

Assurance maladie et aléa de moralité ex-ante : L'incidence de l'hétérogénéité de la perte sanitaire

David Alary¹ et Franck Bien²

Cet article examine l'incidence de l'hétérogénéité de la perte sanitaire sur les contrats d'assurance maladie en présence d'aléa moral. Même si le risque sanitaire peut inciter les assurés à réaliser un effort préventif en présence d'une couverture complète ce qui induit l'absence d'un ticket modérateur, celui s'avère nécessaire pour discriminer entre les assurés lorsque leur hétérogénéité les amène à réaliser des actions préventives différentes.

HEALT INSURANCE AND HAZARD MORAL :
THE INCIDENCE OF TYPE OF HEALTH RISK

This paper investigates the impact of the heterogeneity of loss status on the insurance contracts. If a full coverage is possible with moral hazard, the heterogeneity of loss status lead to a deductible to achieve screening between insured.

Classification *JEL* : D8, I1

La dernière réforme de l'assurance maladie utilise le ticket modérateur comme instrument de régulation des dépenses de santé qui résulteraient d'un comportement de risque moral ex-ante.

Un résultat théorique en économie de l'assurance stipule qu'il est optimal d'adopter une couverture partielle (prenant notamment la forme d'un co-paiement équivalent au ticket modérateur) afin de réduire les comportements de risque moral ex-ante (Shavell [1979] et Arnott et Stiglitz [1988]). On peut donc s'interroger sur la pertinence de l'absence de tickets modérateurs qui prévaut pourtant dans de nombreux cas en assurance maladie et également de la différence de tickets modérateurs pour différentes situations sanitaires.

¹ David.alary@dauphine.fr SDFi, Université PARIS-DAUPHINE Place du Maréchal De Lattre de Tassigny 75775 Paris Cedex 16.

² franck.bien@dauphine.fr. EURISCO-LEGOS, Université PARIS-DAUPHINE Place du Maréchal De Lattre de Tassigny 75775 Paris Cedex 16.

L'assurance maladie est spécifique puisqu'elle couvre un risque qui génère deux effets : un financier et un sanitaire. Bien que le premier effet soit assurable, le second ne l'est pas. Ainsi, la perte sanitaire peut être appréhendée comme un « background risk » dans le sens où il s'agit d'un risque supplémentaire non assurable. Toutefois, ce risque modifie la perception de la richesse sans l'altérer ce qui le différencie d'un « background risk » financier. A la suite de Cook et Graham [1977] qui définissent un bien irremplaçable comme un bien modifiant l'utilité obtenue par la richesse³, nous considérons que le capital sanitaire est un exemple d'un bien dit irremplaçable.

La spécificité du risque sanitaire modifie les caractéristiques des contrats d'assurance en présence d'aléa moral. En effet, une couverture complète du risque financier ne désincite pas toujours l'assuré à fournir un effort préventif car il subsiste un bénéfice lié à la réalisation d'un effort préventif qui modifie la probabilité d'occurrence de la perte sanitaire non assurable. Bien [2004] et Bardey et Lesur [2005] montrent qu'une couverture complète est offerte en présence de risque moral sans générer d'inefficacité, et ce pour différentes spécificités de la fonction d'utilité (respectivement une dérivée croisée⁴ négative et une dérivée croisée nulle).

Dans cet article, nous nous proposons d'apporter quelques éléments de réflexion sur la relation entre la couverture d'assurance maladie, l'effort de prévention et l'hétérogénéité de la perte sanitaire. *A priori*, la différence de perte sanitaire devrait inciter l'assuré à entreprendre des actions préventives différentes puisque le risque maladie les affecte différemment. Par conséquent, l'introduction de l'hétérogénéité de la perte sanitaire conduit à ajouter un problème de sélection adverse à un problème d'aléa de moralité⁵. Pour une couverture complète du dommage financier, les deux types d'agents ne réalisent pas le même effort préventif. Dans un cadre de sélection adverse, les résultats de Rothschild et Stiglitz [1976] (noté par la suite RS) stipulent que la discrimination entre deux types d'agents est réalisée si la condition de croisement unique de Spence-Mirrless⁶ est vérifiée

³ Rey [2003] caractérise les contrats d'assurance optimaux pour une utilité bivariée et dans un cadre d'information parfaite.

⁴ L'utilité retenue est bivariée et elle dépend de la richesse financière et du capital sanitaire de l'agent.

⁵ Pour un état de la littérature des travaux considérant aléa moral et sélection adverse, se référer à Fagart et Chopin [2007]. Nous ne présentons ici que deux travaux afin de comparer nos résultats. Il est à noter que la présence de l'aléa de moralité implique l'absence d'une couverture complète.

Jullien, Salanié et Salanié [2001] montrent que l'hétérogénéité de l'aversion au risque modifie les incitations individuelles. Plus l'agent est averse au risque et plus son effort préventif est élevé. La vérification de la contrainte de croisement unique implique que le contrat est séparateur. Il s'en suit que la couverture de l'agent plus averse au risque est plus faible que celle de l'autre agent puisqu'il présente la probabilité d'occurrence du risque la plus faible.

Fagart et Chopin [2007] montrent que l'introduction de l'hétérogénéité du coût d'autoprotection dans un cadre d'aléa moral conduit à offrir une couverture plus importante aux agents dont le coût est le plus élevé *a contrario* des agents aux coûts les plus faibles puisque ces premiers présentent la probabilité d'occurrence du risque la plus élevée.

⁶ Cette condition énonce que les courbes d'indifférence ne se coupent qu'une seule fois. Cette propriété implique que l'assuré de type haut risque préférera toujours le contrat de l'autre type. Par conséquent, l'équilibre sera séparateur.

et si les assurés se voient offrir des contrats d'assurance différents⁷. Il s'en suit que le ticket modérateur redevient optimal afin de pourvoir à la discrimination entre les types. *A priori*, pour une couverture complète du dommage financier, plus la perte sanitaire est importante plus l'effort de prévention réalisé est grand. Ainsi, le mauvais risque sanitaire devient le bon risque financier puisque sa probabilité d'occurrence du risque sanitaire est la plus faible. Nous en déduisons qu'il serait optimal d'offrir une couverture incomplète au mauvais risque sanitaire et une couverture complète au bon risque sanitaire. Ainsi, la relation négative qui prévaut entre la gravité du risque sanitaire et le ticket modérateur ne serait pas efficace.

L'article est organisé comme suit. Le cadre du modèle sera décrit dans une première section⁸. L'effort optimal sera décrit dans une seconde section. La condition de croisement unique sera étudiée dans une troisième section. La quatrième section caractérisera les contrats d'assurance d'équilibre. La dernière section conclura l'article.

MODELISATION DES CHOIX INDIVIDUELS

Nous considérons une économie comportant des compagnies d'assurance maladie intervenant sur un marché parfaitement concurrentiel dans lequel les individus diffèrent par leur perte sanitaire. Nous posons comme hypothèse qu'une partie de la population de proportion λ subit une perte sanitaire nette après traitement notée H_H et l'autre partie une perte H_L avec $H_L < H_H$. Chaque agent connaît son type de risque sanitaire *a contrario* des compagnies d'assurance : H et L pour respectivement une perte sanitaire H_H et H_L .

En raison de la concurrence, les assureurs offrent des contrats à prime actuarielle notés $C(\pi, q)$ où π désigne le montant de la prime et q le montant des indemnités qui ne peut être supérieur au montant des dommages financiers (absence de sur assurance).

L'utilité des assurés dépend du niveau de leur richesse financière W et de leur état de santé H . $U(W, H)$ est une fonction d'utilité additive et séparable :

$$U(W, H) = u(W) + H \quad (1)$$

Cette fonction est croissante et concave en la richesse. Il est à noter que la dérivée seconde croisée notée U_{12} est égale à 0⁹. Cette hypothèse permet de

⁷ L'assuré de type haut risque obtient une couverture complète et l'autre assuré une couverture partielle.

⁸ Ce travail reprend en partie la logique de résolution de Jullien, Salanié et Salanié [2001]. En effet, comme pour Fagart et Chopin [2007], résoudre un problème d'assurance en présence d'aléa moral et de sélection adverse revient à considérer un problème d'assurance de sélection adverse avec des probabilités endogènes. Dans un premier temps, les contrats d'assurance en présence d'aléa moral sont définis. Puis, l'introduction de la sélection adverse les modifie.

⁹ Cette forme d'utilité est équivalente à une forme additive d'utilité dépendante de l'état de la nature. Karni [1983].

conserver un corps d'hypothèses le plus semblable à celui de Shavell [1979] et de limiter les effets santé à leur plus simple expression¹⁰.

W_0 et H_0 désignent respectivement la richesse initiale et l'état de santé initial de l'assuré, m le coût du traitement curatif.

L'assuré choisit un effort de prévention e qui diminue la probabilité d'occurrence de la maladie. Nous supposons que $e \in \mathcal{R}^+$, $p'(e) < 0$ et $p''(e) > 0$, ce qui signifie que l'effort de prévention réduit la probabilité d'occurrence de la maladie mais que le rendement marginal de l'effort est décroissant. Le coût monétaire d'une unité d'effort est noté c . Nous supposons que ce coût est suffisamment faible, c'est-à-dire que sans assurance un individu est incité à effectuer un effort préventif ex-ante.

EFFORT OPTIMAL

En présence d'aléa moral, l'inobservabilité de l'effort préventif implique que la prime d'assurance ne soit plus indexée sur le niveau d'effort. Le contrat d'assurance offert à un agent i maximise à l'équilibre l'espérance de l'utilité sous contrainte d'incitation individuelle :

$$\begin{aligned} \max_{q_i, e_i} EU_i &= p(e_i)u(W_0 - m + q_i - p_i q_i - ce_i) \\ &+ (1 - p(e_i))u(W_0 - p_i q_i - ce_i) + H_0 - p(e_i)H_i \end{aligned} \quad (2)$$

$$\frac{p'(e_i)}{c} [u(W_0 - m + q_i - p_i q_i - ce_i) - u(W_0 - p_i q_i - ce_i) - H_i] \quad (3)$$

$$= p(e_i)u_1(W_0 - m + q_i - p_i q_i - ce_i) + (1 - p(e_i))u_1(W_0 - p_i q_i - ce_i)$$

avec u_1 et u_{11} respectivement les dérivées premières et secondes, et $i \in (L, H)$.

Résultat 1 *Si l'assuré est suffisamment averse au risque financier (toutes choses égales par ailleurs), une couverture complète peut être offerte tout en incitant à réaliser un effort.*

Preuve : Après réarrangements, les conditions optimales de premier rang calculées avec une couverture complète impliquent :

$$\text{sign} \left[\frac{\partial EU_i}{\partial q_i} \Big|_{q_i = m} \right] = \text{sign} \left[\frac{\partial e_i}{\partial q_i} \right] \quad (4)$$

$$\frac{\partial e_i}{\partial q_i} = \frac{p'u_1(\cdot) - pu_{11}(\cdot)}{p''H_i - u_{11}(\cdot)} \quad (5)$$

L'effort est une fonction croissante de la couverture si l'assuré est suffisamment adverse au risque c'est-à-dire si $-\frac{u_{11}}{u_1} > -\frac{p'}{p}$. Ainsi, avec une couverture complète, l'espérance d'utilité marginale par rapport à la

¹⁰ Absence d'aversion au risque sanitaire.

couverture est une fonction croissante de celle-ci. Par conséquent, la couverture complète est bien la couverture optimale puisque la surassurance est interdite.

CQFD.

Une couverture complète ne désincite pas toujours l'assuré à fournir un effort préventif car il subsiste un bénéfice marginal lié à l'effort préventif qui résulte de la variation de l'état de santé. Pour une couverture complète, ce bénéfice marginal de réalisation de l'effort pour une couverture complète est différent de 0 car le risque sanitaire engendre une perte nette sanitaire après traitement. A l'optimum, ce bénéfice est égal au coût marginal. Il existe donc une valeur non nulle de l'effort préventif qui vérifie la contrainte d'incitation individuelle (3).

Lemme 2 : *En présence d'aléa moral, l'effort de prévention est une fonction croissante de la perte sanitaire.*

Preuve : En dérivant l'équation (3) par rapport à la perte sanitaire, nous obtenons :

$$\frac{\partial e_i}{\partial H_i} = \frac{p'}{\frac{\partial^2 EU_i}{\partial e_i^2}} \quad (6)$$

L'existence d'une solution intérieure ($e_i \neq 0$) implique que le dénominateur est négatif. $p' < 0$ implique que $\frac{\partial e_i}{\partial H_i} > 0$.

CQFD.

L'effort de prévention est donc corrélé positivement à la perte sanitaire. Plus un agent supporte une perte sanitaire importante, plus le bénéfice à réaliser un effort est élevé puisqu'il réduit ainsi la perte sanitaire espérée. Il s'en suit que le haut risque sanitaire (type H_H) est moins risqué (sa probabilité d'occurrence du risque sanitaire est plus faible) que le bas risque sanitaire (type H_L)¹¹. Par conséquent, l'hétérogénéité des agents introduit de la sélection adverse dans le modèle d'aléa moral.

LA CONDITION DE CROISEMENT UNIQUE

L'équilibre séparateur de RS est possible si et seulement si la condition de Spence Mirlees est vérifiée. Dans le cadre RS, elle implique que les agents dont la probabilité d'occurrence du risque financier est plus élevée, sont disposés à payer un montant plus faible par unité monétaire reçue en cas de réalisation de la perte que les agents de type bas risque.

¹¹ Comme Jullien, Salanié et Salanié [2001] et contrairement à Fagart et Chopin [2007], nous obtenons une corrélation négative entre la valeur du paramètre considéré (respectivement aversion au risque et coût de l'effort) et le risque financier évalué par l'espérance du dommage. Cette différence avec le second article s'explique par l'introduction de la santé dans le problème usuel de l'aléa de moralité. En effet, coût de l'effort et perte sanitaire génèrent des effets contradictoires.

Graphiquement, cette propriété est caractérisée par le fait que la pente de la courbe d'indifférence du type bas risque est supérieure à celle du type haut risque.

L'assuré de type haut risque sanitaire réalise plus d'effort de prévention que le type bas risque sanitaire ce qui conduit à une probabilité d'occurrence du risque sanitaire plus faible (bas risque au sens de RS).

Ainsi, l'agent de type haut risque sanitaire préfère des schémas de transfert plus performants¹².

Lemme 3 : Pour un contrat $C(\pi, q)$, la propriété de Spence-Mirrless est vérifiée¹³ si l'aversion absolue au risque financier est une fonction non croissante de la richesse.

Preuve : Cette propriété est vérifiée si la pente de la courbe d'indifférence dans le plan des richesses (état de non occurrence du risque sanitaire, état d'occurrence du risque sanitaire) du type haut risque financier (type H_L) est inférieure à celle du type bas risque financier (type H_H). Pour un contrat $C(\pi, q)$, ladite pente s'énonce :

$$\frac{1 - p(e_i)}{p(e_i)} \frac{u_1(W_0 - \pi - ce_i)}{u_1(W_0 - m + q - \pi - ce_i)} \quad (7)$$

D'après le lemme (2), $e_H > e_L$ et $p(e)$ étant décroissante, nous obtenons que $\frac{1 - p(e_H)}{p(e_H)} > \frac{1 - p(e_L)}{p(e_L)}$. La condition de croisement unique est vérifiée si

$$\frac{u_1(W_0 - \pi - ce_H)}{u_1(W_0 - m + q - \pi - ce_H)} > \frac{u_1(W_0 - \pi - ce_L)}{u_1(W_0 - m + q - \pi - ce_L)} \quad (8)$$

Une condition suffisante est que le rapport des utilités marginales est une fonction décroissante de l'effort. Après dérivation et arrangements, ladite condition est respectée si et seulement si

$$-\frac{u_{11}(W_0 - \pi - ce_i)}{u_1(W_0 - \pi - ce_i)} \leq -\frac{u_{11}(W_0 - m + q - \pi - ce_i)}{u_1(W_0 - m + q - \pi - ce_i)} \quad (9)$$

L'inéquation (9) est vérifiée si pour tout $q \leq m$, l'aversion absolue au risque financier est une fonction non croissante de la richesse.

CQFD.

Le lemme (3) établit que l'ensemble de classes de fonctions d'utilité admettant une solution est moins restrictif que celui de Jullien, Salanié et Salanié [2001] et Fagart et Chopin [2007]. En effet, aux fonctions CARA sont ajoutées les fonctions DARA.

LES CONTRATS D'EQUILIBRE

L'introduction de l'hétérogénéité introduit

¹²Nous supposons que la performance d'un schéma est mesurée par le transfert obtenu dans l'état occurrence du risque.

¹³ Condition suffisante.