement, à lui seul, ne permet pas d'enrayer les comportements de sélection. On observe une certaine convergence dans l'énoncé des autres conditions qui doivent être réunies pour obtenir que la concurrence fonctionne selon les critères souhaités. Nous les présentons ici en soulignant, quand cela sera nécessaire, en quoi les enjeux diffèrent entre les systèmes, initialement concurrentiels, dans lesquels on cherche à limiter les effets des stratégies de sélection, et les systèmes dans lesquels on introduit une forme de concurrence ou de responsabilité financière en cherchant à éviter la sélection.

¹⁷ *optimistes quant aux possibilités techniques d'éradiquer l'écrémage, ..(même si)..la réalisation pratique de cet objectif requiert un effort considérable.*

1.2.1 Définition du panier offert

Dans les deux modèles étudiés, la définition d'un panier standard apparaît comme une nécessité. Les enjeux d'une telle définition sont cependant différents selon le contexte.

Notons tout d'abord qu'imposer une standardisation des garanties constitue en soi un moyen réglementaire pour limiter la sélection des risques (Newhouse, 1994). Nous avons en effet vu qu'en profilant son offre, l'OGR ou l'assureur était en mesure d'attirer des personnes qui présentent un niveau de risque plus ou moins élevé. La standardisation des garanties limite donc cette sélection volontaire ou involontaire par les OGR.

Mais la standardisation est aussi considérée, pour des raisons techniques, comme une condition nécessaire au fonctionnement de l'ajustement au risque (Bowen, 1995, Luft, 1995, Swartz, 1995) même si certains auteurs, comme Bowen, estiment qu'il serait envisageable d'offrir un nombre limité de contrats différents. En effet, pour procéder à l'ajustement des risques, il faut pouvoir distinguer, dans les prix (ou les coûts) d'un OGR, ce qui relève de l'efficacité productive des effets de sélection. Si les contrats proposés par les plans diffèrent par les copaiements ou les niveaux de franchise, il faut pouvoir, en outre, estimer ce qui, dans la différence de prix, relève :

1. du fait que les sommes déboursées par l'OGR sont plus ou moins élevées ;
2. des conséquences sur les frais de gestion de la structure du contrat (les frais de gestion sont moins élevés toutes choses égales par ailleurs en cas de franchise puisque le nombre de dommages remboursés diminue) ;
3. de l'impact des différences de taux de couverture sur le risque moral.

En imposant la standardisation des contrats on peut espérer que le calcul de l'ajustement sera plus juste et - relativement - plus simple.

Les nombreux auteurs qui se penchent sur cette question de l'ajustement au risque s'accordent à dire qu'elle est une nécessité pour promouvoir l'efficacité productive et qu'elle doit s'accompagner d'une standardisation des bénéfices. Ils mettent en avant l'argument selon lequel cette standardisation représente un avantage pour le consommateur qui voit la transparence du marché augmenter et qui peut faire des choix rationnels. En revanche, ils passent sous silence le fait que, par rapport au modèle idéal

de concurrence, elle constitue une contrainte forte sur les choix des individus et se traduit, du point de vue théorique, par une restriction de l'efficacité allocative.

Les pays dans lesquels le contrôle de la sélection accompagne l'introduction d'une dose de concurrence partent en général d'une situation antérieure où toutes les personnes couvertes bénéficient des mêmes garanties. La définition d'un panier standard est une opération qui semble pouvoir soulever moins de difficultés que si l'on part d'un contexte initial où la liberté d'établir son panier couvert est totale. Ceci dit, la définition d'un panier standard nécessite de répondre à plusieurs questions.

1. Dans un premier temps, il faut définir le domaine sur lequel porte la concurrence. On peut, en particulier, exclure certains risques ou certaines populations vulnérables du champ soumis à concurrence. C'est l'option qui a été choisie aux Pays-Bas. Tous les résidents bénéficient de l'AWBZ¹⁸. La réforme inspirée des propositions de Dekker prévoyait d'intégrer l'AWBZ dans le champ concurrentiel mais cette option a finalement été rejetée. Deux arguments essentiels justifient une telle décision : tout d'abord, la qualité est un enjeu beaucoup plus grand pour les soins de long terme et les personnes qui en bénéficient sont probablement, à la fois plus vulnérables à cet égard, et moins à même de se défendre si un OGR ne garantit pas un niveau de qualité suffisant. Ensuite, les études économétriques sur données longitudinales (van Barneveld *et al.*, 1997) ont montré qu'il est difficile de déterminer une capitation ajustée au risque qui gomme toute incitation à la sélection différentielle des patients qui relèvent de ce régime.
2. De façon plus générale, déterminer le panier de biens et services sur lesquels la responsabilité de la collectivité est engagée nécessite de faire des choix (politiques) explicites. Le rapport Santé 2010 souligne que le panier couvert à l'heure actuel en France est très large et que la contrepartie est un taux de remboursement moyen moins élevé que dans un grand nombre de pays. "*Le resserrement de l'éventail pourra conduire à mieux couvrir certains biens au détriment d'autres services dont l'utilité médicale sera jugée moins importante (les cures thermales par exemple)*" (1993, pp.403).

¹⁸ Couverture socialisée et obligatoire des risques médicaux graves (rappel)

3. Enfin, il faut définir les processus de révision du panier au fur et à mesure que des techniques ou des produits nouveaux apparaissent.

La mise en œuvre des modes de financement requis pour limiter la sélection nécessite donc de standardiser les contrats d'assurance et de définir explicitement les biens et services couverts. Dans un contexte au départ concurrentiel, cela signifie que l'efficacité allocative doit être limitée afin d'obtenir l'efficacité productive. Dans un contexte où chacun bénéficie initialement d'une couverture identique, cette obligation peut permettre de redéfinir les priorités, ce qui est un avantage mais qui nécessite de trouver un consensus entre différents acteurs du système (prestataires de services, pouvoirs publics et assurés). Une telle démarche implique des arbitrages politiques souvent délicats. Ainsi aux Pays-Bas, le ministère a tenté d'éliminer les prothèses dentaires du panier standard et a été contraint l'année suivante de les réintroduire¹⁹.

1.2.2 Détermination du périmètre et autres conditions annexes

La seconde condition nécessaire au fonctionnement de l'ajustement au risque concerne la définition du périmètre de cet ajustement. Pour que l'ajustement ait un sens, il faut en effet imposer à un groupe de personnes et aux entités qui les couvrent de participer à l'ajustement au risque.

a) Obligation d'assurance ?

En effet, les assurés qui présentent de faibles niveaux de risque ont intérêt à obtenir la tarification la plus juste de leur risque. De leur point de vue, il est préférable que l'OGR qui les couvre sélectionne les risques. Si l'on admet la nécessité de juguler la sélection des risques afin de promouvoir l'efficacité productive, il faut élaborer un mécanisme de financement dans lequel les signaux prix (s'il y en a) rendent compte de différentiels d'efficacité productive et non de la capacité des OGR à sélectionner. Les modalités exactes de financement détermineront toutefois l'impact final sur les bas risques d'une telle mesure et la nécessité ou non de les contraindre à s'assurer. Si on part d'une situation antérieure où les OGR ne différencient pas spontanément les primes en fonction de critères de risque fins, et si on continue à imposer une forme de *community rating*,

¹⁹ En Israël, le processus de définition et de révision du panier a donné lieu à de nombreux débats. Le sous-titre d'un article qui traite du système israélien est instructif : "*explicit rationing of health care services in a universal system is easier said than done*" (Chinitz et Israeli, 1997).

comme c'est souvent le cas dans les mécanismes de péréquation élaborés aux USA, l'impact pour les bas risques n'est pas nécessairement élevé et peut être compensé à terme par une amélioration de la concurrence qui porte sur la gestion des processus de soins, la prévention... La plupart des systèmes qui émergent aux Etats-Unis s'organisent sur une base largement volontaire, à l'initiative des employeurs ou des centrales d'achat que nous avons évoquées dans le troisième chapitre. Les nombreux avantages que présentent ces formes d'assurance semblent donc, à l'heure actuelle, l'emporter sur les inconvénients pour les bas risques. Toutefois, si le coût net devient trop élevé pour certaines catégories de personnes, par rapport à ce qu'elles pourraient espérer payer dans un autre système que celui qui leur impose une forme de péréquation, et si elles ont la possibilité d'échapper au système, ou de ne pas s'assurer, le mécanisme d'ajustement ne peut fonctionner. On peut alors imposer une obligation d'assurance. Ceci est particulièrement vrai si on veut mettre en œuvre ou maintenir une participation des individus qui n'est pas fonction des risques mais, par exemple, du revenu.

Dans tous les cas, la couverture au sein du groupe concerné, qu'il soit constitué sur une base volontaire ou non, doit être virtuellement universelle (Bowen, 1995).

b) Obligation de participation et contraintes imposées aux OGR

Dans la mesure où les services de plusieurs OGR sont offerts aux membres du groupe d'assurés constitué²⁰, il faut évidemment que tous les OGR participent au mécanisme d'ajustement au risque et que tous soient soumis aux mêmes exigences vis-à-vis de cet ajustement des risques.

Au total, comme le note Bowen (1995, pp.34), l'ajustement du risque doit s'inscrire dans un marché défini. Ce marché peut couvrir tous les individus et tous les OGR, ou il peut être défini de façon beaucoup plus limitative (tous les employés d'une entreprise et tous les OGR qui leur proposent des contrats). En tout état de cause, l'entrée et la sortie du marché doivent être réglementés. Cette contrainte se fonde aussi sur des raisons d'ordre technique : pour que le mécanisme d'ajustement fonctionne, il faut que la durée de tous les contrats soit standardisée et que tous soient souscrits à la même période. Si ces conditions ne sont pas réunies, l'ajustement au risque est compliqué par le fait que la structure de risque de chaque OGR peut varier à tout moment et qu'il doit porter sur une

²⁰ Si les assurés n'ont pas le choix entre plusieurs prestataires, la question de la sélection ne se pose pas.

période différente pour chaque OGR. La durée des contrats est donc en général fixée à un an (mais prévoit des exceptions en cas d'entrée ou sortie pour des raisons justifiées, déménagement notamment). Pour ce qui est du renouvellement, on permet en général aux assurés de manifester leur choix pendant une période d'un mois (dite *d'open enrollment*), au terme de laquelle les anciens contrats arrivent à échéance et les nouveaux prennent effet²¹.

Enfin, pour que l'ajustement au risque prenne tout son sens, les OGR doivent respecter quelques obligations supplémentaires :

- couvrir toute personne du groupe constitué qui en fait la demande (*guaranteed issue*) ;
- renouveler par la suite les contrats des personnes qui le souhaitent (*guaranteed renewal*) ;
- ne pas exclure de la garantie des maladies avérées au moment où la personne entre dans le contrat (*pre-existing conditions*)²².

Ces contraintes réglementaires sont, comme la standardisation des contrats, couramment utilisées pour limiter la sélection par les assureurs, y compris lorsqu'il n'y a pas d'ajustement au risque (Kronick *et al.*, 1995, Newhouse, 1994). Toutefois, elles ne permettent pas à elles seules d'enrayer la sélection et elles restent nécessaires au fonctionnement de l'ajustement au risque.

Malgré les nombreuses contraintes qui leur sont imposées, nous verrons, au travers d'exemples précis de mise en œuvre de mécanismes d'ajustement au risque, que les OGR trouvent un intérêt commercial à offrir des contrats, qui contrebalance les contraintes qui leurs sont imposées et auxquelles ils acceptent en général de se soumettre.

Nous avons, dans cette première partie, présenté les enjeux et les conditions techniques de mise en œuvre d'une gestion de la concurrence efficace. L'élément principal du dispositif est le mode de financement conçu pour éviter que les OGR ne recourent à la sélection. L'objectif ultime est d'obtenir un financement prospectif véritablement actuariel

²¹ Cette tradition américaine, qui simplifie la gestion, avait émergé avant que la question de l'ajustement des risques ne se pose.

²² Il faut toutefois éviter les comportements trop opportunistes des assurés : s'ils ne sont pas obligés de souscrire un contrat et s'ils savent que l'assureur ne peut refuser de les couvrir, ils peuvent attendre pour souscrire l'assurance d'être malades. Luft (1995) suggère dans le cas où l'assurance n'est pas obligatoire, que les maladies avérées puissent être exclues pendant la première année d'assurance (délai de carence), mais qu'elle soient couvertes par la suite, y compris si l'assuré change de plan (il faut alors que l'information soit transmise).

des OGR. Plutôt qu'élaborer une parade réglementaire à chaque modalité de sélection, le régulateur cherche à éliminer les possibilités de réaliser les profits de court terme générés par de telles stratégies.

2. PRATIQUE DE L'AJUSTEMENT AU RISQUE ET DE LA CONCURRENCE TEMPEREE

La seconde partie de ce travail vise à recenser les systèmes dans lesquels un mode gestion de la sélection prévaut, qui s'apparente à celui que nous venons de décrire. Les pays développés de l'hémisphère nord, concernés par ce type de réforme, sont les Etats-Unis, Israël, la Belgique, l'Allemagne, les Pays-Bas, l'Irlande et la Suisse.

En décrivant ces systèmes et leurs évolutions ou réformes récentes, nous verrons en quoi les conditions que nous avons identifiées comme nécessaires au fonctionnement d'une concurrence limitant la sélection sont remplies ou non.

Nous nous attacherons aussi à décrire le fonctionnement concret des différentes méthodes d'ajustement au risque, dont nous n'avons pour l'instant fait qu'esquisser les principes²³. Pour finir, nous illustrerons par cette démarche l'idée selon laquelle l'arbitrage entre les exigences de maîtrise des coûts et la nécessité de lutter contre la sélection, induit une convergence dans la structure des systèmes d'assurance maladie à priori très éloignés. En fait, les deux modèles que nous avons opposés de "limitation de la sélection dans un univers concurrentiel" (modèle américain) et "d'introduction de la concurrence" (faisant plutôt référence à l'expérience des caisses publiques aux Pays-Bas), sont des exemples illustratifs de deux groupes de systèmes que nous étudions successivement :

- les systèmes de type concurrentiel dans lesquels on introduit l'ajustement des risques (2.1) ;
- les systèmes dans lesquels un organisme central va déléguer la gestion des risques à des opérateurs en concurrence (2.2).

²³ Nous renverrons le cas échéant, pour le détail de la présentation, à des encadrés techniques ou à des sous paragraphes clairement identifiés.

2.1 Introduction de l'ajustement en complément de la concurrence

La première série de cas que nous présentons regroupe un ensemble de pays dans lesquels l'assurance est traditionnellement souscrite sur une base volontaire et plutôt concurrentielle. Elle est financée par des primes dont s'acquittent les individus et qui reflètent peu ou prou leur niveau de risque. L'ajustement des risques est progressivement apparu comme une condition nécessaire au bon fonctionnement du marché. L'exemple le plus emblématique est celui des Etats-Unis. L'Irlande, dans un contexte initialement monopolistique, cherche maintenant à promouvoir la concurrence sur le marché de l'assurance volontaire et a mis en œuvre un mécanisme d'ajustement destiné à éviter les effets jugés néfastes de la sélection. Enfin, la Suisse, qui a introduit l'assurance universelle en 1994, tout en maintenant une organisation concurrentielle, a, dans le même temps, mis en place un ajustement des risques.

2.1.1 Les Etats-Unis

Dans le système américain, ces dernières années, la sélection est devenue un enjeu de plus en plus important sur l'ensemble des marchés d'assurance volontaire (a). Des formes d'organisation émergent sur les marchés privés qui apportent une réponse locale et limitée aux problèmes posés par la sélection dans le cadre de la *managed competition* (b).

a) La généralisation du problème de la sélection

Aux Etats-Unis, les marchés sur lesquels la sélection est la plus flagrante sont les marchés individuels et celui des petites entreprises. Une des raisons pour lesquelles ces marchés sont particulièrement vulnérables est que les assurés ne sont pas groupés par d'autres liens que ceux de l'assurance (Gauthier *et al.*, 1995). Lorsque les risques sont élevés, les personnes concernées ne trouvent pas d'assureur qui accepte de les couvrir ou les contrats ne sont pas abordables. Il n'existe pas de mécanisme structurel qui permette de grouper les risques comme dans les grandes entreprises.

Par contraste, jusqu'à une période assez récente, le marché des grandes entreprises était relativement épargné car les risques y sont plus stables et plus homogènes. Les bénéficiaires de l'assurance de groupe ont un emploi stable à plein temps (les autres catégories d'employés sont rarement couvertes), ce qui constitue un premier élément

d'homogénéisation des risques, même s'il ne s'applique pas à l'ensemble des ayants droit (autres membres du ménage). La péréquation entre niveaux de risque est donc plus aisée, parce que les personnes s'assurent de façon solidaire et non à titre individuel. Ceci dit, la sélection est devenue un enjeu au fur et à mesure que les employeurs ont multiplié les choix d'assurance des individus (Jones, 1989, Jones, 1990, Luft, 1995). Même si la couverture reste quasiment universelle au sein des grandes entreprises, dès lors que plusieurs contrats sont offerts, les conditions du marché individuel sont partiellement recréées et nous avons vu que la segmentation spontanée compromettrait le fonctionnement de ces systèmes (cas de Harvard, chapitre 5).

Aux Etats-Unis, sur l'ensemble des marchés de l'assurance volontaire, on cherche à restreindre la capacité des plans à segmenter les risques afin de les obliger à se concurrencer sur les prix, le service et la qualité.

b) L'émergence de la "managed competition"

L'assurance volontaire est en général souscrite à titre individuel dans une gamme de contrats sélectionnés, négociés et partiellement financés par les gros employeurs (privés ou publics). Les Etats encouragent la création de centrales d'achats qui jouent auprès des PME un rôle similaire.

Selon la terminologie proposée par Enthoven (1988, 1994), les employeurs et les centrales d'achat doivent jouer le rôle d'acheteur avisé ou de "sponsor" dans la "concurrence régulée" (*managed competition*). Ce modèle idéal stipule que les consommateurs sont sensibilisés au prix à payer, qu'ils peuvent exercer un réel choix qui incite les assureurs à offrir des contrats (et éventuellement des services) efficaces. Le "sponsor" est un intermédiaire entre assureurs et assurés qui doit, pour mener à bien sa tâche :

- présélectionner un nombre limité de plans en amont (il est en mesure de se doter d'outils plus performants que les assurés pour juger de leur qualité) ;
- prendre des mesures pour que les salariés puissent faire leur choix sur la base de critères simples, transparents, et en nombre limité. Pour ce faire, il peut stimuler ou exiger la standardisation des contrats et imposer des conditions de tarification - en général de type *community rating* - qui sont uniformes ;
- imposer aux plans d'accepter tout membre du groupe (*guaranteed issue et renewal*).

En outre, le sponsor, par définition, effectue une péréquation des risques au sein d'un groupe relativement moins homogène que s'il était constitué sur un marché d'assurance individuelle. Il centralise les demandes et obtient ainsi un pouvoir de marché qui lui permet de négocier les prix, notamment parce qu'il diminue les frais de gestion et de marketing des plans (Diamond, 1994). Pour garantir que la concurrence se déroule sur les bons critères, il faut enfin que les sponsors mettent en place une compensation des risques entre les plans participants.

L'expérience montre que les employeurs ne jouent pas leur rôle de sponsor. Seuls 43% des salariés ont le choix entre plusieurs plans (NEHIS 1993²⁴). Hibbard, Jewett *et al.* (1997) montrent que la plupart des employeurs ignore les indicateurs de résultats médicaux sur les hôpitaux proposés par les plans. En outre, seuls 31% des employeurs diffusent une information aux salariés (dite *performance information*), et encore, cette information privilégie les performances en termes de services (rapidité à gérer les remboursements) et très rarement sur les performances cliniques.

Dans ce contexte, les sponsors ayant mis en place une compensation des risques restent encore assez rares. Un total de cinq exemples est recensé dans la littérature. Nous avons choisi, lorsque l'information est disponible²⁵, d'en présenter le fonctionnement de façon assez détaillée, pour témoigner de la diversité des solutions possibles.

le Health Insurance Plan of California (HIPC)

Le HIPC est une centrale d'achat californienne, destinée aux PME et gérée par le MRMIB²⁶ dont le mécanisme d'ajustement, mis en place en 1996, est décrit et analysé par Shewry *et al.* (1996) et Bertko (1998). Le mécanisme a été élaboré en concertation par le MRMIB, les plans qui participaient à la centrale d'achat et des consultants.

Afin d'en comprendre le fonctionnement, il faut noter au préalable que, pour participer au HIPC, les plans doivent se soumettre à plusieurs obligations (*open enrollement period, guaranteed issue et renewal*). De plus, ils doivent utiliser un *community rating* ajusté sur la base de l'âge, de la localisation, et de la structure de la famille (selon quatre catégories qui ne prennent pas en compte explicitement le nombre d'enfants). Ces facteurs de risque

²⁴ National Employer's Health Insurance Survey

²⁵ Le système de calPers, entité publique qui joue le rôle de sponsor pour les employés de l'Etat de Californie depuis 1962 et celui que l'Etat de New-York a mis en place sur le marché des PME ne seront ainsi pas présentés.

²⁶ Managed Risk Medical Insurance Board

sont donc inclus dans la prime facturée aux assurés et le mécanisme d'ajustement doit porter sur les facteurs non tarifés. Autrement dit, le mécanisme doit gommer, pour un individu résidant dans un endroit particulier et ayant un âge précis, ce qui, dans la différence de prix d'un contrat à l'autre, relève du fait que l'ensemble des personnes couvertes par chaque assureur sont, par exemple, en plus mauvaise santé.

Après concertation, les parties concernées ont choisi d'ajuster les risques en fonction de trois critères non tarifés : le nombre d'enfants couverts, le sexe, et certaines pathologies coûteuses. On calcule pour chaque plan un indice de risque relatif (RAV). S'il est trop éloigné de la moyenne, égale à 1, le plan est taxé ou subventionné.

Encadré 1 Ajustement au risque dans le MRMIB

Pour chaque plan, le MRMIB calcule un *RAV* (*risk assessment value*), fondé sur ces trois composantes, qui reflète la structure de risque relative du plan par rapport à celle de la population totale couverte. Le choix des diagnostics thérapeutiques pris en compte a été effectué sur la base de données de quatre plans représentant 7,2 millions d'années-personnes, et 120 diagnostics ont été sélectionnés sur la base de deux critères :

- ils entraînent des consommations supérieures à la moyenne ;
- leur dénombrement n'est pas trop dépendant des pratiques de codage des plans ou des praticiens. Autrement dit, les possibilités de manipuler l'information pour exhiber une structure de risque plus défavorable qu'elle ne l'est réellement, afin d'obtenir une compensation plus grande, sont limitées. Les chiffres déclarés par les plans sont, en outre, soumis à un audit externe.

Pour finir, seuls des diagnostics entraînant hospitalisation ont été retenus, car les plans n'ont pas d'information suffisamment précise en ambulatoire. Les principales indications retenues sont : SIDA, cancer du sein, maladies pulmonaires chroniques, maladies du foie, maladies ischémiques, arythmies, troubles du pancréas, syndrome de détresse respiratoire.

Une fois la proportion de personnes par diagnostic connue, celle-ci est multipliée par le ratio (coût moyen des soins des personnes atteintes / coût moyen des soins dans la base). On alloue aux individus qui n'ont aucun des diagnostics recensés un poids tel que la moyenne des poids sur la population vaut 1. Ce système revient à allouer aux personnes en bonne santé une dotation déterminée "par défaut" quand on a mis de côté des sommes pour les autres. On pourrait aussi envisager d'allouer aux personnes en bonne santé un poids qui reflète leur coût relatif par rapport aux autres. Cette solution adoptée par exemple aux Pays-Bas (pour d'autres facteurs de risque), semble plus logique sur le plan économique, mais elle favorise moins explicitement les personnes qui génèrent des coûts élevés (ici les malades).

Enfin, on recalculé pour chaque plan en fonction du nombre de personnes atteintes de chaque diagnostic un poids moyen. De façon similaire, on estime un poids relatif de chaque plan qui rend compte de la répartition des sexes et du nombre d'enfants couverts. En multipliant les poids et en normant à 1, on obtient le *RAV* de chaque plan. Il a été convenu de ne mettre en œuvre l'ajustement que pour les plans dits *outliers* (ceux qui sortent de la norme ou qui sont assez éloignés de la moyenne) et dont le *RAV* est éloigné de plus ou moins 5% du *RAV* total qui est par construction égal à 1.

Reste alors à estimer la somme qu'il faudrait transférer ou prélever sur les *outliers* pour les ramener aux bornes de 0.95 et 1.05. Pour ce faire, on calcule une prime moyenne théorique par plan, qui prenne en compte les facteurs qu'il est en droit de tarifer afin que le transfert ne porte que sur les facteurs non tarifables. On calcule un *rated risk factor (RRF)*, qui reflète, pour chaque plan, sa structure âge / région / type de famille, par rapport à celle de l'ensemble de la population couverte. On postule alors que la prime moyenne théorique qu'un plan devrait facturer si sa structure de risque était identique à celle de la population totale est égale à la prime moyenne sur tous les plans (P_M), multipliée par son *RRF*. Compte tenu de sa structure de risque non tarifable, la prime qu'il doit (théoriquement) faire payer pour parvenir à l'équilibre est égale à :

$$RAV_i \times P_M \times RRF_i$$

Si son *RAV* est supérieur à 1.05, la subvention qu'il va recevoir par personne couverte est :

$$(RAV_i - 1.05) \times P_M \times RRF_i$$

Si au contraire son *RAV* est inférieur à la borne limite, il va payer par personne couverte :

$$(0.95 - RAV_i) \times P_M \times RRF_i$$

Des mécanismes sont prévus pour garantir que le fond de péréquation sera finalement à l'équilibre : si on estime par exemple que les sommes prélevées sont insuffisantes pour couvrir les subventions, on prélève des sommes sur les plans dont le *RAV* est légèrement supérieur à 0.95.

En dernier lieu, il reste à expliquer la temporalité de ce mécanisme : l'ajustement calculé une année est en fait mis en œuvre l'année suivante. Par exemple, en t on calcule que le plan i devrait recevoir (ou payer) 10\$ par mois et par personne sur la base des informations collectées en $t-1$. Lorsque la période de renouvellement des contrats arrive, le plan propose pour $t+1$ une prime qui prend en compte ce transfert. Il peut choisir de répercuter intégralement ou non ce transfert dans la prime proposée mais il paiera ou recevra effectivement via le HIPC par personne couverte et par mois la somme arrêtée l'année précédente, transfert compris. Ce mécanisme d'ajustement au risque fonctionne donc de façon comparable à l'*experience rating* : on récupère ou répercute sur l'année suivante les sur- ou sous-dotations de l'année précédente. Pour que le mécanisme fonctionne, il faut donc que les taux de *turn over* ne soient pas trop élevés, même si des ajustements supplémentaires sont effectués en fin d'année ($t+1$). Il n'en reste pas moins que la gestion globale du système est complexe et que l'ajustement au risque est effectué sur la base d'informations collectées deux années auparavant, ce qui complique en particulier la prise en compte de l'entrée ou de la sortie de plans.

Le bilan global de l'expérience après deux années complètes d'existence est positif. La taille de la population totale couverte dans le HIPC a considérablement augmenté et l'ensemble des acteurs en présence est satisfait (Bertko et Hunt, 1998). Quelques défections initiales de plans ont été constatées mais d'autres plans (de taille importante) ont rejoint le HIPC, et le nombre total de plans offerts est resté stable, entre 20 et 25. Les résultats de l'ajustement montrent que les HMO ont en général des structures de risque meilleures (7 sur 22 ont un *RAV* inférieur à 0.95 la première année), et qu'un des deux PPOs offert a dû être compensé assez fortement la première année (\$46 par contrat et par mois). On peut voir dans cet état de fait la conséquence d'un effet d'auto-sélection spontané, qui traduit encore une fois que les plus malades préfèrent les formes plus souples de prise en charge. Au total, la première année, 1% des primes ont été transférées. La seconde année, les transferts totaux ont représenté 0,15% des primes et la

gamme des reversements s'est considérablement rétrécie. Quatre HMO ont reversé des sommes et un PPO a été subventionné. Un des objectifs recherchés, à savoir maintenir un choix aux assurés entre HMO et PPO, a donc été atteint. Les premiers sont plus souvent mis à contribution et les seconds continuent à exiger des primes relativement élevées car l'ajustement est incomplet, mais tous y gagnent en termes de parts de marché et acceptent de jouer un jeu dont le fonctionnement a été élaboré sur la base d'une concertation des acteurs en présence. Tous participent aussi à la recherche pour améliorer le fonctionnement du mécanisme d'ajustement.

le *Buyers Health Care Action Group* (BHCAG)

Le BHCAG du Minnesota regroupe les employés de 28 grosses entreprises et négocie directement l'achat de soins à des groupes médicaux et médico-hospitaliers (systèmes de soins, SdS ci-dessous). Il ajuste les paiements en fonction des risques. Ce système est toutefois assez particulier puisque les SdS sont payés à l'acte et ne sont donc pas "mis au risque" dans l'acception traditionnelle du terme (pas de capitation). L'analyse suivante résume le fonctionnement du système sur la base d'informations collectées par Knutson (1998), Newhouse (1998), Dunn (1998a et 1998b).

Chaque année les SdS proposent un prix mensuel pour prendre en charge un patient "moyen"²⁷. Chaque employeur détermine sa contribution théorique. Celle-ci est en général inférieure au prix le plus faible proposé par les SdS afin d'encourager les employés à faire du prix un critère de choix central. Chaque personne couverte choisit son SdS pour une année et paie à son employeur la différence. Par la suite, l'employeur paie directement les SdS à l'acte.

Une procédure similaire à celle du HIPC est utilisée pour calculer un poids relatif par individu qui reflète son état de santé, puis un poids qui reflète le risque relatif du SdS par rapport à la population totale couverte, dont le risque est normé à 1. Les poids sont plus précisément calculés par groupe de diagnostics ambulatoires (ACG). On pondère la proportion de personnes dans un ACG par le coût moyen de prise en charge observé dans les différents SdS rapporté aux coûts de prise en charge de l'ensemble de la population. Comme les SdS sont payés à l'acte, ils sont capables de faire remonter

²⁷ Hors soins exceptionnels : le BHCAG prend en charge par un autre mécanisme 90% des dépenses qui dépassent un seuil donné.

l'information médico-économique nécessaire. Ces poids sont calculés tous les trimestres sur la base d'informations collectées sur les douze mois précédents.

Le prix mensuel proposé par le plan est multiplié par son risque relatif et constitue pour lui l'objectif à atteindre. Chaque trimestre, on calcule la moyenne des paiements mensuels effectués par le plan sur les douze mois précédents et on la compare à l'objectif calculé pour la même période. Si les paiements sont supérieurs à l'objectif, les honoraires et les prix de l'ensemble des services²⁸ sont ajustés d'un pourcentage tel que, si cette nouvelle grille avait été appliquée, l'objectif aurait été atteint. Inversement, si les paiements moyens sont inférieurs à l'objectif, les prix unitaires sont augmentés. La nouvelle grille de tarifs s'applique pour le trimestre suivant. Autrement dit, si, compte tenu de son niveau de risque, le SdS a utilisé "trop" de soins, il est sanctionné et verra ses prix diminués ; s'il a fait "mieux", la rémunération des prestataires augmente. La différence entre l'objectif et le résultat est mesurée de façon rétrospective mais la sanction ou la récompense sont prospectives.

A l'occasion de la première réévaluation tarifaire, début 1997, on a constaté que les niveaux de risques étaient très différents d'un SdS à l'autre (avec une différence de 35% entre le plus faible et le plus élevé) et les prix ont été ajustés de -10%²⁹ à +7.7%. Le bilan de 1997 a montré que les paiements effectués par les employeurs avaient diminué de 9% par rapport à l'année précédente, le système a donc permis de maîtriser les dépenses. Il a depuis été affiné et il est en cours d'évaluation. Il avait été initialement jugé très complexe, mais il semble que les SdS sont satisfaits (aucun n'a quitté le système et de nouveaux SdS ont même rejoint le BHCAG).

la *Washington State Health Care Authority* (HCA)

La HCA est le sponsor de différentes populations de l'Etat de Washington, dont les employés publics (Dunn, 1998a, Dunn, 1998b, Newhouse et Chairman, 1998, Wilson, et al., 1998). Le mécanisme d'ajustement au risque est en cours de révision. Il ne se fondait avant le début de 1998 que sur un modèle démographique et il intègre progressivement des informations de nature diagnostique. Lorsque ce modèle sera mis en place, les

²⁸ Le groupe médical peut toutefois exclure certains prestataires ou institutions de ce système de point flottant (système utilisé en Allemagne où les honoraires de certains praticiens sont ajustés à la hausse ou à la baisse si leur somme n'est pas égale à une enveloppe macro-économique déterminée). Entre 10 et 30% des prestataires par SdS ne sont pas concernés, mais, dans ce cas, la variation des prix pour les autres doit être plus élevée.

²⁹ Une limite à l'ajustement à la baisse était imposée.

variables d'ajustement seront : l'âge, le sexe, le statut vis à vis de l'assurance (employé, conjoint, dépendant), la structure du ménage et des variables de statut (retraité par exemple), l'information médicale étant prise en compte sous la forme de DRG.

Le déroulement du mécanisme est le suivant : chaque année, les plans proposent une prime par assuré moyen³⁰. Chaque personne effectue son choix de plan et paie la différence entre une contribution de l'employeur déterminée par la HCA et la prime affichée par le plan. Par la suite, les paiements de la HCA versés aux plans sont basés sur les prix proposés par les plans et ajustés au risque. Le modèle utilise des informations démographiques et de diagnostics regroupés par niveaux de coût associés. Par catégorie de risque, on estime une probabilité d'avoir des dépenses positives puis une dépense moyenne quand la personne consomme effectivement. Sur la base de ces informations, un indice de risque relatif pour l'individu est calculé et agrégé au niveau du plan. La contribution totale de la HCA est redistribuée en fonction du risque relatif de chaque plan : le processus d'ajustement consiste à ré-allouer des fonds mais il est neutre au niveau global. Une simulation effectuée en 1997 montre que les assureurs, lorsque l'ajustement est basé sur le modèle démographique, reçoivent entre 90 et 112% de la prime qu'ils proposent dans le processus d'appel d'offre. En ajoutant les informations de diagnostic, la gamme s'élargit (65% à 130%)³¹. Les simulations montrent aussi que ce dernier modèle prévoit mieux les différences de structure de risque entre les coûts.

En conclusion de ces trois exemples, il nous semble intéressant de souligner ou de rappeler plusieurs points de convergence.

1. Dans tous les cas, la motivation principale pour mettre en oeuvre l'ajustement au risque est la crainte de voir se développer la sélection des risques. La population couverte est relativement bien définie, assez large et hétérogène. L'ajustement se fait dans un contexte où les contrats sont standardisés ; les acteurs offrant leurs services sont, chaque année, à la même époque, obligés d'accepter de prendre en charge toute personne qui se présente (*open enrollement*). Ils n'ont pas le droit de tarifier certains facteurs de risque (puisque le *community rating*, voire un prix unique pour tous les assurés sont imposés), mais l'ajustement doit théoriquement les compenser de façon adéquate.

³⁰ Qui varie dans les mêmes proportions pour tous les plans afin de prendre en compte la structure de la famille.

³¹ Ce modèle est mis en oeuvre progressivement et sera utilisé en l'an 2000

2. Pour ce qui est de l'ajustement lui même, les différents programmes semblent tous converger, malgré les difficultés que cela pose, vers une utilisation de l'information de nature diagnostique associée à des coûts relatifs.
3. A l'exception du BHCAG, les modèles se veulent prospectifs. L'information de nature médicale sur le structure de risque est, certes, collectée deux années avant l'ajustement. Mais, pour la période sur laquelle porte l'ajustement :
 - les versements se fondent sur le nombre de personnes effectivement couvertes ;
 - les paiements aux plans sont déterminés ex ante et non en fin de période lorsque les coûts sont observés.

Les modèles reposent donc fortement (et implicitement) sur l'hypothèse que les structures de risque médical des populations couvertes par les différents plans sont stables, ce qui ne va pas sans poser de problèmes lorsque de nouveaux participants rejoignent le champ d'application. En particulier, on peut redouter ce qui se passerait si un nouvel entrant générerait un *turn over* très élevé pour une catégorie de risque particulière (si par exemple, un entrant attirait tous les ménages jeunes et sans enfant). Le BHCAG évite en partie cet écueil car il a adopté un ajustement plus fréquent et il utilise des informations plus récentes (cet ajustement n'est toutefois pas réellement prospectif, comme nous l'avons souligné). Il fait cependant face à des difficultés spécifiques : les ajustements doivent prendre en compte une estimation des dépenses non présentées au remboursement.

4. Quoi qu'il en soit, dans tous les cas, les prestataires de service en présence ont quasiment tous accepté les contraintes qui leur étaient imposées, probablement parce qu'ils étaient largement associés au processus d'élaboration du mécanisme. De plus, de nouveaux prestataires ont proposé leurs services après que le mécanisme a été implanté.
5. Enfin, une étude plus qualitative a permis de montrer que l'existence d'un ajustement diminuait la crainte des plans d'être victimes de stratégies de sélection de leurs concurrents ou d'effets de segmentation spontanée (Knutson, 1998) et pouvait les conduire à limiter leur propre stratégie de sélection (Bertko et Hunt, 1998).
6. Les employeurs et les employés n'ont pas souffert de l'implantation des mécanismes et s'en déclarent globalement plutôt satisfaits (Dunn, 1998b, pp.145).

7. La principale leçon de ces expériences est toutefois que l'application concrète de l'ajustement au risque nécessite : "*a series of trade-offs between risk-adjustment theory, data availability, [and] the willingness of participants to play the game*"³² (Bertko et Hunt, 1998).

2.1.2 Irlande

En Irlande, comme aux Etats-Unis, l'assurance est volontaire. En revanche, elle ne porte pas sur l'intégralité de la couverture comme nous l'avons vu dans le chapitre sept. Néanmoins, elle participe à l'équilibre général du système et le ministère de la santé a jugé bon, pour préserver cet équilibre, de mettre en place un ajustement des risques lorsque le marché s'est ouvert à la concurrence.

a) La fin du monopole de l'assurance volontaire

L'assurance privée volontaire est, depuis 1996, soumise à un système de péréquation. Depuis 1957, une entreprise publique d'assurance maladie, le *Voluntary Health Insurance Board* (VHI) détenait légalement un quasi-monopole sur le marché de l'assurance maladie. Traditionnellement, tous les assurés payaient la même prime et l'assureur ne pouvait refuser de couvrir une personne ou de renouveler son contrat. Ce mode de fonctionnement n'était pas imposé par une réglementation spécifique. En 1992, la troisième directive européenne "non vie" impose aux Etats membres d'ouvrir les marchés d'assurance aux compagnies d'assurance européennes. Le texte prévoit toutefois que certaines dispositions spécifiques mises en place par les Etats sur le marché de la maladie pour protéger le public pourront s'appliquer aux entrants sur le marché. Le *Health Insurance Act* de 1994 fixe explicitement un certain nombre de principes :

- *community rating*. L'assureur ne peut faire varier la prime en fonction du sexe, de l'orientation sexuelle (sic) ou de l'état de santé actuel ou prévisible des assurés. Quelques exceptions sont prévues : les enfants et les étudiants couverts par leurs parents peuvent bénéficier d'un abattement de 50% maximum et les contrats de groupe peuvent obtenir une réduction de 10% maximum.

³² Une série d'arbitrages entre la théorie de l'ajustement au risque, la disponibilité des données et la volonté des participants de jouer le jeu.

- *open enrollment et guaranteed renewal* : toute personne de moins de 65 ans a droit à l'assurance et aucune personne ne peut se voir refuser le renouvellement de son contrat. Les assureurs ont le droit d'imposer les restrictions sur la couverture des pathologies déclarées, mais celles-ci sont strictement réglementées.

Les *Health Insurance Regulations* de 1996 introduisent, à l'initiative du ministère de la santé irlandais, un mécanisme de péréquation des risques entre assureurs afin d'éviter que la concurrence ne passe par la sélection. Avant de présenter ce mécanisme, notons que la réforme du fonctionnement du marché de l'assurance trouve son origine dans la réglementation européenne. Une institution publique en situation de quasi-monopole doit s'ouvrir à la concurrence d'assureurs étrangers. Le gouvernement autorise cette concurrence mais cherche à préserver ce qui est perçu comme un avantage acquis : une tarification de l'assurance indépendante des risques. La dynamique est donc assez différente de celle présente dans d'autres pays : l'introduction de la concurrence n'est ici pas perçue comme une nécessité pour maîtriser les coûts. Ceci s'explique probablement en partie par le fait que l'assurance est volontaire, financée de façon privée, qu'elle ne conditionne pas l'accès au système de soins et qu'elle est plutôt réservée à une partie aisée de la population. Toutefois, nous avons vu que l'assurance privée contribuait à alléger les charges de l'Etat (puisque'elle finance une partie de l'hospitalisation par ailleurs gratuite pour tous). Le maintien du système et de la forme de solidarité qu'il génère entre différents niveaux de risque est jugé primordial.

b) le mécanisme d'ajustement

Un rapport élaboré par un groupe d'experts indépendants et commandité par le ministère de la santé³³ détaille et justifie les choix réalisés dans la méthode d'ajustement puis recommande que certaines améliorations soient apportées le plus rapidement possible. La méthode d'ajustement choisie est résolument rétrospective : l'objectif est de corriger ex post ce qui, dans les différences de coût observées entre assureurs, provient de différences de structure de risque et non de différences d'efficacité. Les éléments de coût qui proviennent du fait que la structure de risque est différente sont corrigés et on remplace la valeur propre observée pour l'assureur par la valeur moyenne dans la population. Les autres éléments ne donnent pas lieu à correction. Pour ce faire, on

³³ *Report of the Advisory Group on the Risk Equalisation Scheme*. Irish Ministry for Health and Children, 8 April 1998.

détermine en première approche les dépenses normatives brutes de chaque assureur. La somme de ces dépenses normatives brutes est normée de façon à ce que les dépenses normatives ajustées totales soient égales aux coûts totaux (le mécanisme est globalement neutre). Il s'articule concrètement autour de la correction des différences de structure par âge et par sexe des populations couvertes (R catégories indicées r) entre I assureurs indicés i .

Les remboursements effectués par l'assureur i avant ajustement peuvent être écrits de la façon suivante :

$$n_i \sum_r \left(P_{ir} \sum_d F_{ird} C_{ird} \right)$$

Avec :

n_i : nombre total de personnes couvertes par l'assureur ;

P_{ir} : proportion de personnes couvertes par l'assureur, appartenant à la catégories âge-sexe r ;

F_{ird} : fréquence d'un type de procédure médicale d dans la catégorie r pour l'assureur i ;

C_{ird} : coût de cette procédure pour l'assureur i .

On estime que P_{ir} reflète la sélection (la distorsion de la structure âge sexe d'un assureur à l'autre) et on le remplace par P_r , la part relative de la catégorie âge-sexe r dans la population totale couverte. A l'opposé, on postule que le coût unitaire C_{ird} relève de la responsabilité individuelle de l'assureur.

Deux scénarios sont ensuite explorés :

- dans le premier, on considère que la fréquence observée des actes médicaux relève de la responsabilité de l'assureur. Les calculs sont alors très simples : le coût moyen par catégorie de risque r de l'assureur n'est pas modifié par l'ajustement et seules les proportions de chaque cellule sont ajustées pour être égales à la proportion moyenne observée dans la population³⁴. Cette méthode revient en fait à assumer que les seuls facteurs qu'il importe de corriger sont l'âge et le sexe, puisqu'on estime que toute hétérogénéité de coût à l'intérieur d'une catégorie r relève exclusivement d'une gestion

³⁴ Les coûts par assureur ainsi calculés sont ajustés d'un facteur α tel que la somme pondérée par ce facteur des coûts théoriques des assureurs est égale à la somme observée des coûts.

plus ou moins efficace, ce qui est jugé par le groupe d'experts comme erroné et potentiellement dangereux ;

- dans le second, on considère que la fréquence reflète des différences de structure de risque et doit donc être remplacée par F_{rd} , fréquence moyenne observée de la procédure dans la population. La validité de cette approche dépend de la nature des "procédures médicales" considérées. Rappelons que l'assurance couvre essentiellement des dépenses hospitalières et des soins de spécialistes en milieu hospitalier. Si on regroupe les dépenses par groupe homogène de malade (*GHM* ou *DRG*), on peut alors postuler que la fréquence observée du DRG dans une catégorie r par un assureur i dépend largement de l'état de santé de la population couverte et que le coût unitaire du DRG est lié à la capacité de gestion de l'assureur. On peut donc remplacer la fréquence du DRG observée par l'assureur par la fréquence moyenne observée dans la population³⁵. Cette approche est recommandée par le groupe d'experts auteurs du rapport d'évaluation du schéma d'ajustement.

L'ajustement mis en œuvre en 1996 constitue une approximation dans laquelle on remplace les DRG par des durées de séjour. Le rapport souligne qu'elle génère des incitations perverses : si un assureur parvient à réduire la durée de séjour pour une pathologie (et donc à être plus efficace), compte tenu du mécanisme d'ajustement, l'opération entraînera une augmentation nette des coûts.

c) Bilan provisoire

Les experts indépendants qui ont mis au point le mécanisme d'ajustement, le jugent imparfait (parce qu'il se heurte à de nombreuses contraintes techniques), mais ils adoptent une position très pragmatique.

Selon eux, dans la mesure où le *VHI* est, pour des raisons historiques, en position dominante sur le marché, les nouveaux entrants, qui se tourneront en priorité vers des catégories relativement peu risquées (entreprises, jeunes...), obtiendront un avantage comparatif qui leur permettra de pénétrer sur le marché. Les défauts du mécanisme d'ajustement serviront donc la concurrence à court terme.

Les experts recommandent néanmoins que l'ajustement sur la base des DRG soit mis en place le plus rapidement possible. On peut enfin remarquer que le rapport préconise une

³⁵ On norme ensuite les dépenses pour que la somme pondérée des coûts théoriques soit égale à la somme observée.

certaine forme de standardisation des contrats. Plus précisément, il recommande que le *community rating* et l'ajustement des risques ne portent que sur un panier de soins jugés essentiels, au motif que l'Etat doit se contenter de garantir l'accès à une couverture de ces soins.

2.1.3 La Suisse

En Suisse, l'assurance maladie était, jusqu'à une période récente, volontaire, servie sur une base non lucrative et très réglementée. Son fonctionnement n'était pas satisfaisant et la loi LAMal, en introduisant l'assurance universelle, a tenté de limiter les effets de sélection, avec des résultats mitigés.

a) Le cadre général : la loi LAMal

Le système Suisse a été profondément réformé en 1994 car l'ancien système générait une inéquité et des problèmes d'accès à l'assurance croissants (OCDE, 1994, pour un diagnostic détaillé). Dans le même temps, les méthodes de tarification limitaient les incitations à l'efficacité : les tarifs étaient calculés en fonction de l'âge à l'entrée en assurance et, passé un certain âge, les assurés étaient de facto captifs de leur assurance.

Rappelons les grandes lignes de la loi LAMal (18 mars 1994) :

- tous les assureurs doivent offrir un contrat qui couvre obligatoirement un panier de services défini précisément par la loi. Une franchise fixée par le Conseil fédéral et une participation de 10% aux frais en deçà, dans une limite annuelle légale, sont prévues au contrat. L'assureur peut proposer une gamme de franchises plus élevées, déterminée légalement, moyennant diminution des primes selon des conditions précises, mais il ne peut offrir plus ou moins de services que ce qui est prévu par la loi ;
- l'assureur ne peut refuser de couvrir ou de renouveler le contrat des personnes qui en font la demande ;
- tous les Suisses doivent s'assurer.

Une des premières conséquences de la réforme a donc été d'introduire une assurance obligatoire et universelle. Le principe du choix de l'assureur et de la concurrence entre caisses qui caractérisaient le système avant la réforme ont été maintenus. Les contraintes qui pèsent sur les assureurs (acceptation obligatoire en assurance) et les assurés

(interdiction de sortir du système si on présente un faible niveau de risque) sont les premiers éléments destinés à garantir une mutualisation large des risques et à éviter la sélection explicite ou l'auto-sélection. En outre, les contrats sont très largement standardisés. Tout assureur qui offre des contrats dits particuliers³⁶, doit aussi offrir le contrat de base et ne peut donc espérer bénéficier d'une sélection favorable massive que s'il est le seul à offrir ce type de contrat. En tout état de cause, il devra tout de même couvrir les personnes qui présenteraient de plus hauts niveaux de risque et qui choisiraient le contrat standard.

Le système de financement renforce ces mécanismes de solidarité entre malades et bien portants et de temporisation de la sélection. Les assureurs sont contraints d'utiliser une forme stricte de *community rating*. Les primes peuvent varier par canton selon le lieu de résidence de l'assuré "s'il est prouvé que les coûts diffèrent selon les cantons et les régions" (LAMal, Titre I Chapitre 5, section 2, Art 61). Sous les mêmes conditions, il peut y avoir jusqu'à trois niveaux de primes par canton. Les enfants (moins de 18 ans) et les étudiants de moins de 25 ans peuvent bénéficier d'une réduction de prime. En dehors de ces conditions, les primes sont uniformes et doivent être approuvées par le Conseil fédéral. Le mode de tarification cherche donc à imposer une solidarité forte entre hauts et bas risques. Dans le même temps, les cantons doivent subventionner les primes des personnes défavorisées.

b) L'ajustement des risques : un essai manqué

Enfin, un mécanisme d'ajustement des risques a été mis en place sans lequel la sélection, du fait des contraintes tarifaires légales strictes, pourrait être assez forte, même si les caisses ou les assureurs qui ont le droit de servir l'assurance légale ne peuvent réaliser un profit sur cette activité. Le mécanisme d'ajustement n'est imposé aux assureurs que pour les dix années qui suivent l'entrée en vigueur de la loi (1996), pour des raisons qui ne sont pas explicitées.

³⁶ Nous avons signalé le contrat à franchise majorée. Les assureurs peuvent aussi offrir des contrats à bonus malus (très réglementés et dans le détail desquels nous n'entrerons pas) et peuvent constituer des HMO (moyennant là encore diminution des primes pour les assurés).

L'ajustement prévu dans l'Ordonnance de compensation des risques dans l'assurance maladie du 12 avril 1995 fonctionne selon les principes suivants :

1. les assurés (de plus de 18 ans) sont répartis en groupes de risque (par tranches d'âge de 5 ans et par sexe) ;
2. on détermine, par canton et par groupe de risque le coût moyen global, et les coûts par groupe de risque ;
3. si, pour un groupe de risque, le coût moyen de prise en charge par l'assureur est inférieur à la moyenne du canton, il paie par assuré une redevance égale à la différence à la moyenne ; inversement, il reçoit une redevance égale à la différence à la moyenne s'il a des coûts plus élevés que celle-ci.
4. la compensation est valable pour l'année courante, elle est calculée sur la base des coûts de cette année là. Une compensation provisoire est fixée en cours d'année (sur la base des informations de l'année précédente) et les règlements définitifs interviennent l'année suivante.

Ce mécanisme d'ajustement, dont la simplicité peut paraître séduisante, semble critiquable à plusieurs égards :

- en premier lieu, il ne prend en compte que l'âge et le sexe, facteurs de risque qui n'expliquent qu'une part infime de la variance explicable (cf. § 1.1.2).
- ensuite, il risque de gommer les incitations à la concurrence puisque l'objectif implicite de l'assureur est d'avoir le même coût moyen par canton que ses concurrents. Si un assureur a la même structure de risque mais est plus efficace qu'un autre, dont le coût est supérieur au coût moyen observé, il doit lui reverser une compensation. A l'inverse, si un assureur n'est pas efficace, il sera compensé. Il n'a donc qu'un intérêt réduit à maîtriser les coûts.
- enfin, comme le soulignent Beck et Zweifel (1995), la limite de 10 ans prévue pour l'ajustement au risque n'a guère de sens, car tant que le mode de tarification légal est maintenu, le risque de sélection est très élevé.

Le fonctionnement de l'ajustement au risque, malgré ces limites, ne semble pas avoir été modifié depuis l'ordonnance de 1995 (en tout cas, aucun texte législatif n'en atteste).

Le bilan actuel de la réforme est mitigé. Quelques créations de HMO ont été abondamment commentées dans la presse. Si les résultats de ces structures en termes de maîtrise des coûts sont plutôt bons, il n'est pas toujours facile de faire la part de la sélection dans les résultats obtenus (Perneger *et al.*, 1995). Un autre effet de la réforme a été d'accélérer la concentration des caisses et on estime, à l'horizon 2005, que trois caisses domineront le marché (Mosca, 1997). Mais la réforme a été mal tolérée par le public. Tout d'abord, les primes ont augmenté très fortement en 1996 (5,4% en moyenne entre 1995 et 1996). Cette augmentation, qui s'explique en partie par l'augmentation du panier de biens et services couverts, était pourtant prévisible et elle correspond à un transfert de charges : les paiements directs des ménages aux prestataires de soins ont diminué. Les observateurs estiment cependant que les assureurs ont largement dépassé les augmentations qui auraient été strictement nécessaires. Le passage au nouveau système a en outre été très douloureux, car le mode de tarification a induit des augmentations individuelles de prime allant jusqu'à 60% (Mosca, 1997). Dans le même temps, de nombreux cantons n'ont pas su mettre sur pied un système de subventionnement qui fasse coïncider l'augmentation générale des primes avec leur réduction pour les personnes à revenu modeste. D'autres, qui ont mobilisé l'intégralité des subventions, n'ont pas réussi à satisfaire l'objectif social de la Loi car les ressources étaient insuffisantes (Dreifuss, 1996). Parallèlement, des réformes portant sur l'offre ont été retardées ou se sont avérées incohérentes, ce qui a renforcé le mécontentement des assurés. Enfin, les assureurs se sont lancés dans une concurrence effrénée selon des méthodes discutables : ainsi, des études (citées par Mosca) ont montré que, dans les cantons où la concurrence était forte, les assureurs sous-estimaient les primes, et prélevaient les ressources complémentaires en majorant les primes sur d'autres régions. Comme la comptabilité des caisses et les informations statistiques disponibles sont très frustrées, il est très difficile de connaître les coûts réels générés par les assurés et donc la pertinence du montant de leurs cotisations. Le système ne semble toujours pas stabilisé. Ainsi, la seconde caisse de la Confédération a décidé de mettre fin en septembre 1998 à son activité d'assurance au titre de la LAMal dans 8 cantons qui comportaient les proportions les plus fortes de personnes âgées. Plus de cent mille personnes se sont donc trouvées dans l'obligation de chercher une autre assurance.

2.1.4 Principales leçons sur les systèmes de tradition concurrentielle

Les exemples issus de ces trois pays permettent de tirer une première série de conclusions.

Tout d'abord, les exemples américains et irlandais, détaillés à dessein, soulignent que la mise en œuvre d'un mécanisme d'ajustement au risque est un processus éminemment complexe. A l'inverse, un système dont la simplicité peut paraître séduisante, ne permet pas de mettre en œuvre les incitations souhaitées (Suisse). Il est vrai que l'expertise est récente et assez peu diffusée, mais il semble nécessaire de rappeler avec Luft que "*We must be aware of the need to adjust for risk differences and not to simply "assume" that an adequate mechanism exists or will be available in time for its effective use*"³⁷.

En second lieu, les exemples américains montrent que la reconnaissance de la nécessité de gérer la concurrence, pour en obtenir des bénéfices en termes d'efficacité productive, peut émerger spontanément. Pour qu'une telle organisation de la concurrence soit mise en œuvre, il faut cependant qu'un sponsor intervienne de façon volontariste et qu'il investisse dans le processus, y compris en termes financiers. Il peut alors parvenir à élaborer un système auquel assureurs et assurés adhèrent sur une base volontaire parce qu'ils y trouvent un intérêt.

Enfin, les systèmes américains et suisse, qui sont plutôt de type libéral, promeuvent implicitement une forme d'équité entre les hauts et les bas risques, objectif qui est, par contre, affiché en Irlande. Les primes payées par les assurés ne tiennent compte que d'un nombre limité de facteurs de risques et l'ajustement a lieu sur d'autres. Dans la mesure où l'information est disponible (mis à part les problèmes de délais de remontée de l'information), on pourrait envisager que le sponsor impose aux assureurs de les utiliser pour offrir des primes mieux ajustées aux risques individuels. Le résultat en terme d'efficacité serait strictement identique, mais l'équité du système y serait moindre. On peut rappeler à cette égard que la péréquation des risques constitue une forme de protection collective contre l'éventualité de devenir un mauvais risque et présente donc aussi un avantage à cet égard.

³⁷ Nous devons avoir conscience de la nécessité d'ajuster les risques et pas seulement faire l'hypothèse qu'un mécanisme adéquat existe ou qu'il sera disponible pour une utilisation effective lorsque le besoin se fera sentir.

Pour finir, il semble nécessaire de rappeler que les systèmes existants restent rares et que leur mise en œuvre est récente. Ils n'ont donc jamais été évalués en termes d'analyse coût-bénéfice. De nombreux employeurs aux Etats-Unis se contentent à l'heure actuelle d'utiliser des outils plus rudimentaires pour limiter la sélection (*guaranteed issue, guaranteed renewal, ..*). La preuve reste encore à faire de l'efficacité plus grande d'un système dans lequel on a introduit un ajustement au risque.

2.2 Introduction de la concurrence et délégation de gestion

Dans la seconde série d'exemples que nous étudions, les entités qui mettent en œuvre l'assurance maladie ne sont pas responsables au premier chef de la détermination du montant qui leur permet de faire face à leurs engagements ou de la collecte des fonds. Une entité intermédiaire, publique, remplit cette fonction. Dans ce type de système, la participation au financement de l'assurance est généralement explicitement déconnectée du risque individuel. Des contributions proportionnelles au revenu (impôts ou cotisations) sont prélevées sur les personnes couvertes ou sur l'ensemble de la population. Ceci permet, en théorie, de régler simultanément les problèmes d'accès à l'assurance et d'équité (entre personnes riches et pauvres et entre personnes malades et en bonne santé).

Les fonds sont alloués aux entités en concurrence sous la forme d'une capitation ajustée au risque. Ce refinancement se veut actuariel. Les entités auxquelles on confie la gestion de l'assurance sont alors responsabilisées sur le plan financier et s'approprient tout ou partie du "profit" de cette activité. Le premier objectif dans ce type de système est d'améliorer l'efficacité productive.

Parmi les pays étudiés, certains partent d'une situation antérieure dans laquelle la concurrence était absente (Pays-Bas, Allemagne) ou dans laquelle les organismes n'étaient pas responsabilisés sur d'éventuels déficits (Belgique et dans une certaine mesure, Israël). Le cas de Medicare est un peu particulier. Dans la situation antérieure, les dépenses des personnes couvertes étaient remboursées directement par la HCFA qui fonctionnait comme une assurance traditionnelle. Depuis quelques années, des plans (HMO) peuvent intervenir comme intermédiaires entre les assurés et l'administration. Ils reçoivent une capitation ajustée au risque et sont chargés de gérer le risque des assurés. L'objectif recherché par l'administration reste, là encore, l'amélioration du processus

d'administration des soins, l'augmentation de l'efficacité productive et à terme, la maîtrise des coûts.

2.2.1 Les Pays-Bas : archétype de l'introduction de la concurrence

La mise en concurrence des caisses publiques aux Pays-Bas est l'exemple le plus connu d'introduction d'une forme de responsabilité financière³⁸.

a) Le contexte et les principes de la réforme

L'assurance obligatoire des soins courants est gérée par une trentaine de caisses d'assurance maladie, qui sont des organismes à but non lucratif. Elles couvrent les salariés dont le revenu n'excède pas un certain seuil et offrent toutes les mêmes prestations, définies par l'Etat. Celui-ci prélève sur la population couverte obligatoirement par ce contrat une cotisation proportionnelle au revenu.

Avant la réforme, les caisses bénéficiaient de monopoles régionaux pour la couverture des personnes relevant de l'assurance maladie obligatoire. Elles recevaient du ZiekenfondsRaad (ZFR) un remboursement intégral des dépenses que leurs affiliés engendraient. En cas de déficit global, les taux de cotisation étaient augmentés ou l'Etat complétait le financement. Les caisses publiques étaient avant tout des organismes de liquidation.

La réforme a consisté à autoriser les bénéficiaires du régime d'assurance publique à choisir leur caisse d'affiliation (mise en concurrence sur une base nationale) et à en changer une fois par an (*open enrollement*), tout en maintenant l'obligation faite aux caisses de prendre en charge toute personne en faisant la demande (*guaranteed issue et renewal*). Parallèlement, leur mode de financement a été fondamentalement bouleversé : les caisses sont désormais financées de façon prospective en fonction du profil de risque de leurs assurés (et non plus remboursées systématiquement de manière rétrospective). Elles sont responsabilisées financièrement et donc en principe incitées à minimiser les coûts de gestion et à gérer efficacement les prestations des assurés³⁹. Le coût de l'assurance pour un individu reste, en revanche, complètement dissocié du risque qu'il présente.

³⁸ Une partie de cette analyse a été précédemment développée dans Couffinhal et Rochaix (1998).

³⁹ La stratégie utilisée par l'Etat néerlandais pour mettre en concurrence les caisses publiques d'assurance maladie relève de cette logique qui était sous-jacente aux propositions de DEKKER (Couffinhal et Rochaix, 1998).

Plus précisément, le financement des caisses est assuré par deux sources :

- une capitation ajustée au risque, versée par le ZFR qui centralise les cotisations ;
- une prime versée directement par chaque assuré à la caisse publique auprès de laquelle il choisit de s'inscrire. La contribution des primes collectées directement par les caisses au financement global de leurs dépenses n'est pas très élevée (moins de 6%). Le montant de la prime est obligatoirement identique pour tous les assurés d'une même caisse mais il est librement déterminé par celle-ci. Elle peut la diminuer si elle réalise un bénéfice et elle doit l'augmenter dans le cas contraire. Cette prime reflète donc théoriquement les différences d'efficacité productive d'une caisse à l'autre. Les assurés sont supposés se fonder sur ce critère pour arbitrer entre les caisses publiques. Comme il s'agit d'un financement marginal, c'est avant tout le niveau relatif des primes entre différentes caisses qui doit ici jouer un rôle incitatif. En 1997, l'écart annuel constaté entre la plus haute et la plus basse des primes est de 300 francs. Il semble augmenter au cours du temps : il était de 30 francs en 1995 et de 90 francs en 1996.

b) La détermination du budget des caisses

En l'année n-1, le gouvernement retient un niveau global de dépenses des caisses (hors primes nominales) pour l'année à venir ainsi que sa répartition entre grandes catégories de soins. Le ZFR calcule un budget par caisse sur la base des informations transmises par celles-ci sur les effectifs prévus par catégorie de risque. Ce budget est composé des capitations individuelles modulées en fonction de la classe de risque. En fin d'année, les caisses transmettent les données concernant les risques effectifs et les dépenses engagées. Le budget est recalculé pour tenir compte des effectifs réels et des ajustements sont effectués pour déterminer la part du décalage entre ce budget et les engagements financiers de la caisse qui demeurera sous sa responsabilité.

Calcul de la capitation ajustée au risque

Pour le calcul de la capitation pondérée au risque, le ZFR dispose des coûts moyens par catégorie de risque calculés sur l'ensemble des caisses pour les années précédentes. Connaissant la représentation dans la population globale des facteurs de risque, il alloue des poids relatifs à chaque catégorie de risque. Ces poids prennent en compte la différence de coût observée entre les catégories de risque et leur représentation dans la population. Ils sont ensuite appliqués au budget de la catégorie de soins, en fonction des

caractéristiques de la population effectivement couverte par chaque caisse, afin de déterminer la part qu'elle recevra.

Dans un premier temps (1993 et 1994), la capitation versée aux caisses se base sur l'âge et le sexe (38 classes), critères qui apparaissent rapidement trop grossiers. Depuis 1995, la capitation prend en compte un critère d'incapacité (de travail) et un critère régional, qui reflète probablement à la fois des effets d'offre (le coût des soins) et de demande (le profil de risque des populations couvertes). Ce dernier critère prend plus précisément en compte le degré d'urbanisation des lieux de résidence des personnes couvertes (cinq catégories). Les variables utilisées pour calculer les capitations ajustées au risque sont donc au nombre de quatre (âge, sexe, région et incapacité). Ces facteurs d'ajustement présentent l'avantage d'être facilement observables et contrôlables. Il est peu vraisemblable qu'ils changeront dans les années à venir. En effet, le modèle d'ajustement au risque, si imparfait soit-il, est suffisamment compliqué d'après le ZFR pour que les caisses ne puissent déterminer objectivement les risques pour lesquels elles seraient éventuellement sur-dotées et qu'elles seraient alors tentées d'attirer. En outre, le ZFR estime que l'imperfection de l'ajustement au risque constitue en soi une incitation pour les caisses à contrôler les coûts. Ces affirmations se fondent aussi implicitement sur l'idée que les caisses ne sont pas habituées à évoluer dans un univers concurrentiel, à évaluer le profil assurantiel de leurs assurés et sont peu capables de développer rapidement des techniques de sélection suffisamment sophistiquées pour n'être pas observables et donc sanctionnables.

La capitation ajustée au risque porte sur l'ensemble des soins de ville. Depuis 1996, les soins hospitaliers (qui représentent 55% des dépenses) sont aussi partiellement inclus dans ce système de capitation. Le budget était jusqu'alors calculé sur une base historique. Il est désormais scindé en deux parties représentant respectivement les coûts fixes (63%), qui restent estimés sur une base historique et les coûts variables, calculés selon quatre indicateurs de volume et soumis pour leur part au modèle de capitation ajustée au risque.

Ajustements ex post

Ce système de financement sur la base d'une capitation ajustée au risque est imparfait et s'accompagne d'une mutualisation obligatoire des hauts risques.

Le système de mutualisation mis en œuvre aux Pays-Bas en 1997 s'inspire du modèle théorique que nous avons évoqué dans la section précédente. Il n'opère cependant pas

une mutualisation des coûts engendrés par les personnes que l'assureur choisit ex ante mais des coûts constatés. Il oblige en fait l'ensemble des caisses à supporter collectivement le surcoût engendré par des personnes dont les dépenses individuelles dépasseraient 4500 Florins par an. Ce mécanisme permet une mutualisation de 90% du surcoût et s'opère dans la limite des budgets déterminés par l'Etat pour l'ensemble des caisses. Il ne génère donc pas de déresponsabilisation collective et constitue simplement une nouvelle réallocation du budget global.

Le tableau ci-dessous présente une simulation de ce mécanisme dont l'objet n'est pas tant le réalisme que l'illustration du mécanisme mis en place en 1997 aux Pays-Bas.

Tableau 1 : Mécanisme de mutualisation des hauts risques aux Pays-Bas

	Caisse A	Caisse B	Total
Coûts totaux observés	100 000	50 000	150 000
Nombre de personnes générant un coût supérieur au seuil de 4 500 NLG	5	0	
Coût de chacune de ces personnes	5 000		
Montant mutualisé	$5 \times (5\,000 - 4\,500) \times 90\% = 2\,250$		
Budget initial normatif de la caisse (après ajustement des risques)	95 000	45 000	140 000
Taux de participation au financement du pool	$2\,250 / 140\,000 = 1,61\%$		
Paiement de chaque fond au pool (1,61% de son budget initial)	- 1 527	- 723	2 250
"Subvention" au fond A	2 250		
Nouveau budget des caisses	95 723	44 227	140 000

Responsabilité finale des caisses

Par ailleurs, la responsabilité financière individuelle des caisses publiques est tempérée par une réglementation qui opère un partage des déficits éventuels entre caisses. Ceci permet de maintenir entre elles de façon transitoire une solidarité et se justifie d'autant que l'ajustement au risque de la capitation est imparfait. Ainsi, en 1997 et 1998, 20% des déficits individuels ont été partagés entre les caisses publiques.

Enfin, les déficits effectivement attribués aux caisses font l'objet d'allègements supplémentaires pour tenir compte du fait que leur marge de manoeuvre pour peser sur

les coûts n'est pas encore suffisante, l'Etat réglementant encore largement l'offre de soins. Cette reprise de déficit par le ZFR se justifie aussi par la volonté de ne pas confronter brutalement ces organismes à leur responsabilité financière, alors qu'ils assumaient jusqu'à un passé récent le simple rôle de liquidateur de prestations. On constate toutefois une augmentation rapide de la part du déficit final assumée par les caisses (5% en 1991 65% en 1998). Cette prise en charge du déficit se traduit par l'augmentation des primes nominales supportées par ses assurés pour l'année suivante.

c) Bilan provisoire des effets de la mise en concurrence des caisses publiques

Quand bien même le recul manque pour juger de l'efficacité d'ensemble de la mise en concurrence des caisses publiques, apparaissent déjà quelques difficultés.

Si l'on constate une amélioration de la qualité du service rendu, par la mise en place (par exemple) de comités permettant d'enregistrer les doléances des usagers (*complaints committees*), on ne peut manquer de noter la pérennité des comportements d'écrémage, malgré la présence d'une capitation ajustée au risque. Alors même que la sélection explicite des individus est interdite, la caisse peut mettre en œuvre des stratégies destinées à attirer les bons risques ou à dissuader les mauvais risques de s'assurer auprès d'elle (des comportements de démarchage sélectif ont été observés). Le pan le plus important des bénéfices attendus de la concurrence entre caisses, à savoir leur capacité à négocier avec les prestataires de soins, demeure pour l'instant embryonnaire. Les caisses sont certes obligées de conclure des conventions avec les prestataires, qui portent sur les prix, les quantités et la qualité, mais la marge de manoeuvre demeure réduite du fait de l'existence de budgets globaux et de la structure peu concurrentielle de l'offre de soins. L'évolution du droit néerlandais de la concurrence (qui a pour origine la mise en conformité du droit néerlandais avec le droit européen) devrait permettre dans quelque temps aux caisses de renforcer leur pouvoir de négociation et de limiter la collusion entre prestataires⁴⁰.

Dans l'immédiat, le conventionnement des prestataires leur permet de collecter de l'information qui constituera un bon moyen de procéder à une surveillance de la qualité et du contenu des services (*utilization review*). De même, ces conventions, en rapprochant les caisses et les prestataires, devraient encourager la coopération et l'introduction

⁴⁰ van de Ven, entretien, Avril 1998.

(rendue particulièrement nécessaire dans un environnement de plus en plus compétitif pour les caisses) de techniques de *managed care* ou de *disease management* pour certaines catégories d'assurés. On peut aussi s'attendre à ce que les caisses soient amenées à mettre en place des mesures de réduction du risque moral côté usagers.

L'hypothèse centrale à la mise en œuvre de la concurrence est que les assurés effectuent des arbitrages rationnels entre caisses (en fonction de la prime), renforçant l'émulation entre elles sur le prix et la qualité des services offerts. On constate aujourd'hui que l'arbitrage demeure timide, en raison du différentiel encore peu significatif entre primes (l'incitation financière que représente la prime nominale n'est, semble-t-il, pas suffisante pour entamer des démarches afin de changer de caisse), mais surtout, en raison de l'inertie des comportements des assurés. Même si certaines associations de consommateurs commencent à publier des informations comparatives sur les prix et les services des caisses, la grande majorité des assurés ne met pas encore à profit cette possibilité de "voter avec les pieds". Enfin, il semble que les individus hésitent à exercer leur droit, par crainte de se voir imposer des conditions moins favorables que celles dont ils bénéficient dans leur caisse d'origine en ce qui concerne l'assurance volontaire, dite "du troisième étage", qui couvre les soins non pris en charge par le contrat public (chapitre 7).

2.2.2 La Belgique

La réforme du système belge mise en œuvre au milieu des années quatre-vingt-dix, s'inspire assez largement de celle des Pays-Bas.

a) le contexte général

La quasi-totalité de la population belge est couverte par une assurance maladie obligatoire⁴¹. Des organismes assureurs (OA) ou mutualités à but non lucratif gèrent le régime d'assurance des travailleurs indépendants et celui de la population générale. La Belgique pratique depuis toujours la pluralité des OA, sans restriction légale à la mobilité des assurés d'une caisse à l'autre. La segmentation de la population entre les six grandes catégories de mutualités, regroupées en unions nationales, se fait avant tout selon des

⁴¹ Principales références : "Responsabilité financière des assureurs : le point". *Interface* (1998), Rapport de la commission pour la responsabilité financière des organismes assureurs, Mars 1994, *Calcul de la responsabilisation financière des organismes assureurs pour l'exercice 1995*. Institut National d'Assurance maladie et d'Invalidité, Note CSS 97/8 et Hermesse et Beekmans (1998)

critères linguistiques, religieux et politiques. La concurrence entre mutualités, qui reste marginale, porte, sur des services annexes non directement liés à l'assurance maladie. Conjointement, les unions détiennent un monopole sur le marché de l'assurance.

Le financement de l'assurance maladie obligatoire est principalement assuré par des cotisations patronales et salariales ainsi que des subsides de l'Etat venant combler le manque à gagner pour les personnes exonérées (chômeurs, veuves...). Traditionnellement, les ressources imputées aux différents OA provenaient des cotisations de leurs propres affiliés. Les assiettes de cotisation étant différentes d'une caisse à l'autre, les mutualités ne recevaient pas une proportion des cotisations égale à la proportion de la population qu'elles couvraient⁴² (Nonneman et van Doorslaer, 1994).

b) La réforme du financement et le mécanisme de responsabilisation

La volonté de maîtriser les dépenses de santé a progressivement conduit les autorités à augmenter la responsabilité financière des caisses. Depuis 1990, le ministère de la santé détermine le budget global de financement des dépenses de santé par secteur de soins (hôpitaux, médicaments...). Depuis 1994, un taux maximal de croissance des dépenses est imposé. La loi du 15 février 1993, a mis en œuvre une responsabilisation progressive des organismes assureurs selon des modalités qui se rapprochent de celles utilisées aux Pays-Bas.

Le mécanisme général

Le budget de l'assurance mentionné ci-dessus est réparti entre les mutuelles selon deux clés :

- une première répartition du budget est calculée en fonction d'une capitation ajustée au risque (calcul normatif) ;
- une seconde répartition est calculée en fonction de la part de chaque OA dans les dépenses de soins totales observées pour l'exercice concerné (calcul sur les prestations de l'année).

⁴² Traditionnellement, les mutualités chrétiennes (44.7% de la population en 1990) recevaient une part des cotisations relativement élevée (48.2%) et les mutuelles socialistes (28.9%), une part inférieure (25.3%). Les transferts implicites nécessaires à l'équilibre global du système entre les Flandres plus riche et la Wallonie francophone plus pauvre sont stigmatisés par certains hommes politiques qui veulent que la gestion de l'assurance maladie soit déléguée aux régions, contre la volonté de ceux qui estiment que le système de couverture social est fédérateur et doit relever de la responsabilité fédérale.

Le budget théorique de chaque OA est une somme pondérée des budgets calculés selon les deux méthodes. De 1995 à 1997, 10% du budget total est versé selon le calcul normatif. Ce taux augmente progressivement et devrait atteindre 30% en l'an 2000⁴³.

Lors de la clôture des comptes⁴⁴, cet objectif est comparé aux dépenses comptabilisées. Chaque OA est responsable à raison de 15% de la différence entre les deux montants (bonus ou malus)⁴⁵. Il est en outre prévu que la responsabilité d'un OA soit limitée à 2% de son objectif budgétaire global. Depuis 1995, afin de faire face à leur responsabilité financière, les OA sont tenus de prélever directement auprès des assurés une cotisation uniforme par OA (dont le montant minimum est fixé légalement). Chaque OA constitue un fond de réserve. S'il réalise un bonus, il est versé à la réserve, dans le cas contraire, l'organisme assureur peut tirer sur la réserve, voire utiliser ses fonds propres ou augmenter la cotisation uniforme.

Le calcul du budget normatif

La capitation sur la base de laquelle le budget normatif est calculé représente la dépense que devrait supporter l'assureur en fonction de sa structure de risque et des caractéristiques de sa population assurée. Elle est calculée séparément pour les deux régimes d'assurance (régime général et des travailleurs indépendants) et pour une série d'états sociaux. Le décret royal du 10 juin 1996 donne la liste des états sociaux pour lesquels on calcule une dépense normative. Pour le régime des salariés, on distingue principalement : les titulaires indemnisables primaires (la majorité de la population), les pensionnés, les invalides ou les handicapés, les veufs, veuves et orphelins, et les personnes non protégées. Les effectifs par OA de chaque état social sont déterminés sur la base des données administratives de l'année pour laquelle on détermine le budget normatif.

⁴³ Pour prendre un exemple, une mutualité dont les dépenses réelles en 1997 s'élèvent à 30% du budget global et dont le budget normatif s'élève à 35% du budget global se verra attribuer : $0,9\% \cdot 30\% + 0,1\% \cdot 35\% = 30,5\%$ de l'objectif budgétaire global. Par construction, le budget global prévu est intégralement alloué.

⁴⁴ En 1997, pour l'année 1995 à titre d'exemple.

⁴⁵ Ce taux était valable de 1995 à 1997. En 1998 et 1999, il est de 20% et de 25% à partir de 2000.

Les dépenses normatives d'un assureur i , pour un état social r , C_{ir} sont égales à :

$$C_{ir} = n_r \left(\frac{C_r}{n_r} + COR_{ir} \right)$$

avec :

n_{ir} : nombre de bénéficiaires couverts par l'assureur i qui sont dans l'état s ;

n_s : nombre total de bénéficiaires dans l'état s ;

C_s : dépenses réelles globales de l'ensemble des OA pour l'état social s ;

COR_{ir} est un terme de correction pour l'état s et l'assureur i . Il reflète, par état social et pour chaque union de mutualités, l'ajustement en fonction d'une série de facteurs de risque (sexe, âge, situation socio-professionnelle, revenus, composition du ménage, taux d'urbanisation, taux de chômage et de mortalité)⁴⁶. L'estimation des paramètres de correction, sur la base de modèles économétriques, a été confiée à deux équipes universitaires pour l'année 1993. L'arrêté royal de 1996 prévoit que ces paramètres par état social seront ajustés chaque année pour tenir compte de l'augmentation des dépenses et de façon à ce que les termes de correction définitifs de chaque OA multipliés par l'effectif qui lui correspond donnent une somme égale à zéro. L'ajustement prend donc en compte la répartition des facteurs de risque à une date antérieure à celle pour laquelle la clé de répartition normative est calculée. L'arrêté royal prévoit cependant que les facteurs d'ajustement seront ajustés dès que possible si les paramètres recalculés diffèrent de façon significative de ceux qui figurent dans l'arrêté.

c) Quelques éléments de résultat

En 1995, toutes les unions de mutuelles ont obtenu un bonus, alors qu'en 1996, toutes ont dépassé l'objectif budgétaire de 6% environ⁴⁷. Cette uniformité souligne l'importance de choix du montant du budget macro-économique : les facteurs qui agissent uniformément sur les dépenses des unions ont une influence non négligeable sur l'équilibre global du système. En 1998, on observe pour la première fois une différenciation des primes entre les unions de mutualités (l'union de mutualité chrétienne

⁴⁶ Cette liste a été obtenue par recoupement à partir de Hermesse et Beeckmans et l'article d'Interface précédemment cité.

⁴⁷ Le malus est calculé sur la base de 15% de 2% (plafond).

propose une prime légèrement inférieure). Toutefois, la valeur absolue de la prime n'étant pas très élevée, ses concurrents doutent que ceci entraînera des transferts massifs d'assurés.

Comme aux Pays-Bas, la principale limite de la réforme réside dans le fait que les marges de manoeuvre des mutualités pour négocier avec les prestataires sont très limitées. Ainsi l'article d'Interface souligne que : *"si le mécanisme de répartition de l'objectif budgétaire...[est défini] de manière précise et technique, il n'en va pas de même ..de la mise à disposition des outils nécessaires à cette responsabilisation individuelle... [Les conventions entre prestataires de soins et organismes assureurs] valent pour l'ensemble des mutualités et il n'est pas possible pour les prestataires de débattre des tarifs individuellement avec un OA, dans le cadre de l'assurance obligatoire"* (pp.12).

2.2.3 L'Allemagne

Nous avons souligné les similitudes entre le système belge et celui des Pays-Bas. L'objectif de court terme visé par l'introduction de la concurrence entre les caisses publiques allemandes dans les années quatre-vingt-dix était de décloisonner un système dans lequel, pour des raisons historiques, elles se trouvaient dans des situations très différentes au regard des risques qu'elles couvraient et des ressources qu'elles pouvaient prélever. A terme, l'objectif visé par la concurrence est pourtant une meilleure gestion des risques.

a) Le contexte de la réforme

En Allemagne, les caisses publiques, qui couvrent 90% de la population, sont traditionnellement organisées sur une base professionnelle. Les personnes qui ne peuvent être affiliées à ce titre, se regroupent par défaut dans des caisses locales. Le système est très morcelé, puisqu'environ mille caisses œuvrent sur le territoire.

Avant 1995, chaque caisse prélève des cotisations proportionnelles au revenu en déterminant le taux qui lui permet d'atteindre l'équilibre financier. Les caisses sont donc responsabilisées financièrement, mais elles sont en position de quasi-monopole. Par construction, la structure des risques couverts varie fortement d'une caisse à l'autre ainsi que la capacité contributive (le revenu moyen) des adhérents. Les caisses locales cumulent les handicaps : le revenu moyen des assurés est inférieur au revenu moyen des assurés des autres caisses, et elles couvrent des proportions importantes de personnes

âgées, d'invalides... La réunification a renforcé ces inégalités : les taux de cotisations sont systématiquement plus élevés dans les nouveaux Länder que dans les anciens (Boccognano, et al., 1998, pp.132).

La réforme Seehofer, initiée en 1992, cherche en premier lieu à harmoniser le système et à introduire entre les caisses une forme de concurrence qui les incitera à l'efficacité. Les éléments de la réforme sont les suivants :

- toutes les caisses sont tenues de couvrir les mêmes biens et services et d'offrir les mêmes taux de couverture ;
- le symbole de l'instauration de la concurrence est l'introduction d'une plus grande liberté de choix pour les assurés depuis 1997. Ainsi, toute personne a maintenant le droit d'adhérer à la caisse de son choix. Cette liberté est pourtant incomplète, puisque les caisses professionnelles conservent la possibilité de ne pas offrir de contrats aux personnes qu'elles ne couvraient pas traditionnellement, ce qui ne correspond à aucune logique économique, mais relève plutôt d'un lobbying efficace de la part des caisses d'entreprise couvrant une population jeune à revenus élevés ;
- afin que les prix reflètent l'efficacité des caisses et non leur structure démographique et économique plus ou moins favorable, il existe maintenant un mécanisme de compensation des risques qui se traduit par des transferts financiers entre caisses, selon des critères d'âge, de sexe et de revenu des assurés.

b) L'ajustement des risques

Ce mécanisme a été élaboré sous la pression des caisses locales (Allgemeine Ortskrankenkassen, AOK) qui étaient très défavorisées par la situation antérieure, avec le soutien du Parti Social Démocrate (SPD) et des ministres sociaux des Länder.

Le mécanisme est original car il ne cherche pas seulement à compenser les caisses pour leur différence de structures de risque mais aussi pour celle résultant des variations de capacités contributives de leurs membres affiliés.

Dans un premier temps (Files et Murray, 1995), on estime la capacité financière de chaque caisse. Pour ce faire, on détermine par personne *couverte* (assurés et ayants droit) un revenu soumis à cotisation⁴⁸. On multiplie ce montant par une estimation du

⁴⁸ Pour un même niveau de revenu disponible, les caisses dans lesquelles le nombre d'ayant droits est très élevé ont donc une capacité contributive estimée plus faible.

taux de cotisation moyen sur la population (dépenses maladie de l'ensemble des caisses/revenu national soumis à cotisation). On détermine ainsi le montant total que la caisse prélèverait si elle avait un taux de cotisation "moyen".

Ensuite, on calcule un coût moyen par personne selon trois facteurs de risque : l'âge, le sexe et l'incapacité. On calcule alors une structure de coût de la caisse en fonction du nombre de personnes qui répondent à ces critères.

Si la capacité contributive est supérieure aux coûts standardisés, la caisse doit contribuer au financement du pool et vice versa. Les sommes versées et collectées sont déterminées ex ante : la caisse reste responsable de son équilibre financier mais elle peut prendre en compte dans le calcul de son taux de cotisation les compensations qu'elle devra payer ou qu'elle recevra.

c) Bilan provisoire

La compensation a été mise en œuvre pour la première fois en 1994 sur la base d'estimations sur données d'enquête de 1993. Les AOK, relayées par les médecins, soutiennent toutefois que cette compensation est insuffisante pour permettre une concurrence juste, qui serait fondée sur une utilisation efficace des ressources et non sur la sélection des risques. En effet, l'état de santé des personnes défavorisées couvertes par les AOK est bien en dessous de la moyenne nationale. La compensation ne prend en compte le revenu que sous l'angle de la capacité contributive, et celui non de la corrélation entre bas revenus et dépenses hospitalières élevées. Les AOK qui, entre autres, couvrent des proportions beaucoup plus élevées de diabétiques, souhaiteraient que des indicateurs de maladie chronique soient pris en compte dans le calcul des dépenses anticipées. Toutefois, même si elle est imparfaite, la compensation a conduit à un rapprochement significatif des taux de cotisation : l'écart de taux moyen en 1997 dans la partie occidentale du pays entre les caisses d'entreprises et les caisses locales était de 1,1 point contre 2,3 points en 1993. La concentration des caisses a aussi été très rapide. Les effets les plus néfastes de la segmentation ont donc été réduits.

De manière générale, la mise en concurrence des caisses n'a pas eu d'effet sur les coûts médicaux ni sur l'efficacité du système, en grande partie parce que la faible mobilité des assurés amoindrit la menace de sanction contre d'éventuels manquements à l'efficacité. Les assurés ont parfaitement le droit de changer de caisse, mais ils ne le font pas

facilement : ainsi, les caisses locales, dont les taux de cotisation sont parmi les plus élevés, n'ont vu leurs effectifs baisser que de 2%. L'autorisation de changer de caisse a apparemment plutôt été utilisée par les assurés qui déménagent pour rester dans la même caisse. Il semble toutefois que les bons risques (populations jeunes et à revenus élevés) soient davantage mobiles que les autres.

Enfin, l'introduction de la concurrence a incité les caisses les plus exposées (les AOK) à s'impliquer -de façon limitée - dans les expérimentations de type *managed care*. Leurs concurrentes, dont la structure de risques est plus favorable, sont plus réticentes à proposer des formules que leur clientèle traditionnelle, relativement aisée, pourrait boudier : les modèles de type *gatekeeper* ont relativement mauvaise presse en Allemagne parce qu'ils symbolisent une médecine bon marché et de mauvaise qualité (Boccognano, et al., 1998).

2.2.4 Israël

Les système israélien avant 1995 présente quelques similitudes avec le système suisse (non universel et concurrentiel), même si les cotisations sont proportionnelles au revenu (comme en Allemagne). La réforme a introduit l'assurance universelle et a mis à plat le système de financement en centralisant les ressources prélevées. Les caisses en concurrence sont désormais théoriquement rémunérées en fonction des risques qu'elles couvrent.

a) La réforme de 1995

En Israël, l'assurance maladie est traditionnellement servie par des caisses à but non lucratif contrôlées pas des syndicats et les individus peuvent théoriquement choisir leur caisse d'affiliation, comme en Belgique. La caisse dominante, qui dépend du principal syndicat ouvrier (70% du marché) fonctionne en *HMO staff model*⁴⁹. Trois caisses mineures se partagent le reste du marché et passent des contrats avec les groupes de médecins. Les caisses et les syndicats prélèvent des cotisations proportionnelles au revenu sur les employeurs et employés. Les taux sont en principe uniformes pour toutes les caisses mais le processus de financement global manque singulièrement de transparence.

⁴⁹ Les médecins sont salariés.

La réforme de l'assurance maladie israélienne de 1995 vise à répondre à une série de problèmes mis en lumière par une commission parlementaire quelques années auparavant (Commission Netanyahu 1990). L'assurance est volontaire et 4% de la population n'y a pas accès. Le problème est beaucoup plus prononcé pour les citoyens arabes - 12% non assurés - et les personnes résidant dans certaines régions. Certaines catégories d'immigrants sont aussi touchées. Par ailleurs, de nombreuses personnes n'ont pas réellement la possibilité de changer de caisse, soit parce qu'elles savent que leur candidature sera rejetée par une autre caisse - l'écroulement se développe au fur et à mesure que les coûts augmentent -, soit parce que dans leur région, seule la caisse dominante est présente. Celle-ci étant obligée statutairement d'accepter toute personne syndiquée qui en fait la demande, voit au cours des années 80, sa structure de risque se dégrader fortement (12% de personnes de plus de 65 ans versus 6% pour les autres caisses, beaucoup plus de malades chroniques). Comme, en outre, elle couvre des personnes de condition modeste, elle reçoit au total (à âge égal) 30% de moins que les autres caisses. Ses déficits sont financés au coup par coup par l'Etat qui s'évertue dans le même temps à transférer une part croissante du financement vers le secteur privé. Le manque de transparence généralisé du système a des conséquences concrètes pour les assurés. Leurs droits ne sont pas clairement définis et même s'ils sont couverts, ils se voient de plus en plus souvent obligés de payer directement des interventions lourdes.

La réforme de 1995 a donc deux objectifs principaux (Chernichovsky et Chinitz, 1995) : d'une part clarifier les responsabilités de chacun (afin d'inciter les caisses à une meilleure gestion) et d'autre part garantir à tous les citoyens, indépendamment de leurs ressources, un accès à un panier standard de soins servi par la caisse de leur choix. Les principaux éléments du texte de loi sont les suivants :

- l'assurance devient universelle ;
- elle couvre un panier de biens et services clairement définis ;
- chacun peut choisir la caisse d'affiliation ;
- les caisses n'ont pas le droit de refuser la couverture à un individu, ou d'exiger un complément de prime pour financer l'assurance légale. Les caisses doivent donc se faire concurrence sur la qualité des réseaux proposés.

b) Le financement prospectif des caisses

- les cotisations existantes sont regroupées en une cotisation santé unique prélevée de façon centralisée et l'Etat s'engage à financer le complément si nécessaire ;
- les fonds sont reversés aux caisses sous la forme d'une capitation. Celle-ci est cependant très grossière, car elle ne dépend que de l'âge (9 tranches). La pondération de chaque tranche d'âge est calculée par le ministère sur la base d'informations externes aux caisses⁵⁰. Des capitations égales au coût moyen du traitement sont versées aux caisses pour les assurés qui souffrent de quatre pathologies (dont l'IRC et l'hémophilie). Ce versement représente 3.3% du total versé en 1996.

Un bilan récent de la réforme (Rosen, 1998) indique que la concurrence entre caisses a augmenté. Certaines régions ou populations ont désormais plus de choix et la qualité des services s'est globalement améliorée. Parallèlement, il semble que les caisses développent des stratégies d'écrémage plus actives. L'allocation de ressources est incontestablement plus juste qu'auparavant mais toutes les caisses doivent faire face à des déficits importants et croissants, ce qui crée des problèmes politiques entre le ministère de la Santé et les caisses d'une part, qui estiment que le secteur de la santé est sous financé et le Trésor, qui considère pour sa part que les caisses n'ont pas déployé assez d'efforts pour être efficaces. Fin 1997, le Parlement envisage d'autoriser les caisses à introduire des copaiements⁵¹ pour les services du panier de base et à prélever une prime uniforme par assuré, les personnes défavorisées pouvant être exonérées totalement ou partiellement. Les tenants de cette réformes voient là un moyen d'instiller une concurrence limitée par les prix et d'augmenter l'autofinancement des caisses. Les détracteurs soulignaient le côté régressif de la taxe uniforme et le fait que les copaiements seraient plus lourds pour les malades chroniques. Après de nombreux débats, le principe de ces réformes a été approuvé mi-1998, mais les mesures concrètes proposées par les caisses doivent être approuvées par les plans et aucune n'a jusqu'à maintenant été acceptée.

⁵⁰ Données d'enquête (pour l'ambulatoire), données hospitalières, et comptes de la santé (pour la répartition entre types de soins par tranche d'âge). Les consommations pharmaceutiques ne sont pas prises en compte.

⁵¹ Certains copaiements étaient déjà prévus par la Loi mais celle-ci stipulait qu'ils ne devaient pas être augmentés et étendus.

Par ailleurs, le parlement a commandité une étude visant à évaluer dans quelle mesure la capitation devrait être améliorée. Le rapport (Habib et Rosen, 1998), après consultation de nombreux experts et évaluation la faisabilité avec les caisses elles-mêmes, recommande plusieurs améliorations marginales (par exemple, réactualisation plus fréquente de la capitation, inclusion de la pharmacie dans le calcul des pondérations par classe d'âge...) dont nous ne présenterons pas le détail. Quant à l'introduction de nouvelles variables, les conclusions sont les suivantes :

1. en ce qui concerne l'état de santé : les indicateurs proposés sont les maladies chroniques (directement ou par le biais des hospitalisations qu'elles entraînent) et/ou l'incapacité. Le choix des indicateurs devrait dépendre de la fiabilité des informations disponibles et la faisabilité n'est pour l'instant pas garantie. Il est aussi proposé que l'Etat prenne en charge une partie (50%) des dépenses individuelles au delà d'un plafond annuel. Cette approche pose aussi un problème car les caisses ne sont pas capables à l'heure actuelle de calculer un coût par personne et par an. Le rapport estime que la recherche visant à inclure l'état de santé dans la capitation doit être prioritaire ;
2. puisque les dépenses liées à la maternité sont prises en charges en dehors de l'assurance de base, il n'est donc pas jugé nécessaire d'inclure le sexe dans la capitation ;
3. l'introduction d'une variable "zone périphérique" ne fait pas l'objet d'un consensus. Les uns estiment qu'en périphérie les niveaux de consommation sont plus faibles et les coûts d'investissement plus difficiles à rentabiliser et qu'on pourrait, en augmentant la dotation, tenter de corriger ces effets. Les autres répliquent que la loi a déjà amélioré la qualité de l'offre dans les régions plus isolées et qu'augmenter la capitation n'est pas l'incitation la plus efficace à l'investissement. Par ailleurs, ils privilégient une capitation plus élevée là où la consommation l'est (débat entre financement des coûts / financement des besoins). En tout état de cause, les avis divergent sur la façon de mesurer "l'éloignement". Au total, le rapport recommande d'attendre quelques années avant de considérer de nouveau l'inclusion d'une telle variable ;
4. une variable indicatrice de la situation par rapport à l'emploi pourrait être incluse pour rendre compte de différences d'état de santé liées au niveau social (en particulier tant que l'état de santé n'est pas pris en compte en tant que tel). Toutefois, à l'heure

actuelle, le délai nécessaire à la validation de l'information est encore trop long pour envisager l'inclusion de la variable.

La redéfinition de la capitation apparaît en Israël (contrairement à la Suisse) comme un objectif prioritaire qui permettrait de continuer à corriger les inégalités de dotations et qui pourrait limiter les stratégies d'écrémage que la mise à plat du système et le renforcement de la responsabilité financière ont avivées. L'approche très pragmatique et progressive adoptée, qui s'inspire du modèle néerlandais, semble pouvoir amener à terme une amélioration du fonctionnement du système.

2.2.5 Medicare

Le dernier système que nous présentons est celui de Medicare. Depuis la fin des années quatre-vingt-dix, chaque assuré peut rester dans le système de paiement à l'acte ou se tourner vers un opérateur privé, de type HMO obligatoirement, qui lui proposera un réseau fermé de praticiens et une filière de soins organisée. En 1997, 14% de la population Medicare était couverte par un HMO. Medicare est administré par la *Health Care Financial Administration* (HCFA) qui, en ouvrant cette possibilité, souhaitait développer un mode de prise en charge et de gestion des soins plus efficace que celui du paiement à l'acte. La HCFA doit en contrepartie jouer le rôle du "sponsor" dévolu à l'employeur dans la concurrence gérée d'Enthoven : choisir les plans et décider du montant de la capitation qu'elle leur verse pour assurer et délivrer les soins aux personnes âgées. Avant 1997, une capitation était calculée par type de bénéficiaire (personne âgée, insuffisant rénal chronique ou invalide) et de soin (partie A et B) sur la base du coût historique moyen de prise en charge dans le système traditionnel de paiement à l'acte dans le comté, ajusté par tranche d'âge, selon le sexe et quelques variables de statut (selon que la personne réside dans une institution - maison de retraite, long séjour -, travaillant, bénéficiaire de Medicaid). 95% de ce "*Adjusted Average Per Capita Cost*" ou AAPCC était versé aux plans. Il est apparu au fil du temps que cette dotation était trop grossière :

- elle encourageait la sélection des risques par les assureurs ou leur donnait une rente trop forte du fait de l'auto-sélection dont ils bénéficiaient (cf. chapitre précédent).
- elle aboutissait à des situations paradoxales car les capitations entre comtés adjacents pouvaient différer jusqu'à 20% et pouvaient varier d'une année sur l'autre (en fonction

des coûts constatés), ce qui ne favorisait pas le développement harmonieux de l'offre par les HMO, particulièrement dans les régions rurales.

En 1997, plusieurs réformes ont été mises en place (Rapport au Congrès, Medicare Payment Policy, 1998). Tout d'abord, la capitation a été modifiée à compter de 1998 :

- une capitation minimum est garantie aux plans, ainsi qu'un taux d'augmentation minimal de 2% par an par comté ;
- la capitation n'est plus calculée désormais sur une base exclusivement locale : elle correspond, pour partie, à la moyenne nationale calculée sur les mêmes critères de risque et, pour partie, à la moyenne locale. En 1998, la proportion du coût local est de 90% et la pondération devrait atteindre 50% en 2003.

La modification du système de paiement doit être neutre pour la HCFA, qui doit déboursier les mêmes sommes que si l'ancien système avait été appliqué. Toutefois, les ajustements permettant de respecter cette contrainte ne sauraient conduire un comté en dessous du minimum garanti.

La loi de 1997 prévoit en outre que la capitation sera ajustée en fonction de l'état de santé à partir de 2000. Par ailleurs, la HCFA cherche à promouvoir une réforme qui contribuerait à renforcer la concurrence entre HMO et à limiter les bénéfices qu'ils tirent de l'écroulement ou de l'auto-sélection. Pour ce faire, la HCFA veut mettre en place une concurrence en deux étapes. En premier lieu, un mécanisme d'enchères permet de fixer le montant de la capitation. La HCFA demande aux opérateurs de fournir leurs prix de manière confidentielle et elle choisit ensuite le montant de la capitation à partir de l'information collectée. En second lieu, les opérateurs démarchent les assurés, en leur proposant obligatoirement le prix qu'ils ont révélé à la HCFA. Cependant, l'instauration d'une telle concurrence pose des problèmes pratiques de mise en place. Tout d'abord, il convient de signaler que les plans sont très réticents et tentent de faire renoncer la HCFA à son projet, ou bien d'en obtenir une version adoucie. Ils craignent en premier lieu qu'en les obligeant à révéler leurs véritables coûts, le mécanisme proposé ne les prive de leur profit (donc d'une forme de rente qu'ils prélèvent sur le système de soins). Ils avancent aussi l'argument selon lequel ce type de procédure à deux étapes augmente notablement les frais d'entrée sur le marché, ainsi que les risques commerciaux. Elle limite aussi la concurrence en introduisant une nouvelle barrière à l'entrée : seuls ceux qui sont passés par le quasi marché ont le droit de proposer leurs services aux assurés. Donc, une fois

l'enchère réalisée, on ne peut autoriser l'entrée de nouveaux plans jusqu'à la prochaine renégociation. Si on agrandit trop l'intervalle entre deux marchés internes, on interdit l'entrée de compétiteurs potentiellement innovants et on accroît la rente de ceux qui sont présents. Si on réduit trop cette période, on augmente notablement les coûts et les risques pour les opérateurs.

Au total, l'expérience de délégation de gestion de Medicare par la HCFA reste connue par les effets de sélection ou de segmentation spontanée qu'elle a suscités pour les personnes concernées. La HCFA cherche à corriger cet effet en limitant la rente dont bénéficient les HMOs qui reçoivent une capitation jugée trop généreuse et en l'ajustant suffisamment bien pour que des personnes qui présenteraient des niveaux de risque élevés puissent, elles aussi, bénéficier des avantages d'une gestion plus globale et plus efficace de leur prise en charge.

2.2.6 Principales leçons des systèmes où la concurrence est introduite

Les cinq systèmes, dont nous venons de détailler les réformes, avaient, à l'origine, des modes de fonctionnement assez différents, comme le rappelle le tableau ci-dessous.

Tableau 2 : Caractéristiques principales des systèmes avant les réformes

	Couverture	Structure du marché	Concurrence	Responsabilité financière
Pays-Bas	universelle*	Monopole	Non	Non
Belgique	universelle*	6 unions nationales	Oui	Non
Allemagne	universelle*	1000 caisses	Non	Oui
Israël	non universelle	4 caisses	Oui	Non
Medicare	universelle*	Monopole	Non	Non

* Pour la population du régime concerné (cf Chapitre 7)

En premier lieu, notons que les réformes ont indéniablement été porteuses d'une harmonisation, puisque, désormais, dans les cinq systèmes, la couverture est universelle, les organismes qui la gèrent sont en concurrence⁵² et sont responsabilisés financièrement de façon explicite et uniforme au sein de chaque système. En outre, à l'intérieur de chaque système, les paniers de biens et services couverts sont identiques.

Au delà de ce premier constat, un certain nombre de points méritent d'être soulignés.

1. S'il n'est pas possible d'estimer l'impact des réformes sur l'équité de tous les systèmes, on peut remarquer qu'aucune ne vise à réduire la solidarité au profit de la responsabilité individuelle ou à remettre en cause ce qui reste considéré comme un acquis social (accès à l'assurance et financement redistributif). Plus précisément :
 - on constate, tout d'abord, que le nombre de personnes couvertes ne diminue dans aucun pays malgré l'introduction de mécanismes concurrentiels. En Israël, il augmente même, mais comme, dans le même temps, le panier de biens et services couvert par l'assurance sociale diminue pour une partie de la population, il est difficile d'estimer l'impact global en terme de couverture. En revanche, désormais, dans ce pays, toute personne est tenue de s'assurer, ce qui signifie en particulier que les bas risques ou les hauts revenus n'ont plus la possibilité d'échapper à un financement solidaire de l'assurance sociale. Dans les autres systèmes, l'assurance demeure obligatoire et les entités en concurrence n'ont pas le droit de refuser de couvrir une personne qui en fait la demande, ce qui permet de garantir une couverture universelle ;
 - avant les réformes, les systèmes étaient déjà essentiellement financés par des cotisations (Pays-Bas, Israël, Allemagne, Belgique) ou des impôts (Medicare) proportionnels au revenu. Ce principe est aussi maintenu. En Allemagne, la concentration des caisses et le mécanisme de compensation - qui vise explicitement à compenser les différences de capacités contributives - ont plutôt amélioré l'équité du financement par la réduction des écarts entre les taux de cotisations. Les réformes ont aussi conduit à une uniformisation des taux de cotisation en Israël et en Belgique ;

⁵² Pour Medicare, cette concurrence se déroule entre les HMOs qui participent à la délégations de gestion. La HCFA n'est pas à proprement parler en concurrence avec les HMOs.

- les Pays-Bas et la Belgique ont ouvert aux organismes en concurrence la possibilité de prélever des primes uniformes sur tous les assurés pour équilibrer leurs budgets. Un tel mode de financement est, par construction, inéquitable. Dans l'esprit initial de la réforme aux Pays-Bas, il était prévu que cette prime représenterait 25% de la dépense, mesure qui a été abandonnée pour préserver l'équité du système. A l'heure actuelle, dans les deux pays, la prime reste très symbolique et constitue, en théorie, un signal de l'efficacité productive relative d'un organisme par rapport à un autre.
2. L'introduction d'une possibilité pour les assurés de choisir l'opérateur gestionnaire de leur couverture a eu un impact très différent selon la "culture" d'origine.

Aux Etats-Unis, où les assurés sont habitués à choisir des contrats et à changer d'assureur, ils semblent utiliser pleinement les possibilités qui leur sont offertes. Plusieurs exemples en témoignent. On constate ainsi que les assurés passent, en fonction de leur état de santé, d'un système à l'autre : quand ils sont en bonne santé, ils choisissent le système HMO et ils reviennent vers le système traditionnel si leur état de santé se dégrade (système dit des "portes de saloon", Morgan et al. 1997⁵³). Dans le même ordre d'idée, les HMOs proposant une période d'adhésion mensuelle, se sont trouvés confrontés au problème du "*hit and run*" : une proportion importante de la clientèle change tous les mois d'assureur, soit pour bénéficier d'une promotion ponctuelle, soit, plus vraisemblablement, pour contourner les plafonds de dépense sur les médicaments (l'assureur les remboursant dans les limites d'un plafond annuel). Quand l'assuré a atteint ce plafond, il se tourne vers un nouveau HMO⁵⁴.

A l'inverse, dans les autres pays, la réforme n'a semble-t-il pas incité les personnes à changer d'organisme gestionnaire. En Israël et en Belgique, une forme de concurrence existait en théorie déjà, mais le choix d'une caisse ou d'une mutualité était, et reste, plus l'expression d'un choix politique, d'une appartenance culturelle que d'un arbitrage économique. Aux Pays-Bas et en Allemagne, les assurés ont plus mis à profit la possibilité qui leur était offerte pour rester dans la même caisse, là où, avant la réforme, ils auraient été contraints de changer (à cause d'un déménagement ou d'un changement d'emploi). Les mécanismes d'incitation par les prix n'ont semble-t-il pas eu un impact très élevé, ce qui n'est pas étonnant en Belgique et aux Pays-Bas où les

⁵³ La question demeure entière de savoir s'ils sont incités par les HMO à passer de l'un à l'autre.

⁵⁴ Entretien avec Arlen Collins, consultant en Managed Care, 1998.

primes sont assez faibles et encore peu différenciées, mais qui est plus surprenant en Allemagne, où les taux de cotisation peuvent varier d'une caisse à l'autre. En tout état de cause, il est probablement trop tôt pour que les habitudes aient changé et il serait intéressant de suivre l'évolution des comportements dans les prochaines années.

3. Pour ce qui est du comportement des OGR en concurrence, il faut, là encore, distinguer les Etats-Unis des autres systèmes.

Les HMOs auxquels on délègue la gestion de Medicare sont familiers avec le fonctionnement d'un marché concurrentiel et, par définition, avec la gestion du risque. Le versement par la HCFA d'une capitation trop peu ajustée au risque leur a permis de réaliser des profits d'autant plus élevés qu'il leur suffisait d'intégrer ces nouveaux patients dans les réseaux déjà existants et qu'il ne leur a pas été nécessaire de réaliser de nouveaux investissements pour améliorer la gestion des risques.

Dans les autres pays, exception faite d'Israël où le fonctionnement en réseau est implanté dans les moeurs, mais où les problèmes auxquels le système est confronté sont de nature plus politique et dont le mode de gestion n'est pas à l'heure actuelle bien défini (Rosen, 1998), l'évolution semble très lente, pour deux grands types de raisons.

- Tout d'abord, aux Pays-Bas et en Belgique, mais aussi en Allemagne, les caisses doivent opérer un véritable changement culturel. Dans les deux premiers pays, elles n'avaient auparavant jamais été responsabilisées et leur rôle se limitait quasiment à la liquidation des prestations. En Allemagne, où la concurrence n'existait pas, la situation était comparable. Elles doivent donc progressivement constituer des systèmes d'information et se doter des outils de gestion du risque, développer une culture de service... Une telle évolution ne peut s'opérer que progressivement et l'on en n'observe pour l'instant que les prémices (accueil d'une plus grande convivialité...).
- Mais surtout, les conditions réglementaires qui pourraient permettre aux organismes de gérer le risque ne sont pas réunies. Dans l'ensemble de ces pays, les marges de manoeuvre des OGR sont restées très limitées et il semble que les Etats répugnent à abandonner des outils de maîtrise des dépenses éprouvés au bénéfice d'une gestion purement marchande. Ils doivent aussi faire face à la résistance des prestataires de soins. On peut cependant noter qu'aucune réforme n'a été stoppée et que certains pays

envisagent d'aller plus loin. Ainsi, en Allemagne⁵⁵, la Ministre Andrea Fisher souhaite confier d'ici 2003, la gestion des investissements hospitaliers aux caisses et non plus aux régions. Les caisses devraient aussi avoir le droit de passer des contrats avec des médecins ou des groupes de médecins.

Ceci dit, il semble encore trop tôt pour tirer un bilan définitif des réformes entreprises dans ces pays qui dépasse le constat simple d'une harmonisation des modalités de concurrence.

⁵⁵ "Réforme de la santé : les caisses Allemandes devraient sortir gagnantes" *Espace Social Européen*, 18-24 juin 1999, pp. 12.

CONCLUSION DU CHAPITRE ET DE LA PARTIE

Au terme de ce chapitre, il semble donc évident que les modes de gestion de la concurrence convergent, comme le tableau 3 le souligne (page suivante). Même si les réformes ou les évolutions constatées répondent à des objectifs prioritaires différents, on peut estimer que toutes prennent en compte le fait que *"the only time risk assessment and risk adjustment are not necessary is when there is a unitary plan enrolling the entire eligible population"*⁵⁶ (Luft, 1995, pp. 23).

Certaines différences fondamentales entre les systèmes traditionnellement concurrentiels ou privés et les systèmes plus socialisés ne sont pourtant pas gommées :

- le mode de prélèvement auprès de l'assuré, sous forme de prime relativement uniforme d'un côté, sous forme de contribution proportionnelle au revenu de l'autre ;
- le degré de contrainte sur les participants. Dans la quasi-totalité des systèmes, assureurs et assurés doivent respectivement participer à l'ajustement au risque et s'assurer. Pour obtenir que la concurrence porte sur l'efficacité productive, on s'éloigne du modèle dans lequel l'assurance est volontaire, au sens où nous l'avons entendu dans le chapitre 7. Deux pays se distinguent à cet égard. Aux Etats-Unis, les assurés ne sont pas tenus de s'assurer, même s'il est toujours plus avantageux de participer au système mis en place par leur employeur que de s'assurer à titre individuel⁵⁷. Les assureurs, pour leur part, choisissent d'offrir leurs services au groupe d'assurés constitué, mais doivent, dans ce cas, accepter l'ajustement au risque. Il est intéressant de rappeler que la réforme Clinton prévoyait, outre l'universalisation de l'assurance, l'introduction d'un mécanisme de péréquation des risques entre les OGR. L'autre exception est l'Irlande, où tout assureur souhaitant entrer sur le marché devra participer à la péréquation mais où les personnes ne sont pas contraintes de s'assurer. Il faut rappeler qu'en Irlande, l'assurance ne concerne pas l'intégralité du risque et que les plus pauvres bénéficient de l'accès gratuit aux soins.

⁵⁶ La seule situation dans laquelle l'ajustement au risque ou son évaluation n'est pas nécessaire, est celle où un opérateur unique couvre l'intégralité d'une population définie.

⁵⁷ Il peuvent cependant préférer ne pas s'assurer ou bénéficier du contrats d'un conjoint en tant qu'ayant droit.

Tableau 3 : Principales caractéristiques des huit systèmes étudiés

	Objectif principal de la réforme	Droits et obligations des assurés	Obligations des assureurs	Mode de prélèvement des ressources	Financement des OGR
Allemagne	Réduction des inégalités et du morcellement A terme : amélioration de l'efficacité productive	Assurance obligatoire Choix de la caisse	Obligation d'accepter en assurance et de renouveler le contrat Pas d'exclusion des maladies déclarées	Cotisation proportionnelle au revenu qui dépend de la caisse	Cotisation perçue directement +/- compensation
Belgique	Amélioration de l'efficacité productive	Assurance obligatoire Choix de la caisse	Obligation d'accepter en assurance et de renouveler le contrat Pas d'exclusion des maladies déclarées	Cotisation proportionnelle au revenu uniforme	Capitation ajustée au risque partielle Financement rétrospectifs
Israël	Universalisation de l'assurance Amélioration de l'efficacité productive	Assurance obligatoire Choix de la caisse	Obligation d'accepter en assurance et de renouveler le contrat Pas d'exclusion des maladies déclarées	Cotisation proportionnelle au revenu uniforme	Capitation ajustée au risque partielle
Medicare	Amélioration de l'efficacité productive	Assurance obligatoire Choix entre le système traditionnel ou un HMO Choix du HMO	Obligation d'accepter en assurance et de renouveler le contrat Pas d'exclusion des maladies déclarées	Impôts et cotisation uniforme pour la partie B, indépendamment du choix du mode de gestion de la couverture	Capitation ajustée au risque
Pays-Bas	Amélioration de l'efficacité productive	Assurance obligatoire Choix de la caisse	Obligation d'accepter en assurance et de renouveler le contrat Pas d'exclusion des maladies déclarées	Cotisation proportionnelle au revenu uniforme Prime uniforme par caisse	Capitation ajustée au risque partielle Financement rétrospectifs
Irlande	Contrainte exogène (Europe) Préserver le mode de tarification équitable entre hauts et bas risques A terme : amélioration de l'efficacité productive	Assurance facultative Choix de l'assureur	Obligation d'accepter en assurance et de renouveler le contrat Réglementation des conditions d'exclusion des maladies déclarées	Prime Community Rating	Primes collectées +/- ajustement
Suisse	Universalisation de l'assurance Amélioration de l'efficacité productive	Assurance obligatoire Choix de la caisse	Obligation d'accepter en assurance et de renouveler le contrat Pas d'exclusion des maladies déclarées	Prime Community Rating	Primes collectées +/- ajustement
USA centrales d'achat	Amélioration de l'efficacité productive	Assurance facultative Choix de l'assureur	Obligation d'accepter en assurance et de renouveler le contrat Pas d'exclusion des maladies déclarées	Prime Community Rating	Primes collectées +/- ajustement

Toutefois, il ne serait pas rigoureux de déduire de cette analyse que les systèmes de santé eux-mêmes convergent. Comme le souligne Polton : "*la convergence des outils de régulation ne signifie pas la convergence des systèmes, qui demeurent très spécifiques en termes d'institutions et de jeux d'acteurs*"⁵⁸.

Au delà de ce constat, illustré dans la seconde section de ce chapitre, il faut rappeler qu'un élément fondamental de différenciation reste le panier de biens et services (uniformisé) sur lequel porte la concurrence "tempérée". Le chapitre 7 nous a conduit à décrire les différences entre les systèmes à cet égard. On peut prolonger l'analyse en montrant en quoi le choix de ce panier a un impact considérable sur la dynamique globale du système par rapport à la problématique qui nous intéresse ici.

1. Tout d'abord, la nature des biens et des services concernés par la concurrence tempérée aura un impact sur les modalités de recherche de l'efficacité productive, voire sur le mode de prise en charge des patients. Si cette concurrence ne portait que sur les dépenses hospitalières, l'OGR aurait intérêt à faire pression sur les hôpitaux, pour diminuer par exemple des durées de séjour, mais aussi à reporter la prise en charge des problèmes de santé des personnes couvertes vers la ville. Ceci n'est pas nécessairement synonyme d'une diminution de la qualité, si tant est que cette prise en charge est organisée. Or, dans une telle configuration, l'OGR n'aurait pas nécessairement les moyens ou la motivation suffisante pour le faire. Plus généralement, l'amélioration de l'efficacité globale passe souvent par la substitution de services et la réorganisation de la prise en charge. Il faut donc éviter d'instaurer des barrières de nature économique à cette substitution et le choix du panier couvert est à cet égard central. Cette préoccupation fait écho à un débat français ravivé à l'occasion des possibilités ouvertes par le plan Juppé de lancer des expérimentations dans le domaine des réseaux de soins : la non fongibilité des enveloppes de dépenses constitue une barrière importante à leur mise en œuvre⁵⁹.

⁵⁸ D. Polton, 1999, *la convergence des systèmes de santé*, Editorial de la lettre du collège des économistes de la santé.

⁵⁹ Les budgets hospitaliers, ceux des cliniques et les honoraires de différentes professions sont soumis à des enveloppes globales, sans possibilité de transfert des unes vers les autres.

2. Ensuite, la conclusion selon laquelle la forme de concurrence étudiée peut permettre de maintenir l'équité du système, voire de l'augmenter, dépend implicitement de l'étendue du panier de biens et services couverts et de l'articulation du système réglementé avec l'assurance volontaire.

Plus le panier sur lequel porte la concurrence tempérée est restreint, plus l'assurance volontaire joue un rôle important dans l'accès aux soins. La sélection dans ces différentes acceptions n'étant pas tempérée sur ce marché privé, on peut retomber dans un système générateur d'exclusion. En Israël, le choix du panier de biens sur lequel porte la concurrence tempérée est controversé pour cette raison.

3. Il reste enfin à souligner que l'assurance volontaire peut constituer un frein au bon fonctionnement de la concurrence tempérée. Nous avons souligné qu'une des raisons pour lesquelles les personnes hésitent à changer de caisse aux Pays-Bas, est qu'elles craignent de ne pas pouvoir souscrire une assurance "du troisième étage" auprès d'une nouvelle caisse. Si les mêmes entités gèrent l'assurance "de base" et l'assurance "volontaire", les assureurs peuvent utiliser cette dernière pour améliorer leur profil de risque et pour sélectionner les assurés. En refusant l'assurance volontaire à une personne, ils sont moins susceptibles de l'attirer pour l'assurance de base. Si l'ajustement au risque sur la couverture de base était parfait, ce qui n'est jamais le cas, les assureurs n'auraient qu'un intérêt limité à utiliser la sélection car ils n'en tireraient de bénéfice que dans le cadre de l'assurance volontaire. En réalité, plus le champ de l'assurance volontaire est large par rapport à celui de l'assurance "de base", plus cette incitation est forte. Elle disparaît, toutefois, si, comme en Allemagne, les caisses chargées de la gestion de l'assurance de base ne sont pas autorisées à couvrir les biens exclus du panier.

Enthoven (1994, pp. 1417) estime que les ventes jointes d'assurance devraient être interdites et que le panier de biens et services, objet de la concurrence tempérée, devrait être suffisamment large pour qu'il n'y ait pas de marché d'assurance volontaire⁶⁰.

⁶⁰ *"Tie-in sales should not be allowed. I recommend a fairly comprehensive package of services ... so that there will be no market for supplemental coverage for covered services"*

Ceci permet notamment d'instaurer des copaiements qui limitent le risque moral du côté des consommateurs, incitation qui ne fonctionne que si ces copaiements ne sont pas pris en charge par une assurance complémentaire⁶¹.

La dernière interrogation transversale que l'on peut soulever est celle de l'évaluation coût-bénéfice de l'ensemble de ces réformes. Il n'est pas question ici d'y apporter une réponse puisque d'une part, ces réformes ne sont pas abouties et que d'autre part, quand bien même un tel exercice serait possible, il dépasserait largement le cadre de ce travail. Néanmoins, il est apparu clairement que la réglementation de la concurrence nécessitait a minima de mettre en œuvre des mécanismes d'ajustement au risque complexes et coûteux. La gestion du risque par OGR nécessite de procéder à des investissements spécifiques (systèmes d'information, mise en place de nouvelles compétences...). Les gains en productivité sont-ils suffisant pour compenser ces coûts ? Plus généralement, le présupposé théorique selon lequel la gestion des risques est mise en œuvre de façon plus efficace par des organismes en concurrence reste très controversé (Boccognano et al., 1998).

La conclusion de ce travail nous conduira à mettre en lumière l'ensemble des questions qui restent ouvertes et qui sont autant de voies à explorer.

⁶¹ Fonction 3 dans le chapitre 7.

CONCLUSION GENERALE

CONCLUSION GENERALE

La théorie économique suggère que la concurrence, pour une distribution initiale des ressources donnée, permet d'obtenir l'efficacité de l'allocation et de la production des biens et des services échangés dans l'économie. Partant de ce postulat, nous avons cherché à déterminer dans ce travail si la concurrence entre assureurs sur la couverture des risques en santé était spontanément porteuse d'efficacité.

Nous avons entamé cette démarche en tentant d'écarter d'emblée la position selon laquelle l'assurance "santé", portant sur un bien sans prix, ne saurait relever d'un mécanisme privé de couverture. En appliquant à la santé les principes assurantiels issus de la théorie, nous avons montré que le service rendu par les assureurs dans ce domaine ne diffère pas fondamentalement du service d'assurance pour d'autres types de dommages. Nous avons dans le même temps souligné que l'assureur semble être en position d'intervenir dans la gestion des soins et d'améliorer ainsi l'efficacité productive du système de soins. Comme la maîtrise des dépenses de santé est un enjeu central dans les pays développés, cette première analyse laisse à penser que la concurrence entre assureurs pourrait contribuer à la réalisation de cet objectif.

Puis, nous nous sommes tournés vers l'étude du fonctionnement du marché concurrentiel du service considéré. Nous avons rappelé qu'en présence d'antisélection, la concurrence, loin d'améliorer le fonctionnement du marché de l'assurance, est susceptible au contraire d'en menacer l'existence. Une analyse théorique plus poussée nous a cependant permis de montrer qu'une intervention calibrée de l'Etat permet de surmonter cette difficulté. En procédant à une analyse de la littérature empirique pour déterminer si l'antisélection est effectivement constituée sur le marché de l'assurance santé, nous avons été amenés à montrer que le véritable frein à l'efficacité est la sélection, par le risque et par le revenu. En effet, celle-ci a pour première conséquence d'exclure de l'assurance des personnes ayant un revenu modeste et présentant des niveaux de risque élevés. Quelles que soient les raisons pour lesquelles ces personnes ne sont pas couvertes, les assureurs n'ont aucun intérêt à gérer leur risque et le champ sur lequel portent les incitations à l'efficacité productive est restreint. Ceci est d'autant plus préoccupant que sur un marché concurrentiel, sélection et efficacité productive

constituent une alternative. Il est même probable que, sur un marché non régulé, l'assureur mis face à sa responsabilité financière a plus intérêt à procéder à une sélection différentielle des risques qu'à tenter de les gérer.

Nous avons alors montré, dans la troisième partie, que la gestion de la sélection, au nom de l'efficacité productive, est une préoccupation commune à de nombreuses réformes entreprises ces dernières années, visant à organiser la concurrence entre des assureurs ou des caisses publiques d'assurance maladie. Plus encore, nous avons constaté une convergence dans l'organisation de cette concurrence, qui ne porte pas tant sur la couverture du risque que sur sa gestion, qui est porteuse d'une certaine forme d'équité et qui est d'autant plus efficace qu'elle concerne l'ensemble de la population. Obtenir cette concurrence tempérée nécessite, en fait, de limiter l'efficacité allocative, autrement dit la liberté des consommateurs de se couvrir et de choisir leur niveau de couverture. Il est aussi probable que la gestion du risque elle-même passe par un encadrement des choix des assurés quant à la façon dont ils seront pris en charge par le système de soins (choix des points d'entrée, des prestataires, des procédures,...).

En définitive, ce travail permet de relativiser l'opposition entre équité et efficacité qui a dominé les débats depuis les années soixante-dix, pour insister sur la nécessité de réaliser un arbitrage entre efficacité allocative et efficacité productive.

Pour autant, les évolutions et les réformes étudiées sont loin d'être abouties. Leurs promoteurs et leurs observateurs soulignent avant tout la nécessité de parfaire l'outil central de gestion de cette concurrence qu'est l'ajustement au risque. Les questions qui restent ouvertes sont d'ordre empirique (élaboration et choix des critères d'ajustement...) mais aussi technique (configuration des systèmes d'information, mise en place des transferts...). Dans une perspective plus théorique, alors que l'étude de l'antisélection donne lieu à des raffinements dont les implications pratiques sont parfois difficiles à saisir (Rothschild et Stiglitz, 1997), il semble que la réflexion sur la sélection demeure assez embryonnaire. Nous avons souligné que les deux phénomènes, simultanément présents sur le marché de l'assurance santé, entretiennent des liens complexes. Ils peuvent cependant être clairement séparés sur le plan conceptuel et mériteraient de l'être dans l'analyse théorique. L'étude de la sélection pourrait être approfondie en mobilisant, d'une part, la théorie des jeux (déjà utilisée dans l'analyse de l'antisélection) et, d'autre part, les outils de l'économie industrielle qui propose une analyse des incitations que

l'Etat doit mettre en œuvre dans la délégation de gestion. Les récents travaux de Gasmi F., Laffont J.J. et Sharkey W. (1998, par exemple) sur concurrence et obligation de service universel sont particulièrement intéressants à cet égard.

Au delà de ces éléments, qui s'inscrivent naturellement dans le prolongement direct de notre travail, il nous semble que d'autres questions restent ouvertes.

En premier lieu, le modèle de référence implicite en matière de gestion des risques est le *managed care* qui s'est développé sur le marché de l'assurance aux Etats-Unis. L'objet principal de notre étude étant l'assurance, nous avons naturellement adopté cette approche. Il convient à ce stade de rappeler que cette fonction n'est pas nécessairement remplie par des assureurs.

Par exemple, les Britanniques, durant les années quatre-vingt-dix, se sont tournés vers un modèle radicalement différent qui faisait du médecin généraliste le pivot du système. Les généralistes gestionnaires de budget (*General Practitioners Fundholders*, GPFH) organisaient l'ensemble de la prise en charge du patient, sur laquelle ils étaient responsabilisés financièrement, jouant donc bien le rôle "d'entrepreneurs de soins". Cette volonté de confier la gestion du risque à des producteurs de soins a été confirmée par la suite, dans le cadre du *Total Purchasing* (achat total) : le soin d'acheter l'ensemble des biens et services médicaux pour une population donnée a été confié à des associations de cabinets de groupe.

Dans le même ordre d'idée, aux Etats-Unis, on observe une autonomisation de plus en plus grande des prestataires de soins ou, du moins, des groupes médico-hospitaliers par rapport aux assureurs. Ils proposent directement leurs services aux employeurs et assument, d'une certaine façon, la fonction d'assurance, en plus de celle de production et de gestion des soins. Cette dernière étant au cœur des réformes des systèmes d'assurance maladie que nous avons étudiées et non la fonction d'assurance stricto sensu, il semble important de s'interroger sur la forme institutionnelle qui serait la plus efficace pour la mettre en œuvre.

Enfin, l'expérience des différents pays conduit à s'interroger sur les éléments de la concurrence qui sont véritablement porteurs d'efficacité. Nous avons souligné que la responsabilisation financière est un mécanisme d'incitation central. Dans le même temps, les arbitrages de la part des assurés qui sont libres de changer de caisse d'assurance restent, dans certains pays, assez timides.

En poussant le raisonnement plus loin, on peut se demander si des formes de concurrence fictive par comparaison ou même d'enchères (Mougeot, 1999) pour servir des groupes d'individus ne pourraient pas permettre d'améliorer l'efficacité productive sans que plusieurs institutions soient en concurrence directe pour servir la même population (et donc en enrayant les possibilités de sélection).

Au Royaume-Uni, ainsi, depuis avril 1999, des GPFH et les expériences de *Total Purchasing* ont disparu au profit d'un modèle plus coopératif. Au terme de la nouvelle réforme, des entités appelées *Primary Care Groups*, couvrant des populations disjointes, achèteront les soins pour cette population, dans le cadre d'une enveloppe globale ajustée au risque qui leur sera confiée et sur laquelle elles seront responsabilisées financièrement. Cette nouvelle réforme pose des questions de nature économique : la concurrence multiplie les transactions, ce qui a un coût, tout comme la gestion d'une concurrence qui évite la sélection. Une solution basée sur des incitations financières strictes, sans véritable concurrence, ne pourrait-elle pas s'avérer plus efficace ?

En continuant à suivre les réformes en cours et les évolutions des différents pays et en les analysant à la lumière de la théorie économique, il devrait être possible d'apporter des éléments de réponse à ces questions.

L'analyse développée au fil de cette thèse alimente enfin le débat français dans lequel l'opposition entre partisans et détracteurs de la concurrence reste souvent empreinte d'idéologie et semble parfois manichéenne. Elle contribue aussi, nous l'espérons, à souligner que le passage vers un système concurrentiel, si une telle option devrait être choisie, nécessiterait de modifier fondamentalement les modes de régulation actuels et de se doter d'outils particuliers, dont certains restent encore à inventer.

BIBLIOGRAPHIE
TABLE DES MATIERES

BIBLIOGRAPHIE

- Akerlof G. A. (1970) "Market for Lemons: Quality Uncertainty and Market Mechanism". *Quarterly Journal of Economics* 84 (3): 488-500.
- Arrow K. J. (1953) "Le rôle des valeurs boursières dans la répartition la meilleure des risques". *Econométrie* 11 : 41-47.
- Arrow K. J. (1963) "Uncertainty and the welfare economics of medical care". *American Economic Review* LIII (5): 941-973.
- Arrow K. J. (1964) "The role of securities in the optimal allocation for risk bearing". *Review of Economic Studies* 31 : 91-96.
- Arrow K. J. (1971) *Essays in the Theory of Risk-Bearing*. Markham, Chicago.
- Arrow K. J. (1996) "The theory of risk-bearing: Small and great risks". *Journal of Risk and Uncertainty* 12 : 103-111.
- Arrow K. J., Lind R. C. (1970) "Uncertainty and the evaluation of public investment decision". *American Economic Review* 60 : 364-378.
- Ash A., Porell F., Gruenberg L., et al. (1989) "Adjusting medicare capitation payments using prior hospitalization data". *Health Care Financing Review* 10 (4): 17-29.
- Atkinson A. B. (1995) "Social Insurance"(eds), *Incomes and the Welfare State*. Cambridge University Press, Cambridge.
- Altenstetter C., Bjorkman J. W. (1997) Ed., *Health Policy Reform, National Variation and Globalization*. Macmillan Press Ltd., London.
- Bague-Forst A. (1995) "Innovation et assurance santé". *Risques* 21 (Janv-Mars): 95-102.
- Bakker F., Van Vliet R. (1993) "The effect of deductibles on premiums in health insurance. A case study on prescription drugs". *2d European Workshop on Econometrics and Health Economics*, Center for Health Economics, University of York.
- Baxter R. J., Mechanic R. E. (1997) "The Status Of Local Health Care Safety Nets". *Health Affairs* 16 (4): 7-22.
- Beck K., Zweifel P. (1995) "Cream skinning in Deregulated Social Health Insurance: Evidence from Switzerland". *Third European Conference on Health Economics*, Stockholm.
- Becker G. (1965) "A theory of the allocation of time". *Economic Journal*, Sept. .
- Benzeval M., Judge K., Whitehead M., Eds. (1995), *Tackling Inequalities in Health: An agenda for action*.
- Berk M. L., Schur C. L. (1998) "Access to Care : How Much Difference does Medicaid Make ?". *Health Affairs* 17 (3): 169-180.
- Bertko J., Hunt S. (1998) "Case Study: The Health Insurance Plan of California". *Inquiry* 35 (Summer): 148-153.
- Besley T., Gouveia M. (1994) "Alternative systems of health care provision". *Economic Policy* October : 200-258.
- Blanchet D. (1996) "La référence assurantielle en matière de protection sociale : apports et limites". *Economie et Statistique* 291-292 (1): 33-45.
- Blomqvist A. (1997) "Optimal non-linear health insurance". *Journal of Health Economics* 16 : 303-321.
- Boccognano A., Couffinhal A., Grignon M., Mahieu R., Polton D. (1998) *Mise en concurrence des assurances dans le domaine de la santé. Théorie et bilan des expériences étrangères*. CREDES,
- Borch K. (1962) "Equilibrium in a reinsurance market". *Econometrica* 30 (3): 424-444.
- Borch K. H. (1990) *Economics of insurance*.
- Boulard J. C. (1998) *Pour une couverture maladie universelle de base et complémentaire*. Aout.

- Bowen B. (1995) "The Practice of Risk Adjustment". *Inquiry* 32 (Spring): 33-40.
- Browne M. J. (1992) "Evidence of Adverse Selection in the Individual Health Insurance Market". *Journal of Risk and Insurance* 59 : 13-33.
- Browne M. J., I. D. H. (1993) "Information Asymmetries and Adverse Selection in the Market for Individual Medical Expense Insurance". *Journal of Risk and Insurance* 60 (2): 300-312.
- Calcul de la responsabilisation financière des organismes assureurs pour l'exercice 1995*. (1997) Institut National d'Assurance maladie et d'Invalidité, Bruxelles, Note CSS 97/8.
- Cartwright W., Hu T., Huang L. (1989) "Demand for Medigap Insurance by the Elderly: A Micro-Simulation Analysis". *Applied Economics* 21 : 1325-1339.
- Cass D., Chichilnisky G., Wu H.-M. (1991) *Individual risk and mutual insurance: a reformulation*. Columbia University, November.
- Cass D., Chichilnisky G., Wu H.-M. (1996) "Individual risk and mutual insurance". *Econometrica* 64 (2): 333-341.
- Caussat L. (1994) "Sécurité sociale : pour l'assurance". *Droit Social* 11 : 902-907.
- Chakraborty G., Ettenson R., Gaeth G. (1994) "How Consumers Choose Health Insurance". *Journal of Health Care Marketing* 14 (1): 21-33.
- Chassagnon A. (1996) *Sélection Adverse: Modèle générique et Applications*. Doctorat Nouveau Régime, EHESS.
- Chernichovsky D., Chinitz D. (1995) "The political economy of health system reform in Israel". *Health Economics* 4 : 127-141.
- Chinitz D., Israeli A. (1997) "Health Reform and Rationing in Israel". *Health Affairs* 16 (5): 205-210.
- Clark D., Olsen J. A. (1994) "Agency in health care with an endogenous budget constraint". *Journal of Health Economics* 13 : 231-251.
- Cochrane J. H. (1995) "Time-consistent health insurance". *Journal of Political Economy* 103 (3): 445-473.
- Costes P. (1995) "Le système de santé espagnol". *Revue Française des Affaires Sociales* Mars.
- Couffignal A., Rochaix L. (1998) "Concurrence et antisélection en assurance maladie : l'expérience des Pays-Bas". *Etudes et Recherches en Economie Publique* 2 : 35-65.
- Crocker K. J., Snow A. (1985) "A Simple Tax Structure for Competitive Equilibrium and Redistribution in Insurance Markets with Asymmetric Information". *Southern Economic Journal* 51 ((4)): 1142-1150.
- Crocker K. J., Snow A. (1985) "The Efficiency of Competitive Equilibria in Insurance Markets with Asymmetric Information". *Journal of Public Economics* 26 : 207-219.
- Culyer A. J. (1973) "Is Medical Care Different ?". In Michael Cooper, H. and Anthony J. Culyer (eds), *Health Economics*. Oxford University Press, Oxford.
- Cutler D. M. (1994b) *Public Policy for Health Care*. NBER, Working papers : 5591.
- Cutler D. M. (1994c) "A Guide to Health Care Reforms". *Journal of Economic Perspectives* 8 (3): 13-29.
- Cutler D. M., Reber S. (1996) *Paying for health insurance: the tradeoff between competition and adverse selection*. NBER, Working papers : 5796.
- Dahlby B. G. (1981) "Adverse selection and Pareto improvements through compulsory insurance". *Public Choice* 37 : 547-558.
- Dardanoni V., Wagstaff A. (1987) "Uncertainty, Inequalities in Health and the Demand for Health". *Journal of Health Economics* 6 : 283-290.
- Debreu G. (1959) *Theory of value*. Wiley & Sons, New-York.
- Diamond P. (1992) "Organizing the Health Insurance Market". *Econometrica* 60 (6): 1233-1254.

- Diamond P. A. (1994) "Two Improvements on the Clinton Framework". *Journal of Economic Perspectives* 8 (3): 61-66.
- Diamond P., Rothschild M. (1989) Ed., *Uncertainty in economics*. Academic Press, San Diego.
- Dionne G., Doherty N. (1993) "Insurance with Undiversifiable Risk: Contract Structure and Organizational Form of Insurance Firms". *Journal of Risk and Uncertainty* 6 : 187-203.
- Dowd B., Feldman R. (1992) "Insurer Competition and Protection from Risk Redefinition in the Individual and Small Group Health Insurance Market". *Inquiry* 29 (2): 148-157.
- Dowie J. (1975) "The portfolio approach to health behavior". *Social Science and Medicine* 9 : 619-631.
- Dreifuss R. (1996) "Le printemps de la LAMal". *Revue Médicale de la Suisse Romande* 116 : 549-552.
- Drèze J. H. (1971) "Market allocation under uncertainty". *European Economic Review* 2 : 133-165.
- Dufourq N. (1994) "Sécurité sociale : le mythe de l'assurance". *Droit Social* 3 : 291-297.
- Dunn D. L. (1998a) *Case Studies of Health Risk Adjustment: What Can Be Learned From Experience To Date*. Alpha Centre, January.
- Dunn D. L. (1998b) "Applications of Health Risk Adjustment: What Can Be Learned From Experience To Date". *Inquiry* 35 (Summer): 132-147.
- Duriez M., Lequet D. (1996) "Les services nationaux de santé de l'Europe du Sud: nouvelles références et risques de dérive". *Solidarité Santé* 2 : 61-72.
- Eeckoudt L. (1992) "Théorie du risque, assurance et sécurité sociale". *XIIIe congrès des économistes de la santé français*, FUCAM.
- Eeckoudt L., Godfroid P., Marchand M. (1998) "Risque de santé, médecine préventive et médecine curative". *Revue d'économie politique* 108 (Mai-Juin): 321-337.
- Ehrlich I., Becker G. (1972) "Market Insurance, Self Insurance, and Self Protection". *Journal of Political Economy* July/Aug.
- Eliashberg J., Winkler R. L. (1981) "Risk sharing and group decision making". *Management Science* 27 (11): 1221-1235.
- Ellis R. P., Guire M., G. T. (1990) "Optimal Payment Systems for Health Services". *Journal of Health Economics* 9 : 375-396.
- Enthoven A. C. (1988) "Managed Competition ; An Agenda for Action". *Health Affairs* Summer : 25-47.
- Enthoven A. C. (1994) " On the ideal market structure for third-party purchasing of health care". *Social Science and Medicine* 39 (10): 1413-1432.
- Ettner S. L. (1997) "Adverse selection and the purchase of Medigap insurance by the elderly". *Journal of Health Economics* 16 : 543-562.
- Evans R. G. (1983) "The Welfare Economics of Public Health Insurance : Theory and Canadian Practice". In Lars Söderström (eds), *Social Insurance. Contributions to Economic Analysis*. North-Holland Publishing Company.
- Evans R. G., Stoddart G. L., Barer M. L., Eds. (1994) Ed., *Why are some people healthy and others not ?*
- Evans T. (1998) *The UK's Independant Heath Care Sector: Key to Reforms*.
- Feldman R., Escribano C., Pellisé L. (1998) "The role of government in health insurance markets with adverse selection". *Health Economics* 7 : 659-670.
- Feldman R., Manning W. G. (1995) "A simple Formula for the optimal coinsurance rate in a health insurance policy". *Séminaire international. Economie de la santé*, Ministère de l'Economie et des Finances, Paris.
- Feldstein P. J. (1993) *Analyse économique de la santé*.
- Files A., Murray M. (1995) "German Risk Structure Compensation : Enhancing Equity and Effectiveness". *Inquiry* 32 (300-309).

- Fuchs V.-R. (1993) *The future of health policy*. Harvard University Press, Cambridge et Londres.
- Gasmi, F., Laffont, J. J., Sharkey, W. W. (1998) "Competition Policy and Universal Service", *New Challenges in Network Industries*, Toulouse, 16-17 oct.
- Gauthier A. K., Lamphere J. A., Barrand N. L. (1995) "Risk Selection in the Health Care Market: A Workshop Overview". *Inquiry* 32 (Spring): 14-22.
- Genier P. (1998) "Assurance et recours aux soins". *Revue Economique* : 809-819.
- Genier P., Jacobzone S. (1998) "Peut-on parler d'une gestion globale de son capital santé ?". *Economie et Statistique* 311 (1): 37-55.
- Gérard-Varet L. A. (1997) "Principes de l'assurance et types de risques : un cadre pour la gestion de l'analyse du "risque maladie""(eds), *Protection sociale et régulation économique*. MIRE.
- Gollier C. (1992) "Economic theory of risk exchange: A review". In Dionne G. (eds), *Contributions to Insurance Economics*. Kluwer Academic Publishers, Boston.
- Grossman H. I. (1979) "Adverse selection, dissembling, and competitive equilibrium". *The Bell Journal of Economics* 10 : 336-343.
- Grossman M. (1972) *The Demand for Health: A Theoretical and Empirical Investigation*. NBER, occasional papers.
- Habib J., Rosen B. (1998) *Discussion of the Various Requests for Changing the Capitation Formula*. JDC-Brookdale Institute.
- Heffley D. R. (1982) "Allocating Health Expenditures to Treatment and Prevention". *Journal of Health Economics* 1 : 265-290.
- Heflebower R., B. (1980) *Cooperatives and mutuals in the market system*. University of Wisconsin Press.
- Hellander I., Moloo J., Himmelstein D. U., Woolhandler S., Wolfe S. (1995) "The Growing Epidemic of Uninsurance : New Data on the Health Insurance Coverage of Americans". *International Journal of Health Services* 25 (3): 377-392.
- Hellinger F. J. (1995) "Selection Biases in HMOs and PPOs: A review of the Evidence". *Inquiry* 32 (Summer): 135-142.
- Hellwig M. F. (1987) "Some Recent Developments in the Theory of Competition in Markets with Adverse Selection". *European Economic Review* 31 : 319-325.
- Henriet D., Rochet J. C. (1987b) "Price Regulation in Insurance Markets with Asymmetric Information". *Economic Letters* 24 : 327-329.
- Henriet D., Rochet J. C. (1991) *Microéconomie de l'assurance*. Economica, Paris.
- Henriet D., Rochet J. C. (1997b) *Régulation et intervention publique dans les systèmes de santé*. Note pour le groupe santé du Conseil d'Analyse Economique, 11 déc.
- Henriet D., Rochet J. C. (1998) *Is Public Health Insurance an Appropriate Instrument for Redistribution*. GREQAM et GREMAQ.
- Henriet D., Rochet J.-C. (1999) "Régulation et intervention publique dans les systèmes de santé". In Conseil d'Analyse Economique (eds), *Régulation du système de santé*. Documentation Française, Paris.
- Hermesse J., Beekmans J. (1998) "L'assurance soins de santé : organisation, financement et remboursement". *Journal d'Economie Médicale* 16 (1): 5-20.
- Hertzman C., Evans R. G., Frank J. (1994) "Heterogeneities in Health Status and the Determinants of Population Health". In Robert G. Evans, Greg L. Stoddart and Morris L. Barer (eds), *Why are some people healthy and others not ?*
- Hibbard J. H., Jewett J. J., Legnini M. W., Tusler M. (1997) "Choosing a health plan : do large employers use the data ?". *Health Affairs* 16 (6): 172-180.
- Howland J., Strokes III J., Crane S. C., Belanger A. J. (1987) "Adjusting capitation using chronic disease risk factors". *Health Care Financing Review* 9 (2): 15-23.

- Ippolito (1981) "Information and the life cycle consumption of hazardous goods". *Economic Inquiry* 19 (4): 529-558.
- Jack W. (1997) *Managing competition in the health insurance market*. Centre for Economic Policy Research.
- Jacobzone S. (1996) "Les politiques de santé face aux propriétés incitatives et redistributives des systèmes d'assurance maladie". *Economie et Statistique* 291-292 : 49-70.
- Jones A., Zanola R. (1995) "Agency and Health Care: A Suitable Case for Public Policy ?". *Catania Conference in Public Policy*. Sept.
- Jones S. (1989) "Can Multiple Choice Be Managed To Constrain Health Care Costs ?". *Health Affairs* Fall : 51-59.
- Jones S. (1990) "Multiple Choice Health Insurance: The Lessons and Challenges to private Insurers". *Inquiry* 27 : 161-166.
- Knutson D. (1998) "Case Study: The Minneapolis Byuers Health Care Action Group". *Inquiry* 35 (Summer): 171-177.
- Kronick R., Zhou Z., Dreyfus T. (1995) "Making Risk Adjustment Work for Everyone". *Inquiry* 32 (Spring): 41-55.
- Laffont J.-J., Tirole J. (1993) *A theory of incentives and procurement and regulation*. MIT Press, Cambridge.
- Lamers L. M. (1997) "Capitation payments to competing Dutch sickness funds based on diagnostic information from prior hospitalization". *Third European Conference on Health Economics*, Stockholm.
- Le Clairche C. (1996) *Revenu minimum, justice sociale, niveau de vie et bien-etre*. Doctorat ès Science Economique, Université Paris-IX Dauphine.
- Le Pen C. (1988) "Demande de soins, demande de santé". *Revue d'économie politique* 4 : 445-507.
- "Les régimes de santé dans les pays nordiques" (1995) *Espace Social Européen, Documents et Synthèses* 7 (Mars).
- Levy E. (1988) "La demande en économie de la santé". *Revue d'économie politique* 4 : 445-507.
- Levy E., Le Pen C., Rochaix L. (1993) *Responsabiliser..Autrement, de la participation individuelle au financement des soins médicaux*. Cahiers du C.E.R.E.S.A., Université de Paris Dauphine.
- Liaropoulos L. L. (1995) "Health services in Greece: a role for private health insurance". *Health Policy* 34 : 53-62.
- Lohr K. N., al. e. (1986) "Effect of cost sharing on use of medically effective and less effective care". *Medical Care* 24 (Supplement): 31-38.
- Luft H. S. (1995) "Potential Methods to Reduce Risk Selection and Its Effects". *Inquiry* 32 (Spring): 23-32.
- Ma C.-t. A. (1994) "Health Care Payment Systems: Cost and Quality Incentives". *Journal of Economics and Management Strategy* 3 (1): 93-112.
- Maarse H., Paulus A. (1998) "Health-Insurance Reforms in the Netherlands, Belgium and Germany: A Comparative Analysis". In R. Leidl (eds), *Health Care and its Financing in the Single European Market*. IOS Press, Amsterdam.
- Malinvaud E. (1972) "The allocation of individual risks in large markets". *Journal of Economic Theory* 4 : 312-328.
- Malinvaud E. (1973) "Markets for an exchange economy with individual risks". *Econometrica* 41 (3): 312-328.
- Manning W. G., Marquis S. M. (1989) *Health Insurance: the Trade-Off Between Risk Pooling and Moral Hazard*. Rand Corporation.

- Manning W. G., P. N. J., Duan N., Keeler E. B., Leibowitz A., Marquis S. (1987) "Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment". *American Economic Review* 77 (3, June): 251-277.
- Mapelli V. (1995) "Private Health Insurance in Italy". In Markus Schneider (eds), *Complementary Health Schemes in the European Union*, Augsburg.
- Margolis H. (1982) *Selfishness, Altruism and Rationality*. Cambridge University Press, Cambridge.
- Marquis M. S. (1992) "Adverse Selection with a Multiple Choice among Health Insurance Plans: A Simulation Analysis". *Journal of Health Economics* 11 : 129-151.
- Marquis M. S., Phelps C. E. (1987) "Price Elasticity and Adverse Selection in the Demand for Supplementary Health Insurance". *Economic Inquiry* XXV (April): 299-313.
- Marshall J. M. (1974) "Insurance theory: reserves versus mutuality". *Economic Inquiry* 12 : 476-492.
- Masson A. (1995) " Assurance sociale et assurance privée face à la crise de l'état providence". *Risques* 24 (oct.-déc.): 135-156.
- Maynard A. (1995) "Health care reforms: Don't confuse me with facts, Stupid!". *Four Country Conference on Health Care Reforms and Health Care Policies in the US, Canada, Germany and the Netherlands*, The Hague.
- Ministère fédéral du Travail d. l. S. e. d. A. s. (1998) Ed., *La santé publique en Autriche*. Vienne.
- Miyazaki H. (1977) "The rat race and international labor markets". *The Bell Journal of Economics* 8 : 394-418.
- Morgan R. O., Virnig B. A., DeVito C. A., Persily N. A. (1997) "The MEDICARE-HMO revolving door - the healthy go in and the sick go out". *The New England Journal of Medicine* 337 (3): 169-175.
- Mosca G. (1997) "Les failles du système de santé suisse". *Questions de sécurité sociale* 5 (Mai): 21-27.
- Mougeot, M. (1999) Régulation du système de santé, Rapport du Conseil d'Analyse Economique, Documentation Française, Paris.
- Neudeck W., Podczeck K. (1996) "Adverse selection and regulation in health insurance markets". *Journal of Health Economics* 15 : 387-408.
- Newhouse J. P. (1994) "Patients at Risk: Health Reform And Risk Adjustment". *Health Affairs* Spring (I): 132-146.
- Newhouse J. P. (1994) "Symposium on Health Care Reform". *Journal of Economic Perspectives* 8 (3): 3-11.
- Newhouse J. P. (1996) "Reimbursing Health Plans and Health Providers: Efficiency in Production Versus Selection". *Journal of Economic Literature* XXXIV (September): 1236-1263.
- Newhouse J. P., Chairman, 1998, "Restructuring Medicare Choices", *Final Report of the Study Panel on Capitation and Choice*, National Academy of Social Insurance, Avril.
- Newhouse J. P., Manning W. G., Keeler E. B., Sloss E. M. (1989) "Adjusting capitation rates using objective health measures and prior utilization". *Health Care Financing Review* 10 (3): 41-54.
- Nonneman W., van Doorslaer E. (1994) "The role of sickness funds in the Belgian health care markets". *Social Science and Medicine* 39 (10): 1483-1495.
- OCDE (1994) Ed., *La réforme des systèmes de santé : Etude de dix-sept pays de l'OCDE*. Paris.
- OECD (1997) *OECD Economic Surveys : Germany*.
- Okma G. H. (1997) *Studies on Dutch Health Politics, Policies and Law*. Ph.D. Thesis.
- Pannequin F. (1992) *Théorie de l'assurance et sécurité sociale*. Doctorat en sciences économiques, Université Paris I- Panthéon Sorbonne.
- Pauly M. V. (1968) "The Economics of Moral Hazard: Comment". *American Economic Review* 58 : 531-537.

- Pauly M. V. (1984) "Is cream skimming a problem for the competitive medical market ?". *Journal of Health Economics* 3.
- Pauly M. V. (1986) "Taxation, Health Insurance, and Market Failure in the Medical Economy". *Journal of Economic Literature* XXIV (June): 629-675.
- Pauly M. V. (1988) "Market Power, Monopsony, and Health Insurance Markets". *Journal of Health Economics* 7 : 111-128.
- Peele P. B. (1993) "Evaluating welfare losses in the health care market". *Journal of Health Economics* 12 : 205-208.
- Pellisé L. (1995) *Regulationg competition in the spanish health sector. Capitation and risk selection in MUFACE*. Universidad Carlos III.
- Perneger T., Etter J.-F., Schaad B. (1995) "Evaluation de l'HMO de l'université de Genève". *Etudes et Recherches* Université de Genève (6).
- Phelps C. E. (1973) *Demand for health insurance: a theoretical and empirical investigation*. Rand Corporation, July.
- Poterba J. M. (1994) *Government Intervention in the Markets for Education and Health Care: How and Why ?* NBER, Working papers : 4916.
- Price J. R., Mays J. W. (1985) "Biased selection in the Federal Employees Health Benefits Programs". *Inquiry* 22 : 67-77.
- Private Health Insurance. Facts and Figures 1997/1998* (1999) Verband der privaten Krankenversicherung, Köln.
- Propper C. (1992) *The NHS reforms and the UK private sector*. June.
- Rapport de la commission pour la responsabilité financière des organismes assureurs* (1994), Bruxelles (Mars).
- Raynaud D. (1998) *Santé et accès aux soins*. GREMAQ.
- Report of the Advisory Group on the Risk Equilisation Scheme* (1998) Irish Ministry for Health and Children, Dublin (April).
- Report to the Congress: Medicare Payment Policy* (1998) MedPAC (Eds.), Washington DC.
- "Responsabilité financière des assureurs : le point" (1998) *Interface, mensuel d'information de la mutualité socialiste*, Juin : 8-12.
- Rice T. (1993) "Demand curves, economists, and desert islands: A response to Feldman and Dowd". *Journal of Health Economics* 12 : 201-204.
- Rice T. (1997) "Can Markets Give Us the Health System We Want ?". *Journal of Health Politics, Policy and Law* offprint by Duke University.
- Riley G., Lubitz J., Rabey E. (1991) "Enrollee Health Status under Medicare Risk Contracts : An Analysis of Mortality Rates". *Health Services Research* 26 (2): 137-163.
- Riley G., Tudor C., Chiang Y.-P., Ingber M. (1996) "Health Status of Medicare Enrollees in HMOs and Fee-for-Service in 1994". *Health Care Financing Review* 17 (4): 65-76.
- Riley J. G. (1979) "Informational Equilibrium". *Econometrica* 47 : 331-359.
- Robinson J. C. (1996) "Dynamics and Limits of Corporate Growth in Health Care". *Health Affairs* 15 (2): 155-169.
- Rochaix L. (1995) "Le financement par les particuliers : la boîte de Pandore". *Revue d'Economie Financière*, numéro spécial santé, Oct.
- Rochaix L. (1996) "L'analyse du marché des soins médicaux : quelle place pour l'économie de la santé". *Revue d'Epidémiologie et de Santé Publique* 44 : 498-510.
- Rochaix L., Jacobzone S. (1997) "L'hypothèse de demande induite : un bilan économique". *Economie et Prévision, numéro spécial économie de la santé* 129-130 (Juil. - Sept.).

- Rose-Ackerman S. (1996) "Altruism, nonprofits, and economic theory". *Journal of Economic Literature* XXXIV (June): 701-728.
- Rosen B. (1998) *Price competition and the 1998 Budget Arrangements Law*. JDC-Brookdale Institute, July, 27.
- Rossiter L. F. (1987) "Risk-Based Capitation Payments for Health Care: A Survey of Literature". *Journal of Health Administration Education* 5 (4): 571-593.
- Rothschild M., Stiglitz J. (1976) "Equilibrium in competitive Insurance Markets: an essay on the economics of imperfect information". *Quarterly Journal of Economics* 90 (4): 639-649.
- Rothschild M., Stiglitz J. (1997) "Competition and Insurance Theory Twenty Years Later". *The Geneva Papers on Risk and Insurance Theory* 22 : 73-79.
- Rovira J., Mompo C., de Vildt K., Schneider M., Blasco I. (1998) "Comparing Cost-Sharing in European Union Member States: A System-Oriented Framework". In R. Leidl (eds), *Health Care and its Financing in the Single European Market*. IOS Press, Amsterdam.
- Santé 2010* (1993) Commissariat Général du Plan (Ed.). La Documentation Française, Paris.
- Schlesinger H. (1981) "The Optimal Level of Deductibility in Insurance Contracts". *Journal of Risk and Insurance* 48 : 465-481.
- Schmit J. T. (1986) "A New View of the Requisites of Insurability". *Journal of Risk and Insurance* 53 (2): 320-329.
- Schneider M. (1995) "Complementary Health Schemes in the E.C. in Comparison to Compulsory Schemes : Overview". In Markus Schneider (eds), *Complementary Health Schemes in the European Union*, Augsburg.
- Schneider M. (1995) Ed., *Complementary Health Schemes in the European Union*. Augsburg.
- Schneider-Bunner C. (1996) *Economie et Justice sociale: l'organisation et la régulation des systèmes de santé face à l'équité*. Doctorat ès Science Economique, Université de Bourgogne.
- Schneider-Bunner C. (1997) "Équité et économie de la santé : entre positivité et normativité". 10e Journées des économistes français de la santé.
- Shewry S., Hunt S., Ramey J., Bertko J. (1996) "Risk Adjustment : The Missing Piece of Market Competition". *Health Affairs* 15 (1): 171-181.
- Smyth D., O'Faherty P. (1995) "Private Health Insurance in Ireland: Recent Developments". *Third European Conference on Health Economics*, Stockholm.
- Spence M. A. (1978) "Product differentiation and performance in insurance markets". *Journal of Public Economics* 10 : 427-447.
- Swartz K. (1995) "Reducing Risk Selection Requires more than Risk Adjustments". *Inquiry* 32 (Spring): 6-10.
- Tobin J. (1970) "On limiting the domain of inequality". *Journal of Law and Economics* 13 (Octobre): 263-277.
- van Barneveld E. M., van de Ven W. P. M. M., van Vliet R. C. J. A. (1996) "Mandatory High Risk Pooling: An Approach to reducing Incentives for Cream Skimming". *Inquiry* 33 (Summer): 133-143.
- van Barneveld E. M., van Vliet R. C. J. A., van de Ven W. P. M. M. (1997) "Risk-adjusted capitation payments for catastrophic risks based on multi-year prior costs". *Health Policy* 39 (123-135).
- van de Ven W. P. M. M. (1996) "Market-Oriented Health Care Reforms: Trends and Future Options". *Social Science and Medicine* 43 (5): 655-666.
- van de Ven W. P. M. M., van Vliet R. C. J. A. (1992) "How can we prevent cream skimming in a competitive health insurance market ?". In H.E. Frech, III and P. Zweifel (eds), *Health Economics Worldwide*. Kluwer Academic Publishers, Amsterdam.
- van de Ven W. P. M. M., van Vliet R. C. J. A. (1993) "Capitation payments based on prior hospitalizations". *Health Economics* 14 : 177-188.

- van de Ven W. P. M. M., van Vliet R. C. J. A. (1995) "Consumer information surplus and adverse selection in competitive health insurance markets: An empirical study". *Journal of Health Economics* 14 : 149-169.
- van de Ven W. P. M. M., Schut, F. T., Rutten, F. (1994a) "Forming and Reforming the Market for Third-Party Purchasing of Health Care », *Social Science and Medicine*, 39(10), 1405-1412.
- van de Ven W. P. M. M., van Vliet R. C. J. A., van Barneveld E. M., Lamers L. M. (1994b) *Capitation payments: the relation between potential profits and additional variance explained by insurer. Annex to "Risk Adjusted Capitation Payments: The Achillesheel of Market-oriented Health Care Reforms"*. July.
- van de Ven W. P. M. M., van Vliet R. C. J. A., van Barneveld E. M., Lamers L. M. (1995) "Risk Adjusted Capitation Payments : The Achillesheel of Market-oriented Health Care Reforms". Séminaire international. *Economie de la santé*, Ministère de l'Economie et des Finances, Paris.
- van Doorslaer E., Wagstaff A. (1993) "Equity in the finance and delivery of health care: concepts and definitions". In Eddy Van Doorslaer, Adam Wagstaff and Frank Rutten (eds), *Equity in the Finance and Delivery of Health Care. An International Perspective*. Oxford University Press, Oxford.
- van Doorslaer E., Wagstaff A. (1993) "Equity in the finance of health care: methods and findings". In Eddy Van Doorslaer, Adam Wagstaff and Frank Rutten (eds), *Equity in the Finance and Delivery of Health Care. An International Perspective*. Oxford University Press, Oxford.
- van Doorslaer E., Wagstaff A., Rutten F. (1993) Ed., *Equity in the Finance and Delivery of Health Care. An International Perspective*. Oxford University Press, Oxford.
- van Vliet R. C. J. A., van de Ven W. P. M. M. (1992) "Towards a capitation formula for competing health insurers. An empirical analysis". *Social Science and Medicine* 34 (9): 1035-1048.
- Wasem (1995) "La régulation des marchés privés d'assurance maladie". *Séminaire international. Economie de la santé*, Ministère de l'Economie et des Finances, Paris.
- Weinick R. M., Zuvekas S. H. e., Drilea S. K. (1997) "Access to Health Care - Sources and Barriers". *Research Findings AHCPH (Agency for Health Care Policy and Research)* 98-001 (Oct).
- Wilson C. (1977) "A Model of Insurance Market with Incomplete Information". *Journal of Economic Theory* 16 (2): 167-207.
- Wilson R. (1968) "The theory of syndicates". *Econometrica* 36 (1): 119-132.
- Wilson V. M., Smith C. A., al e. (1998) "Case Study: The Washington State Health Care Authority". *Inquiry* 35 (Summer): 171-177.
- Winter R. A. (1992) "Moral Hazard and insurance contracts". In Dionne G. (eds), *Contributions to Insurance Economics*. Kluwer Academic Publishers, Boston.
- Wolfe J., R., Goddeeris J., H. (1991) "Adverse selection, moral hazard, and wealth effects in the Medigap insurance market". *Journal of Health Economics* 10 (4): 433-459.
- Zeckhauser R. (1995) "Insurance and Catastrophes". *The Geneva Papers on Risk and Insurance Theory* 20 : 157-175.
- Zweifel P., Breyer F. (1997) *Health Economics*. Oxford University Press, New-York.

TABLE DES MATIERES

INTRODUCTION GENERALE	9
-----------------------------	---

PARTIE I : LE SERVICE D'ASSURANCE EN SANTE

INTRODUCTION	15
CHAPITRE 1 : PRINCIPES FONDAMENTAUX DE L'ASSURANCE	17
1. FORMALISATION DU RISQUE	18
1.1 Bien et prix contingents	18
1.2 Risque individuel et risque social	19
1.2.1 Définition de l'aléa au niveau individuel	20
1.2.2 Distinction des états collectifs et des états sociaux	20
1.2.3 Distinction du risque social et du risque individuel	21
2. PARTAGE DU RISQUE INDIVIDUEL	22
2.1 L'optimum	22
2.2 L'équilibre des marchés Arrow-Debreu	23
2.3 La couverture réciproque	24
2.3.1 Définition des contrats et programme du consommateur	24
2.3.2 Solution	25
2.4 Le mécanisme d'assurance	27
2.4.1 Hypothèses et définition du contrat d'assurance	28
2.4.2 Résolution et interprétation	29
3. PARTAGE DU RISQUE SOCIAL	31
3.1 L'optimum	32
3.1.1 Position du problème	32
3.1.2 Résolution et analyse	32
a) Egalité des taux marginaux de substitution	33
b) Monotonie des consommations individuelles quand la richesse agrégée augmente	33
c) Mise en commun des ressources agrégées	33
d) Caractérisation de la règle de partage	34
3.2 L'équilibre	35
3.2.1 Position du problème	36
3.2.2 Interprétation des résultats	36
3.3 Le principe de mutualité	37
4. LE SERVICE D'ASSURANCE	40
4.1 L'offre et la demande d'assurance	40
4.1.1 Les conditions d'assurabilité d'un risque	40
4.1.2 La demande d'assurance	41
4.2 L'endogénéisation des risques	43
4.2.1 Auto-protection, auto-assurance et assurance	43
a) L'auto-assurance	43
b) L'auto-protection	44
4.2.2 Le risque moral	45
CONCLUSION	47

CHAPITRE 2 : LE RISQUE ET LA REPARATION DU DOMMAGE EN SANTE	48
1. FACTEURS DE RISQUE ET NATURE DE L'ALÉA EN SANTÉ	49
1.1 L'approche déterministe	49
1.1.1 Le déterminisme comportemental	49
1.1.2 Le déterminisme environnemental.....	53
1.2 L'approche stochastique.....	55
1.2.1 Incertitude initiale	55
1.2.2 Maladie et dommage	57
a) <i>Probabilité de réalisation</i>	58
b) <i>Nature du choc</i>	58
Ampleur et réversibilité.....	58
Choc sur le stock ou choc sur le taux.....	60
1.3 Synthèse.....	62
2. DÉTERMINANTS DE L'UTILISATION DES SOINS	64
2.1 Fonction de production et demande de soins	64
2.2 Caractérisation des outputs de la fonction de production.....	66
2.2.1 Les soins ne produisent pas nécessairement de la santé.....	66
2.2.2 Les soins ne produisent pas uniquement de la santé	67
2.2.3 Les soins produisent des externalités.....	69
a) <i>Les externalités physiques</i>	69
b) <i>Les externalités psychologiques</i>	70
2.3 Souveraineté de la décision de consommer des soins	71
2.3.1 Hypothèse de rationalité individuelle.....	71
2.3.2 Transparence du marché	71
2.3.3 Choix non éclairés et rôle du producteur	72
CONCLUSION	74
CHAPITRE 3 : DECLINAISON A LA SANTE DES PRINCIPES ASSURANTIELS	75
1. OBJET DE L'ASSURANCE SANTÉ	76
1.1 L'assurance santé porte sur les soins	76
1.2 Le caractère limitatif de ce principe	78
2. PRISE EN COMPTE DES COMPORTEMENTS STRATÉGIQUES PAR L'ASSUREUR	81
2.1 Le risque moral.....	81
2.1.1 Assurance et prévention : le risque moral ex ante.....	82
2.1.2 Assurance et effet prix : le risque moral "ex post"	83
2.2 Vers une approche élargie	86
2.2.1 Les transactions sur le marché des soins	87
2.2.2 La gestion directe des transactions par l'assureur	89
3. TARIFICATION DE L'ASSURANCE.....	92
3.1 Tarification actuarielle du risque en santé.....	93
3.1.1 Le principe.....	93
3.1.2 La périodicité des contrats.....	95
3.2 Interprétation et conséquences pour les assurés	96
3.2.1 La pertinence limitée du mécanisme d'assurance	96
3.2.2 La question de l'assurance de long terme.....	98
a) <i>Les attentes des individus</i>	98
b) <i>Les solutions de marché partielles</i>	99
CONCLUSION	101
CONCLUSION DE LA PARTIE I.....	102

PARTIE II : DE L'ANTISELECTION A LA SELECTION :
LES LIMITES A L'EFFICACITE

INTRODUCTION	104
CHAPITRE 4 : ASSURANCE PRIVEE ET ANTISELECTION : APPROCHES THEORIQUES	106
1. OPTIMUM ET ÉQUILIBRE AVEC ANTISÉLECTION	108
1.1 Cadre général de l'analyse	108
1.1.1 Les principales hypothèses	108
1.1.2 Les principes de l'analyse graphique	110
<i>a) Les droites actuarielles</i>	<i>110</i>
Etude d'une droite	111
Les trois droites représentées	111
<i>b) Les préférences des individus</i>	<i>112</i>
Les courbes d'indifférences	112
Le taux marginal de substitution (TMS)	112
Calcul du TMS	112
1.2 Caractérisation des optima en présence d'antisélection	113
1.2.1 Problématique	113
1.2.2 Résolution	114
<i>a) Conditions nécessaires</i>	<i>114</i>
<i>b) Contrats des bas risques</i>	<i>115</i>
<i>c) Détermination des optima de second rang</i>	<i>117</i>
1.3 Equilibre des marchés d'assurance	118
1.3.1 Le modèle de Rothschild et Stiglitz	119
<i>a) Principe</i>	<i>119</i>
<i>b) Limites</i>	<i>121</i>
1.3.2 La sophistication du modèle de RS	122
<i>a) Modification des comportements stratégiques des acteurs</i>	<i>122</i>
<i>b) La redéfinition de la contrainte de profit</i>	<i>125</i>
2. ANTISÉLECTION ET INTERVENTION DE L'ÉTAT	129
2.1 Intervention de l'Etat comme assureur	131
2.1.1 Assurance publique partielle et assurance complémentaire privée	131
<i>a) Principe</i>	<i>131</i>
<i>b) Robustesse du résultat</i>	<i>133</i>
2.1.2 Assurance publique et option de sortie	135
<i>a) Détermination du montant et de la méthode de financement du déficit public</i>	<i>135</i>
<i>b) Comportement des individus et construction de l'équilibre</i>	<i>137</i>
Conclusion partielle	138
2.2 Réglementation du marché de l'assurance par l'Etat	138
2.2.1 Assurance minimale	138
2.2.2 Définition par l'Etat de contrats standards	140
<i>a) Contrat standard d'assurance complète</i>	<i>141</i>
Un équilibre pooling ne peut émerger	141
L'équilibre séparateur est aussi contestable	142
La proposition d'Henriet et Rochet	143

<i>b) Contrat standard d'assurance partielle</i>	143
<i>c) Analyse critique de ces résultats</i>	144
Conclusion du raisonnement précédent	144
Remise en cause de la robustesse de cette conclusion	144
Discussion	145
<i>d) Le mécanisme de Crocker et Snow</i>	146
2.3 Conditions nécessaires au succès de l'intervention de l'Etat: une proposition	147
CONCLUSION	150

CHAPITRE 5 : ASSURANCE SANTE ET ANTISELECTION : APPROCHES EMPIRIQUES	151
1. SEGMENTATION DES RISQUES	152
1.1 Etudes sur données d'enquête	152
1.2 Expériences naturelles	156
1.3 Analyse critique	157
2. PREDICTIONS DU MODÈLE DE ROTHSCHILD ET STIGLITZ	160
2.1 Asymétrie d'information en faveur de l'assuré	160
2.2 Contrainte pesant sur les bas risques	164
2.2.1 Limitation de la quantité d'assurance	165
2.2.2 Un critère de jugement différent ?	166
2.3 Equilibre du marché avec antisélection	167
2.3.1 Nature de l'équilibre.....	168
2.3.2 Absence d'équilibre	169
<i>a) Le cas de Harvard</i>	170
<i>b) Une analyse de micro-simulation</i>	171
2.4 L'asymétrie d'information est elle surmontable ?	172
3. INTERVENTION DE L'ÉTAT : L'EXEMPLE DES PAYS-BAS	176
3.1 L'exemple des régimes à adhésion optionnelle	176
3.2 Le contrat standard du marché privé	178
CONCLUSION	180

CHAPITRE 6 : LA SÉLECTION EN ASSURANCE SANTE	182
1. ACCÈS A L'ASSURANCE ET REVENU	184
1.1 Accès aux soins, assurance volontaire et filet de sécurité	184
1.2 Modèle de base.....	187
1.2.1 Assurance et contrainte budgétaire	187
1.2.2 Renoncement à l'assurance	188
<i>a) Existence de trois catégories de ménages</i>	189
<i>b) Comportement de renoncement</i>	191
1.3 Participation au marché du travail.....	193
1.3.1 Comportement des individus	194
1.3.2 Introduction de l'assurance universelle	195
1.4 Tendances actuelles des réformes.....	198
2. ACCÈS A L'ASSURANCE ET SÉLECTION DIFFÉRENTIELLE	200
2.1 Définition de la sélection différentielle des risques	201
2.1.1 Les circonstances de la sélection.....	201
<i>a) Paiement prospectif imparfaitement ajusté au risque</i>	201
<i>b) Réglementation des primes</i>	203
2.1.2 Les méthodes de sélection	207

a) Refus d'assurance	207
b) Techniques de marketing.....	209
c) Calibrage des garanties	209
d) Modulation de l'offre de soins	210
Conclusion partielle	210
2.1.3 Sélection différentielle et antisélection : une clarification conceptuelle	211
a) Quelques ambiguïtés.....	211
b) Sélection et antisélection : une tentative de clarification	213
c) Difficile identification empirique.....	217
2.2 La sélection différentielle est-elle préjudiciable ?	219
2.2.1 Accès à l'assurance et aux soins	219
2.2.2 Sélection et efficacité	220
CONCLUSION	223

CONCLUSION DE LA PARTIE II.....	224
---------------------------------	-----

PARTIE III : CONCURRENCE ENTRE ASSUREURS
DANS LES SYSTEMES DE SANTE

INTRODUCTION	228
--------------------	-----

CHAPITRE 7 : FONCTIONS DE L'ASSURANCE PRIVEE
DANS LES SYSTEMES MIXTES **230**

1. UNE APPROCHE METHODIQUE DES ARTICULATIONS PUBLIC-PRIVÉ	232
1.1 Cadrage macro-économique.....	232
1.2 Pour une typologie fonctionnelle	235
1.3 Méthodologie	238
1.3.1 Définition opérationnelle de l'assurance privée	238
1.3.2 Choix méthodologiques	241
2. DÉCLINAISON DE LA TYPOLOGIE FONCTIONNELLE DE L'ASSURANCE PRIVÉE.....	243
2.1 Couverture "au premier franc": fonction 1.....	244
2.1.1 États-Unis	244
2.1.2 Allemagne et Pays-Bas.....	245
2.1.3 Synthèse sur la fonction 1	248
2.2 Couverture de biens exclus de la couverture sociale : fonction 2.....	250
2.2.1 Assurance volontaire marginale.....	250
a) Pays-Bas, Belgique, Allemagne	250
b) Deux systèmes en transition : Israël et Suisse.....	252
2.2.2 Assurance volontaire d'une gamme étendue de biens	254
2.2.3 Synthèse sur la fonction 2	255
2.3 Refinancement des copaiements publics : fonction 3	256
2.3.1 La France.....	256
2.3.2 Medicare.....	258
2.3.3 Les pays scandinaves	259
2.3.4 Synthèse sur la fonction 3	260

2.4 Accès à une qualité supérieure : fonction 4	261
2.4.1 Effet doublon complet et système de santé parallèle	261
a) <i>Grande-Bretagne</i>	262
b) <i>Europe du sud</i>	262
c) <i>La Finlande</i>	264
2.4.2 Effet doublon atténué et système de santé intégré	264
a) <i>Autriche</i>	265
b) <i>Irlande</i>	266
2.4.3 Synthèse sur la fonction 4	267
CONCLUSION	269

CHAPITRE 8 : UNE CONCURRENCE ENTRE ASSUREURS POUR L'EFFICACITE PRODUCTIVE..... 271

1. RÉGLEMENTATION DE LA CONCURRENCE : ENJEUX ET PRINCIPES	273
1.1 L'ajustement au risque	273
1.1.1 Critère d'évaluation de la qualité d'un ajustement au risque	274
1.1.2 La tarification du risque en pratique et ses limites	276
a) <i>Les contraintes techniques de l'ajustement au risque</i>	276
b) <i>Quelles variables utiliser dans l'ajustement au risque ?</i>	278
Variables socio-démographiques	278
Indicateurs de l'état de santé	279
c) <i>Bilan et perspectives</i>	282
1.1.3 Les systèmes de paiement mixtes	285
1.2 Les contraintes supplémentaires	291
1.2.1 Définition du panier offert	292
1.2.2 Détermination du périmètre et autres conditions annexes	294
a) <i>Obligation d'assurance ?</i>	294
b) <i>Obligation de participation et contraintes imposées aux OGR</i>	295
2. PRATIQUE DE L'AJUSTEMENT AU RISQUE ET DE LA CONCURRENCE TEMPEREE	297
2.1 Introduction de l'ajustement en complément de la concurrence	298
2.1.1 Les Etats-Unis	298
a) <i>La généralisation du problème de la sélection</i>	298
b) <i>L'émergence de la "managed competition"</i>	299
<i>le Health Insurance Plan of California (HIPC)</i>	300
<i>le Buyers Health Care Action Group (BHCAG)</i>	303
<i>la Washington State Health Care Authority (HCA)</i>	304
2.1.2 Irlande	307
a) <i>La fin du monopole de l'assurance volontaire</i>	307
b) <i>le mécanisme d'ajustement</i>	308
c) <i>Bilan provisoire</i>	310
2.1.3 La Suisse	311
a) <i>Le cadre général : la loi LAMal</i>	311
b) <i>L'ajustement des risques : un essai manqué</i>	312
2.1.4 Principales leçons sur les systèmes de tradition concurrentielle	315

2.2 Introduction de la concurrence et délégation de gestion	316
2.2.1 Les Pays-Bas : archétype de l'introduction de la concurrence	317
a) <i>Le contexte et les principes de la réforme</i>	317
b) <i>La détermination du budget des caisses</i>	318
Calcul de la capitation ajustée au risque	318
Ajustements ex post.....	319
Responsabilité finale des caisses	320
c) <i>Bilan provisoire des effets de la mise en concurrence des caisses publiques</i>	321
2.2.2 La Belgique.....	322
a) <i>le contexte général</i>	322
b) <i>La réforme du financement et le mécanisme de responsabilisation</i>	323
Le mécanisme général	323
Le calcul du budget normatif	324
c) <i>Quelques éléments de résultat</i>	325
2.2.3 L'Allemagne	326
a) <i>Le contexte de la réforme</i>	326
b) <i>L'ajustement des risques</i>	327
c) <i>Bilan provisoire</i>	328
2.2.4 Israël	329
a) <i>La réforme de 1995</i>	329
b) <i>Le financement prospectif des caisses</i>	331
2.2.5 Medicare.....	333
2.2.6 Principales leçons des systèmes où la concurrence est introduite.....	335
 CONCLUSION DU CHAPITRE ET DE LA PARTIE III	 340
 CONCLUSION GENERALE	 346
 BIBLIOGRAPHIE	 340
 TABLE DES MATIERES	 360

Vu : Le Président
M.

Vu : les suffragants
M.

Vu et permis d'imprimer :

le Vice-Président du Conseil Scientifique chargé de la Recherche de l'Université Paris-IX DAUPHINE.

Résumé

Depuis plusieurs années, la nécessité de contenir les dépenses de santé financées par la collectivité, sans altérer la qualité des soins, a conduit de nombreux pays à introduire dans les systèmes publics d'assurance maladie, des mécanismes de marché réputés porteurs d'une plus grande efficacité. Dans le même temps, la quasi-totalité des pays développés, a cherché à maintenir ou à généraliser l'accès à une couverture maladie.

Nous montrons dans ce travail que ces évolutions traduisent la poursuite d'un même objectif : améliorer la gestion du risque en santé. De plus, elles entraînent, dans des systèmes à l'origine très différents, une harmonisation des modalités de la concurrence.

Dans un premier temps, nous présentons les caractéristiques du service que peut rendre un assureur privé, à la lumière de la théorie de l'assurance et compte tenu de la nature spécifique du risque santé.

Puis, nous développons une analyse théorique qui montre en quoi la concurrence non réglementée, sur le marché de ce service, génère des formes de sélection, par le risque et par le revenu, qui limitent fortement les bénéfices attendus de la concurrence en termes d'efficacité productive. Il est donc logique que la plupart des réformes s'articulent, explicitement ou non, autour de la régulation des pratiques de sélection.

Enfin, nous procédons à une analyse comparative des systèmes de santé qui nous conduit à mettre en évidence :

- que l'assurance volontaire privée joue un rôle plus marginal que ce que l'analyse théorique aurait pu laisser prévoir. Elle reste fondamentalement complémentaire de l'assurance publique en remplissant des fonctions différentes selon les caractéristiques de ce service public ;
- que les réformes ou les évolutions spontanées des systèmes d'assurance traduisent l'émergence d'un consensus autour d'un mode de gestion particulier de la concurrence : celle-ci ne porte pas tant sur la fonction d'assurance que sur la gestion du risque. Une analyse détaillée permet de dégager les enjeux techniques de ces réformes et d'en tirer un premier bilan critique.

Cette démarche confirme que, pour promouvoir la maîtrise des coûts, l'organisation de la concurrence nécessite de réaliser un arbitrage au détriment de l'efficacité allocative et peut conduire à augmenter l'équité du système, afin de limiter la sélection.

MOTS CLES : ASSURANCE SANTE, CONCURRENCE, EFFICACITE, REGLEMENTATION, SELECTION.

Competition in health insurance: the balance between efficiency and selection

Over the last fifteen years, the necessity of curbing publicly-financed health expenditures while maintaining the quality of care, has led several developed countries to introduce market incentives into their public health insurance systems. At the same time, most developed countries have tried to maintain or increase their populations' health insurance coverage.

This study shows that these changes are driven by a common objective: to improve the management of health-related risks. Moreover, despite different traditions, they have all converged towards similar means of managing competition. The first part uses standard insurance theory to describe the kind of services private insurers could offer, while taking into account the nature of health-related risks.

In the second part, a theoretical analysis explains how unregulated competition in the health insurance market generates selection on the basis of income and risk level. Selection makes the market less efficient, thereby reducing the benefits of competition. Therefore, it is logical that these reforms try, whether explicitly or not, to limit selection.

The third part compares several health care systems as well as some of their recent reforms and shows that :

- private voluntary insurance plays a limited but important role in covering health-related risks. It complements public systems by filling gaps in insurance coverage or by providing access to higher quality care;
- reforms and spontaneous changes in health insurance systems reflect a consensus on a particular model of competition. This form of competition concerns the risk management function much more than the insurance function. A detailed analysis of these reforms leads to a description of the practical conditions that must be met in order to implement them. It also assesses the preliminary results of these reforms.

This analysis confirms that, in order to overcome selection, the promotion of cost containment needs to be done at the expense of allocative efficiency and may lead to an increase in equity.

KEYWORDS : COMPETITION, EFFICIENCY, HEALTH INSURANCE, REGULATION, SELECTION.