

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 θ ou D .

Sans perte de généralité, nous poserons $U(0) = \theta$.

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.

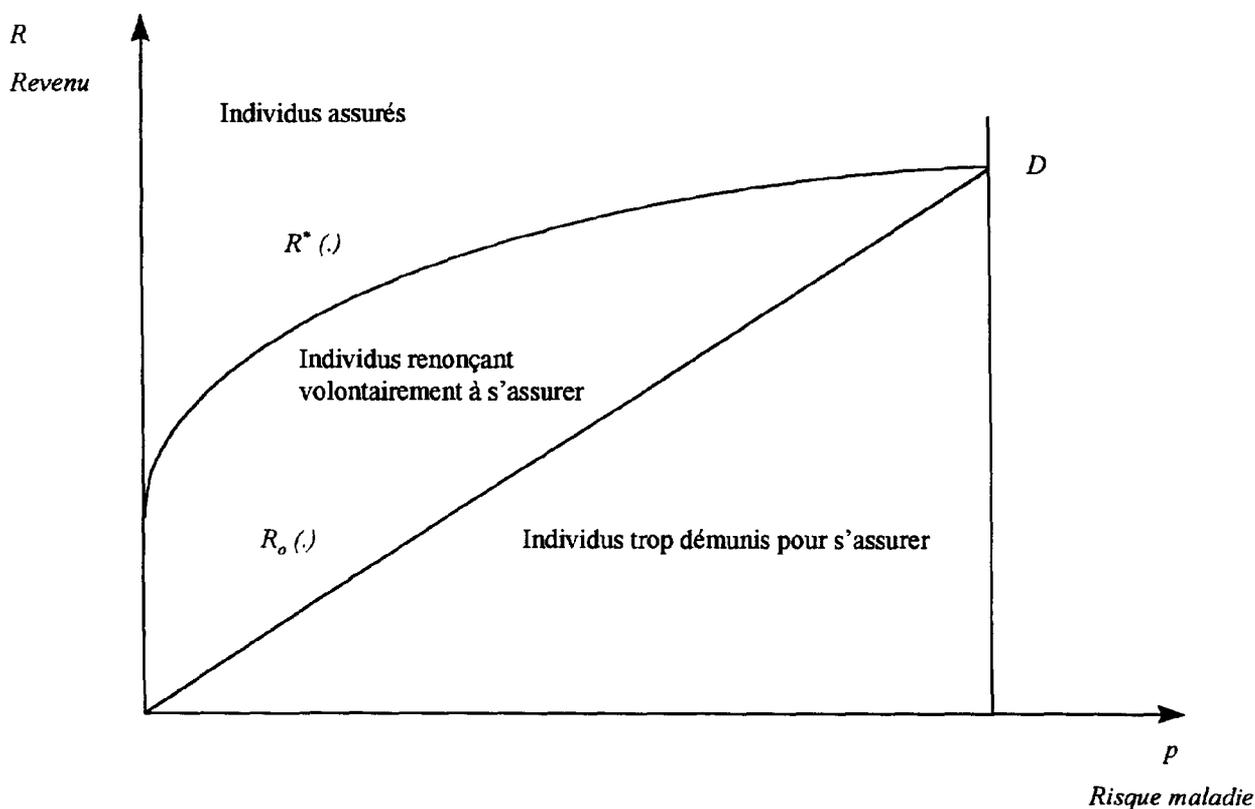
En $p = 0$:

$$\Phi(R, p) \sim p[DU'(R) - U(R)] = 0 \Leftrightarrow \frac{U}{U'}(R^*(0)) = D$$

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 :

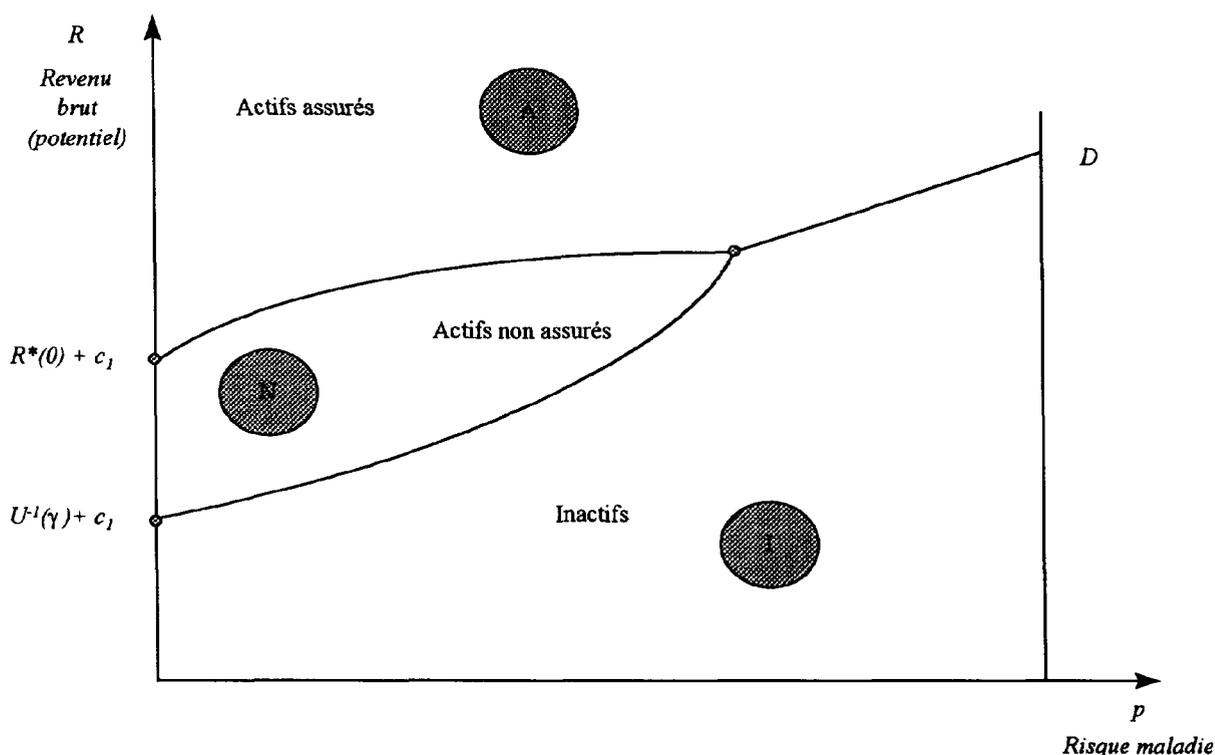
$U_A \geq U_N$ (caractérisée par $R - c_I \geq R^*(p)$, voir proposition 1) et

$U_A \geq U_I$ 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) :

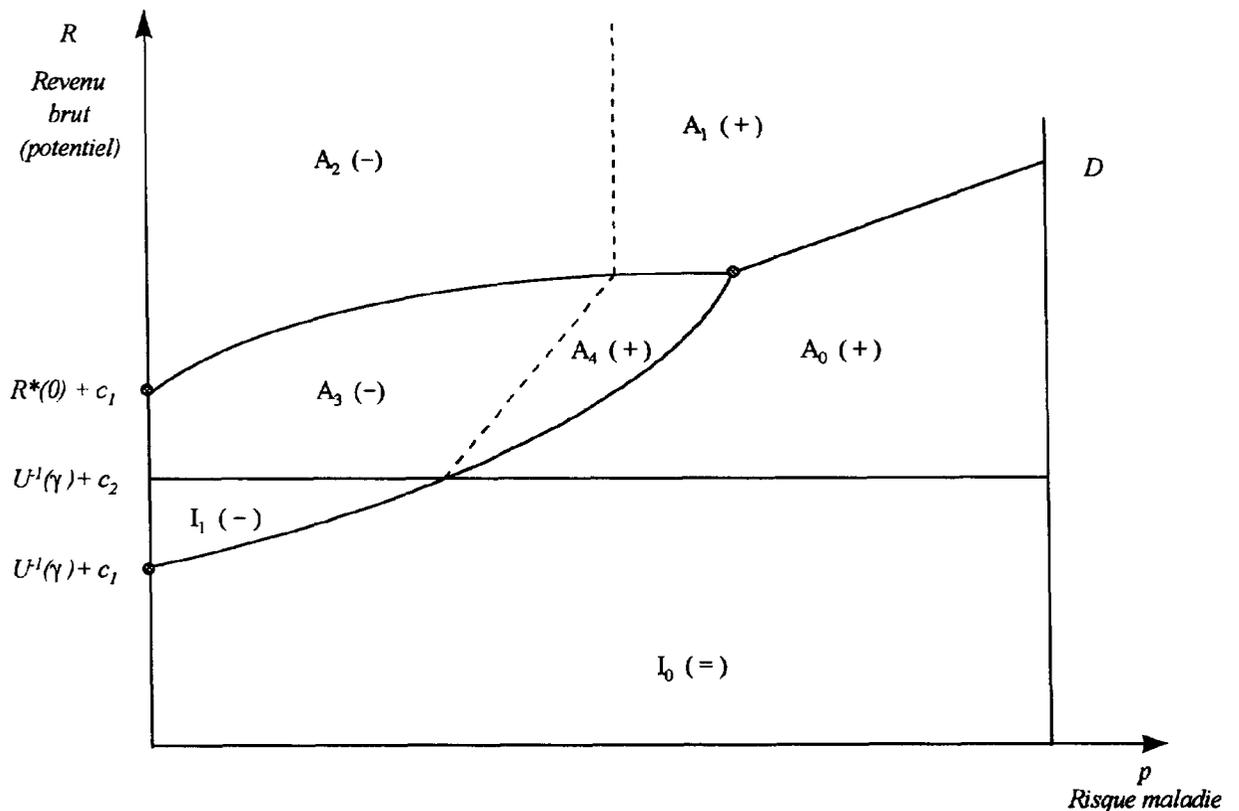
$$p_m D = c_2 (1 - G(c_2 + U^{-1}(\gamma)))$$

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).

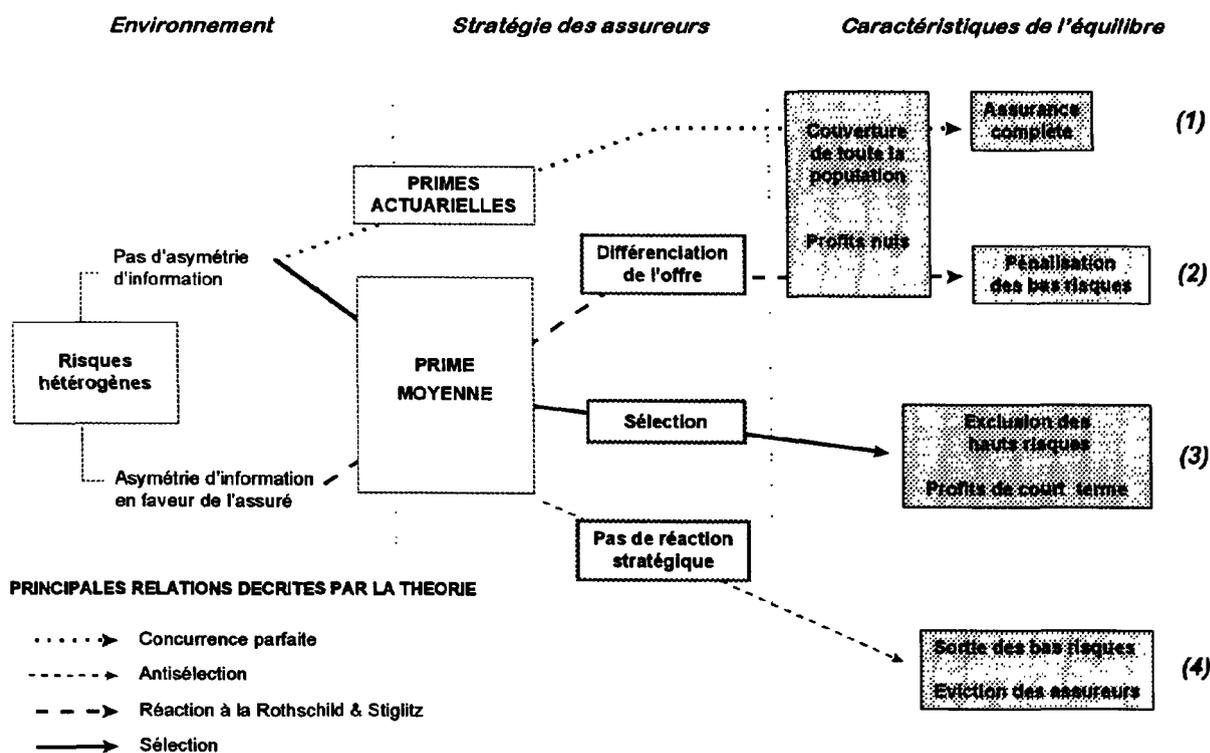
Dans cette optique, on pourrait penser que antisélection et sélection sont mutuellement exclusives puisqu'elles reposent sur des asymétries d'information de sens inverse.

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.*