UNIVERSITÉ DE PARIS II - PANTHÉON ASSAS

École Doctorale de Science Économiques et de Gestion

Université Panthéon-Assas

ANALYSE DU RISQUE EN ASSURANCE AUTOMOBILE : NOUVELLES APPROCHES

THÈSE

pour l'obtention du grade de Docteur de l'Université de Paris II

Discipline : Sciences Économiques

Présentée et soutenue publiquement par

Meriem KOUKI - ZEKRI

JURY

M. Alexis Direr	Professeur à l'University d'Orléans Rapporteur
M. Guillaume Carlier	Professeur à l'Université de Paris Dauphine Rapporteur
M. Damien Gaumont	Professeur à l'Université Panthéon-Assas
M. Michel Grun-Réhomme	Maître de conférence à l'Université Panthéon-Assas Directeur de recherche
M. Eric Langlais	Professeur à l'Université Paris Ouest-Nanterre
M. Alain Trognon	Inspecteur général et Directeur des enseignements et de la recherche de l'INSEE

À la mémoire de mon très cher Papa,

À ma très chère Maman, ma soeur et mon frère,

À mon Amour,

Remerciements

Que toutes les personnes m'ayant aidée ou soutenue dans ce travail, de quelques manières que ce fût, trouvent dans ces lignes le signe de ma reconnaissance.

En premier lieu, je tiens à remercier Michel Grun-Réhomme d'avoir accepté la direction de cette thèse. Ses conseils avisés, sa rigueur, son exigence et la confiance qu'il a su m'accorder toutes ces années ont contribué à mener ce parcours à terme.

Mes remerciements vont à Damien Gaumont de m'avoir accueilli au sein de l'ERMES. Je désire lui exprimer toute ma gratitude pour sa disponibilité et les si précieux conseils qu'il m'a prodigués. J'ai pu apprécier son goût de la recherche lors de notre collaboration scientifique. Je lui suis très reconnaissante d'avoir accepté d'être membre du jury malgré ses nombreuses responsabilités.

Je tiens ensuite à remercier Alexis Direr et Guillaume Carlier d'avoir bien voulu rapporter cette thèse. Je leur suis reconnaissante de m'avoir fait cet honneur. Merci également à Eric Langlais et Alain Trognon d'avoir accepté de participer au jury. Compte tenu de leurs emplois du temps chargés, je suis d'autant plus honorée qu'ils aient accepté de lire ce travail.

Je souhaite remercier tous les membres de l'ERMES avec qui de multiples discussions informelles m'ont permis d'avancer dans le processus de cette thèse. Je remercie particulièrement Georges Bresson pour ses remarques éclairantes qui m'ont été d'une grande aide. Je tiens également à remercier Badi Baltagi pour ses conseils généreux lors de ses passages à l'ERMES.

Ce travail doit également beaucoup aux retours que j'ai pu avoir lors de mes présentations à l'extérieur de l'ERMES, en colloques ou en séminaires. Je tiens à remercier les participants des Journées de Microéconomies Appliquées et du Congrès International du Risque et Assurance. Je tiens ainsi à témoigner de ma gratitude à Georges Dionne rencontré à Singapour. En me faisant partager ses visions sur les problèmes informationnels sur le marché d'assurance, Georges Dionne a contribué à alimenter ma reflexion. Je remercie également Alexis Direr pour ses observations judicieuses dont j'ai bénéficié. Je le remercie d'avoir accepté d'être rapporteur de cette thèse. Je remercie aussi Serge Garcia pour ses remarques pertinentes.

Je tiens à remercier tous mes frères d'armes doctorants et neo-docteurs de l'ERMES qui ont rendu plus agréables de fastidieuses et très longues journées de travail. Par ordre alphabétique, je remercie Myriam Abdelmoula, Mohamed Amraoui, Roger Antoun, Noureddine Benlagha, Amani Ben Rejeb, Emilio Carrera, Jean-Marie Cayemitte, Nicolas Dumas, Adélaïde Fadhuile, Narjess Khili, Elie Kolakez, Mériem Maatig, Soumaya Rekayya, Nesrine Samet, Anna Sess, Aguibou Tall et Emanuel Valat. Mes remerciements vont aussi à Naima Baba-Aissa et Josette Valentin pour leur gentillesse, leur bonne humeur et leur grande serviabilité.

Enfin, et surtout, milles mercis à tous mes proches. Une reconnaissance particulière à l'homme qui partage ma vie, Walid Zekri, pour m'avoir soutenue avec tout son amour, d'avoir supporté mes humeurs pendant ces dernières années, de m'avoir accompagnée pendant des nuits interminables de travail. Parce que lui aussi est doctorant, je promets de le soutenir de tout mon coeur pour que la liste des docteurs *Kouki/Zekri* s'allonge.

Je désire exprimer ma très profonde gratitude à Mamma, Kammour, Doudou et ma tante Aicha. Même si cette dernière année, une longue distance nous a bien séparés, je n'ai jamais manqué de leur encouragement et leur soutien perpétuel. Je les remercie d'avoir su me donner à chaque moment de doute le souffle de motivation qui m'a permis de réaliser mon rêve.

C'est à ta mémoire, mon très cher Papa, que je dédie cette thèse. Toi qui connaissais si bien le domaine de l'assurance, tu m'avais transmis ton sens du détail, l'ouverture d'esprit, la droiture et la curiosité. J'espère que tu es fier de ta fille chérie !

Table des matières

Table des figures

Liste des tableaux

Introduction Générale

Quant aux deux passions de l'incertitude, ce sont la peur et l'espérance...

Salvatore Veca

« Le contrat aléatoire est une convention réciproque dont les effets, quant aux avantages et aux pertes, soit pour toutes les parties, soit pour l'une ou plusieurs d'entre elles, dépendent d'un événement incertain. Tels sont : le contrat d'assurance, le prêt à grosse aventure, le jeu et le pari, le contrat de rente viagère. Les deux premiers sont régis par les lois maritimes ». Article 1964 du Code Civil.

Ainsi, d'un point de vue juridique, le contrat d'assurance est un contrat aléatoire dans lequel chacune des parties accepte le caractère incertain de l'événement. Il donne naissance à l'obligation de l'assureur d'indemniser la victime au cas où surviendrait cet événement incertain. En l'absence d'aléa, le contrat est nul.

Si le risque ne faisait pas partie du quotidien de chacun d'entre nous, le principe même de l'assurance n'existerait pas. Le risque se définit comme l'ensemble des évènements possibles considérés comme un *mal* ou un *dommage*, ainsi que par la probabilité associée à chacun de ces évènements. D'après Knight (1921), le mot « risque » fait référence à une situation, où « la distribution du résultat parmi un ensemble de cas est connue, soit par le calcul *a priori*, soit par des statistiques fondées sur les fréquences observées »[1].

Les comportements des agents économiques face au risque sont pris en compte, « de façon relativement synthétique »[2], et formalisés par la théorie de la décision en environnement incertain. C'est un préalable indispensable à la compréhension des mécanismes du marché de l'assurance. On peut ainsi donner un sens précis aux concepts de prime de risque, d'aversion pour le risque et modéliser le choix des agents économiques.

Les modèles, très largement adoptés par les économistes, rendant possible la description des choix individuels en environnement risqué sont des modèles d'espérance d'utilité. Cela est dû, en premier lieu à Bernouilli (1738), à von Neumann et Morgenstern (1947) et à Savage (1954). La théorie de la décision a connu un développement considérable par la suite avec les approches initiées par Arrow (1971), Mossin (1968), Ehrlich et Becker (1972). Depuis, les travaux dans le domaine ont été très nombreux et se sont orientés dans différentes directions et, dont la principale est, l'analyse des problèmes d'asymétrie d'information sur le marché de l'assurance.

L'assurance est formellement un accord entre deux parties pour lequel le principal (l'assureur) accepte, moyennant un paiement d'une prime, de verser à l'agent

[1]F. Knight, *Risk, uncertainty and profit*, page 233

[2]M. Denuit et A. Charpentier (2005), *Mathématiques de l'assurance non-vie, Tome2 : Tarification et provisionnement*, Economica, page 294.

(l'assuré) un remboursement à la survenance d'une perte bien spécifique. Cet accord se matérialisant par un contrat d'assurance s'effectue en spécifiant les droits ainsi que les obligations réciproques des deux parties.

Un contrat d'assurance automobile permet de couvrir son souscripteur contre les risques d'accidents routiers. Face à des agents à différents degrés d'aversion pour le risque, l'assureur propose une variété de contrats avec plusieurs garanties. En assurance automobile, la couverture minimum est obligatoire. Elle couvre le tiers lésé en cas d'accident responsable. D'autres garanties optionnelles peuvent s'ajouter, pour s'assurer contre les dégâts causés à son propre véhicule. Afin de proposer le meilleur contrat d'assurance qui répond aux besoins de l'assuré, l'assureur doit disposer de toutes les informations indispensables à l'explication du risque de l'individu. L'information doit, bien évidemment, être disponible et fiable. Des classes de risque sont constituées à partir de ces facteurs comme l'ancienneté de permis, l'usage du véhicule, la zone d'habitation, la puissance du véhicule, la gamme du véhicule, etc. Ces facteurs doivent être faciles à mesurer, robustes, conformes à la réglementation, explicatifs de la sinistralité, respecter la vie privée et être acceptés dans la société.

L'assureur doit répartir la charge de sinistralité de façon équitable entre tous les assurés, en même temps qu'il mutualise les risques entre les assurés qui présentent des caractéristiques semblables. La segmentation sur les marchés d'assurance devient alors un moyen d'adapter la tarification aux caractéristiques de risque de chaque individu.

La mutualisation d'un portefeuille s'avère adéquate pour des risques indépendants, dispersés et de faible intensité. Toute la segmentation est axée sur l'élimination de la solidarité subventionnelle et vise à ne garder que la solidarité aléatoire.

Estimer la valeur des effets de différents facteurs sur la sinistralité à partir d'observations semble, en général, possible dans la mesure où l'assureur dispose d'un portefeuille suffisamment grand. Il importe aussi de connaître le degré de confiance que l'on accorde à l'évaluation du pouvoir explicatif de ces différents facteurs (caractéristiques du conducteur et du véhicule). La quantification de ces effets pose des problèmes liés aux propriétés du modèle utilisé : hypothèses sous jacentes, validité et choix du modèle, robustesse des paramètres, etc.

Dans la constitution des classes de risque, un équilibre doit être trouvé entre la granularité et la robustesse. Si la granularité (ou la segmentation) est trop grossière, certes la robustesse temporelle des indicateurs de sinistralité est assurée, mais la mutualisation est trop large et un concurrent peut très bien attirer les bons risques de cette classe en proposant une cotisation plus faible grâce à une segmentation plus fine. A l'inverse une granularité trop fine ne permet pas d'avoir cette robustesse.

Pour que l'assureur trouve cet équilibre entre la granularité et la robustesse, il faut que l'information sur le risque intrinsèque de l'individu soit disponible et fiable. Dans le cas contraire, des problèmes d'asymétrie d'information compromettent l'optimalité de l'équilibre. L'assureur essaye donc d'appréhender, d'estimer les risques supplémentaires liés à cette asymétrie d'information. Deux principaux problèmes existant sur le marché d'assurance ont attiré jusqu'ici l'attention des économistes : la sélection adverse et l'aléa moral.

Objet de la Thèse

Cette thèse propose une contribution à l'analyse du risque sur le marché de l'assurance automobile. L'objet de ce travail tourne autour de trois nouveaux axes :

- Le premier axe consiste à déterminer les contrats d'équilibre dans un contexte de double asymétrie d'information : d'une part, la sélection adverse porte sur l'aversion à l'effort de l'agent qui est une caractéristique inobservable par l'assureur, et d'autre part, l'aléa moral se caractérise par l'effort de prévention entrepris après avoir eu connaissance des termes du contrat d'assurance. Considérer l'aversion à l'effort comme le risque intrinsèque de l'agent différencie notre cadre théorique des études antérieures.

- Le deuxième axe s'inscrit dans un champ de recherche qui est actuellement en plaine expansion. C'est celui d'analyser d'un point de vue empirique les problèmes d'asymétrie d'information. Cette étude est appliquée sur le marché d'assurance français. La prise en compte de la sinistralité passée en utilisant des données en coupe instantanée constitue l'originalité de cette étude.

- Enfin, la troisième question soulevée concerne l'application de la surprime aux jeunes conducteurs. L'idée de l'étude s'est dégagée après un questionnement sur le fait que les jeunes conducteurs payent plus que les conducteurs expérimentés. Est-il légitime pour les assureurs d'imposer une surprime aux jeunes conducteurs ? Est-il favorable pour les assureurs de supprimer cette majoration et fidéliser cette éventuelle jeune clientèle ?

Le premier chapitre traite d'un point de vue théorique les problèmes informationnels sur le marché d'assurance, à savoir la sélection adverse et l'aléa moral. Les deux premières sections du chapitre abordent séparément les propriétés de l'équilibre d'un marché de l'assurance concurrentiel en présence d'une unique source d'asymétrie d'information. Tout d'abord, le modèle d'équilibre en présence d'antisélection à la Rothschild et Stiglitz (1976) est exposé dans la première section. Ensuite, la seconde section décrit l'équilibre en présence d'aléa moral (Arnott et Stiglitz (1988, 1991), Arnott (1992)). Dans la troisième section qui compose l'objet principal de ce premier chapitre, nous proposons un modèle de principale-agent qui couple à la fois la sélection adverse et l'aléa moral. Il existe certes plusieurs travaux théoriques allant dans ce sens, qui considèrent simultanément les deux problèmes d'asymétrie d'information. Mais, à notre connaissance, aucune de ces modélisations n'a pris en considération le paramètre de l'aversion à l'effort comme caractéristique individuelle des assurés, inobservable par l'assureur. Cette section est issue d'une collaboration avec Damien Gaumont (Gaumont et Kouki (2010)).

Le deuxième chapitre a trois principaux objectifs :

– Le premier porte sur la construction d'un nouvel indicateur de la sinistralité passée à partir de données en coupe transversale. Nous estimons son effet prédictif sur la sinistralité future par rapport à la variable du bonus malus .

– Cette variable décrivant la sinistralité passée permet de créer une *proxy* de l'effort que nous utilisons pour tester empiriquement nos prédictions théoriques que nous avons établies dans le premier chapitre.

– Et enfin, le point le plus important concerne l'étude empirique des problèmes informationnels en utilisant ces nouvelles variables dans l'approche de la corrélation conditionnelle entre le risque et la couverture d'assurance.

Cette étude se fonde sur un fichier de 50.000 observations concernant l'année 2004. Ces données proviennent d'une grande mutuelle d'assurance française.

Le troisième chapitre a pour objectif de comparer empiriquement la sinistralité entre deux catégories de conducteurs : les jeunes conducteurs et les conducteurs expérimentés. Plusieurs méthodes d'estimation sont appliquées sur le nombre de sinistres déclarés, d'une part, et leurs coûts, d'autre part, afin de dégager les variables les plus déterminantes de la sinistralité chez les jeunes conducteurs et chez les conducteurs expérimentés. L'intérêt de cette comparaison, et plus spécifiquement celle appliquée sur les coûts des sinistres, est d'étudier la suppression de la majoration appliquée quasi systématiquement par les assureurs sur les primes des conducteurs novices par rapport aux conducteurs expérimentés. En effet, Les compagnies d'assurance en France ont la possibilité de rajouter une surprime pour les jeunes conducteurs sur leur contrat d'assurance qui peut atteindre un maximum de 100% de la cotisation de base.

Chapitre 1

Les contrats d'assurance et l'impact de l'aversion à l'effort

1.1 Introduction

Le marché des assurances est un contexte privilégié d'information incomplète entre l'agent (l'assuré) et le principal (l'assureur). L'assureur ne connaît pas tout de l'assuré, et en particulier les risques financiers qu'il lui fait courir à travers sa sinistralité. Il s'en suit des opportunités de comportements stratégiques résultant d'une évaluation asymétrique des risques. La littérature économique distingue classiquement deux types d'asymétrie d'information : l'aléa moral et la sélection adverse.

L'aléa moral concerne l'influence de la couverture d'assurance sur le comportement de l'assuré. En situation d'aléa moral, les compagnies d'assurances ne connaissent pas le niveau d'auto-protection des assurés après la souscription du contrat. Les agents peuvent être incités à diminuer leur prévention du risque fondamental quand ils bénéficient d'une bonne couverture d'assurance. Dans ce cas, le contrat d'assurance a pour effet d'accroître l'exposition au risque.

Concernant le phénomène de sélection adverse ou d'antisélection, l'agent qui choisit le contrat d'assurance qu'il considère le plus attrayant, dispose d'un avantage certain en matière d'information par rapport à l'assureur. Confronté à des individus hétérogènes, l'assureur se trouve dans l'incapacité de distinguer les agents selon leur degré de risque. L'assureur ne dispose donc pas d'une information suffisante lui permettant de réaliser une adéquation entre les termes du contrat et les risques individuels. Cette situation a pour conséquence qu'un même contrat d'assurance peut être offert à des agents aux risques différents. Une tarification uniforme désavantage les agents de moindre risque, qui paieraient plus cher que ne l'exigent les impératifs liés au coût de leur risque.

La suite de ce chapitre est organisée comme suit. La seconde section présente une synthèse des principaux travaux théoriques. Les résultats relatifs à la caractérisation de l'équilibre d'un marché de l'assurance lorsqu'il existe une unique source d'asymétrie d'information sont décrits dans les troisième et quatrième sections. Nous présentons brièvement les résultats des travaux fondateurs portant sur l'antisélection, notamment ceux de Rothschild et Stiglitz (1976). Nous exposons ensuite les propriétés de l'équilibre en présence d'aléa moral en donnant une synthèse des modèles d'Arnott (1992) et d'Arnott et Stiglitz (1988).

La cinquième section de ce chapitre[1] s'inscrit dans un cadre considérant un marché d'assurance automobile concurrentiel dans lequel les problèmes d'antisélection et d'aléa moral coexistent. C'est le coeur de ce chapitre. Les agents se différencient par leur niveau d'aversion à l'effort. Cette dernière est une caractéristique individuelle que l'assureur n'observe pas. L'aversion à l'effort constitue ainsi le paramètre de sélection adverse. L'aversion à l'effort est soit très forte (par exemple, l'agent ne respecte pas le code de la route, il conduit en état d'ébriété...), soit nulle c'est-à-dire l'agent n'a aucune aversion à l'effort. A un contrat donné, en présence d'aléa moral, l'agent peut choisir un niveau d'autoprotection ou d'effort qui peut influencer sa probabilité d'accident *ex post*. Par la suite, son niveau d'effort peut changer après la conclusion du contrat d'assurance. Le niveau d'effort n'est observable qu'une fois l'accident surgit. Sous les contraintes d'incitation et de participation, nous montrons que le marché propose à l'équilibre soit un contrat séparateur qui incite les agents à fournir de l'effort de prévention, soit un contrat mélangeant avec une forte couverture d'assurance souscrit à la fois par les agents à effort nul et les agents qui sont incités à faire de l'effort .

Enfin, la dernière section conclut le chapitre.

[1]Une partie de ce chapitre a fait l'objet d'un travail en collaboration avec Pr. Damien Gaumont de l'Université de Paris 2

1.2 Les études théoriques de l'asymétrie d'information

Les travaux fondateurs de Rothschild et Stiglitz (1976) font partie des principales analyses décrivant le fonctionnement du marché d'assurance concurrentiel en présence de sélection adverse. Dans le modèle de Rothschild et Stiglitz (RS), les agents se différencient par leur probabilité d'accidents, sur laquelle porte l'antisélection. Les deux auteurs considèrent cette variable comme exogène, une caractéristique individuelle que chaque agent ne peut pas modifier et que l'assureur n'observe pas. La probabilité d'accident peut être soit faible soit élevée. Par conséquent, deux types d'agents existent sur le marché : des bas risques et des hauts risques, que l'assureur n'est pas capable de distinguer. Les auteurs supposent que les agents averses au risque sont demandeurs d'assurance. Chaque assureur propose un seul contrat d'assurance. L'équilibre peut ne pas exister. S'il existe, l'équilibre RS est défini comme un menu de contrats tels que chaque contrat est non déficitaire et qu'il n'existe aucun contrat offert en dehors de ce menu, qui serait profitable. L'équilibre RS, en présence de sélection adverse, ne peut pas être mélangeant. Il est séparateur. Les hauts risques s'autosélectionnent et obtiennent leur contrat d'assurance à coût élevé, leur offrant une couverture totale. Les agents à bas risque, quant à eux, choisissent le contrat leur offrant une couverture partielle à prix plus faible.

Les propriétés de l'équilibre concurrentiel sur un marché d'assurance en présence d'aléa moral sont décrites principalement dans une série de travaux de Arnott et Stiglitz (1988, 1991) et de Arnott (1992). Les agents supposés identiques peuvent modifier leur probabilité d'accidents en choisissant un niveau d'effort après avoir choisi leur contrat d'assurance. Cette action de prévention ou niveau d'effort est une variable non observable de l'assureur. En situation d'aléa moral, plus l'agent est couvert et éventuellement par plusieurs assureurs, moins il fait d'effort. Par conséquent, l'assureur contraint chaque agent par une clause d'exclusivité qui stipule qu'il est son seul assureur pour le risque considéré. A l'équilibre, le marché propose un contrat d'assurance dont la couverture dépend du niveau du coût de la prévention : Si le coût de la prévention est faible, l'agent peut être incité à fournir de l'effort en lui offrant un contrat à couverture partielle. Sinon, il devient très coûteux d'inciter l'agent à l'effort et dans ce cas, le contrat d'équilibre proposé est un contrat à couverture totale.

Les travaux théoriques que nous venons de décrire supposent un seul problème d'information sur le marché d'assurance : soit la sélection adverse, soit l'aléa moral. Or, dans les faits, les deux phénomènes d'asymétrie d'information peuvent exister simultanément. Plusieurs travaux théoriques ont été développés pour étudier le marché d'assurance où coexistent l'antisélection et l'aléa moral. Nous citons principalement les travaux de Stewart (1994), Chassagnon et Chiappori (1997, 2005), Fagart et Kambia-Chopin (2003), De Meza et Webb (2001), Chassagnon (2005) qui développent l'aspect concurrentiel. Julien, Salanié et Salanié (2007) s'intéressent, quant à eux, au cadre monopolistique. Stewart (1994), Chassagnon et Chiappori (1997, 2005), Fagart et Kambia-Chopin (2003) et Chassagnon (2005) supposent que les agents diffèrent par leur coût de prévention (l'antisélection porte sur le coût de prévention), tandis que De Meza et Webb (2001) et Julien, Salanié et Salanié (2007) considèrent l'aversion au risque des agents comme caractéristique de la sélection adverse (les agents diffèrent par leur aversion vis-à-vis du risque).

Stewart (1994), en adoptant le concept d'équilibre proposé par Riley (1979), met en évidence des propriétés d'équilibre proche de celles de Rothschild et Stiglitz

(1976). L'équilibre est séparateur et totalement révélateur. Tout contrat d'équilibre réalise un profit nul. Chassagnon et Chiappori (1997, 2005) proposent une extension au modèle de Rothschild et Stiglitz (1976), en introduisant de l'aléa moral. Ils y caractérisent l'équilibre d'un marché concurrentiel où il existe les deux sources d'asymétrie d'information. Ils définissent deux types d'agents caractérisés par leur probabilité d'accident *ex ante*, avant le choix du contrat. La probabilité d'accident n'est pas exogène. Elle dépend du niveau d'effort choisi par chaque agent, pour un contrat donné. Le niveau d'effort de prévention n'est pas observée par l'assureur. Pour un niveau d'effort donné, l'agent le plus risqué *ex ante* a une probabilité d'accident supérieure à l'agent le moins risqué *ex ante*. Ils supposent que le coût de prévention dépend à la fois du niveau d'effort et du type de l'agent[2]. En supposant que l'agent peut influencer sa probabilité d'accident, pour un coût d'effort donné, l'agent le plus risqué *ex ante* n'est pas forcément le plus risqué *ex post*. Les principaux résultats trouvés par Chassagnon et Chiappori (1997, 2005) se caractérisent comme suit. L'équilibre, lorsqu'il existe, vérifie certaines propriétés de l'équilibre RS : il est séparateur, les agents les plus risqués

[2]Chassagnon et Chiappori (1997) supposent dans leur modélisation théorique que la fonction d'utilité est séparatrice entre la richesse W et l'effort e. Supposer que l'utilité sépare la richesse et l'effort de prévention constitue une hypothèse courante dans la littérature de la théorie des contrats d'assurance. Plus précisément, Chassagnon et Chiappori (1997) utilisent une utilité de la forme :

$$u^{\theta}(W,e) = v(W) - c^{\theta}(e)$$

où v est la même pour tous les types agents et c est la fonction du coût de prévention associée à chaque type d'agent θ. Cependant, d'autres modèles combinant l'aléa moral et la sélection adverse, supposent que les efforts de prévention sont constants pour tous les types d'agents (De Meza and Webb (2001), Julien, Salanié and Salanié (2007)).

ex post obtiennent la franchise la plus faible. Cependant, la propriété du « *single crossing* » ou condition de Spence-Mirrlees n'est plus vérifiée. Plusieurs équilibre RS peuvent coexister. L'existence de l'équilibre ne dépend pas de la proportion de n'importe quel type de risque, contrairement à l'équilibre RS qui existe si et seulement si la proportion des hauts risques dépasse un certain seuil. Ces nouveaux résultats apparaissent suite à la non vérification de la condition de *single crossing*.

De Meza et Webb (2001) proposent un modèle qui explique pourquoi la plupart des études empiriques n'arrivent pas à vérifier les prédictions théoriques traditionnelles obtenues avec des modèles d'asymétrie d'information. Les auteurs supposent qu'il existe deux types d'agents sur le marché d'assurance : les timides et ceux qui osent. Le timide fournit de l'effort pour réduire sa probabilité de perte, chose que ne fait pas l'autre type d'agent. Le timide représente l'agent averse au risque et donc le bas risque. Celui qui ose représente l'agent neutre au risque. Les auteurs montrent que lorsque les agents diffèrent par leur aversion vis-à-vis du risque, les courbes d'indifférence des agents se coupent deux fois, et le marché d'assurance peut admettre un équilibre mélangeant. L'équilibre séparateur existe aussi.

L'approche de Fagart et Kambia-Chopin (2003) est proche de celle de Chassagnon et Chiappori (1997), où leur contribution se situe dans un cadre concurrentiel. Fagart et Kambia-Chopin considèrent plusieurs agents averses au risque et des compagnies d'assurance neutres au risque. Elles supposent que la probabilité d'accident et le coût de prévention subit par l'agent sont à la fois affectés par le niveau de l'effort et le type de l'agent. Fagart et Kambia-Chopin réduisent leur modèle de double asymétrie d'information en une forme plus simple dans laquelle n'apparaît que le problème de sélection adverse. Les auteurs montrent que l'équilibre conserve les propriétés de Rothschild et Stiglitz (1976). L'équilibre est séparateur et unique. Contrairement à Chassagnon et Chiappori (1997), l'hypothèse de *single crossing* est vérifiée.

1.3 L'antisélection

Rothschild et Stiglitz (1976) étudient le fonctionnement des marchés d'assurance en présence d'antisélection. Les auteurs supposent que les assureurs font face à des agents hétérogènes du point de vue de leur risque. L'hétérogénéité des individus associée à leur niveau de risque est représentée par leur probabilité d'accident, exogène et non observable par l'assureur. Ceci constitue la source du problème d'antisélection : la coexistence de « bons » risques et de « mauvais » risques indiscernables *a priori*.

Chaque agent est doté d'une richesse initiale W. Sa fonction d'utilité de Von Neumann et Morgenstern U est strictement croissante, strictement concave et deux fois continûment differentiable ; $U' > 0$ et $U'' < 0$. Un agent est soumis à un risque d'accident causant un dommage noté S réduisant sa richesse initiale. A chaque type d'agent, est associé une probabilité d'accident : p_h est la probabilité d'accident des hauts risques et p_l est celle des bas risques, telle que $p_h > p_l$. La structure de la population, qui est connue de l'assureur, est composée d'une proportion λ_h d'agents à hauts risques et λ_l d'agents à bas risques, avec $\lambda_h + \lambda_l = 1$. La probabilité moyenne de sinistre est donc

$$\overline{p} = p_h \lambda_h + p_l \lambda_l.$$

L'espérance d'utilité d'un agent de type i non assuré est :

$$V_i = p_i U(W - S) + (1 - p_i) U(W), \quad i \in \{l, h\}.$$

La compagnie d'assurance propose un contrat $C(\pi, R)$ constitué d'une prime de risque π versée par l'assuré et d'une couverture R en cas d'accident n'excédant pas le montant S de dommage ($0 \leq R \leq S$).

L'espérance d'utilité de l'agent i assuré est donc :

$$V_i(C_i) = p_i U(W - S - \pi + R) + (1 - p_i)U(W - \pi).$$

Notons W_0 la richesse nette de l'assuré en absence de sinistre, et W_1 la richesse nette lorsque l'individu subit un sinistre. Soit :

$$W_0 = W - \pi \quad \text{et} \quad W_1 = W - S - \pi + R.$$

Nous considérons donc la courbe d'indifférence de l'agent i dans le plan (W_0, W_1) définie par

$$p_i U(W_1) + (1 - p_i)U(W_0) = K, \quad \text{pour un } K \text{ donné.}$$

Il s'en suit par différentiation :

$$p_i U_1^{'} dW_1 + (1 - p_i)U_0^{'} dW_0 = 0,$$

avec $U_0 = U(W_0)$ et $U_1 = U(W_1)$.

La pente de la courbe d'indifférence est donc :

$$\frac{dW_1}{dW_0} = -\frac{(1 - p_i)U_0^{'}}{p_i U_1^{'}}.$$

Comme $p_h > p_l$, la courbe d'indifférence du bas risque est plus pentue que celle du haut risque[3] dans le plan (W_0, W_1). Les deux courbes d'indifférence se croisent donc une seule fois. C'est la propriété du *single crossing* dans le plan (W_0, W_1).

[3] Il est important de noter que ceci dépend de l'hypothèse que les deux types d'agents présentent la même fonction d'utilité U. Si ceci n'est pas le cas, le résultat n'est pas valide.

Le profit de la compagnie d'assurance qui offre le contrat d'assurance $C(\pi, R)$ pour un assuré de type i est :

$$\Pi(C) = \pi - p_i R$$

La droite de profit nul dans le plan (W_0, W_1) satisfait la relation suivante : $W_1 = -\frac{1-p_i}{p_i} W_0 - S + \frac{W}{p_i}$.

Chaque assuré de type i choisit un contrat solution du programme suivant :

$$V_i(C_i) = \begin{cases} max V_i(C), \\ C \in \mathbb{E}. \end{cases}$$

Soit $\mathbb{E}$ l'ensemble des contrats simultanément offerts sur le marché. La contrainte d'incitation de l'agent i (appelée aussi contrainte d'*auto-sélection* selon laquelle le contrat d'un type de risque i donné ne doit pas attirer l'autre type de risque j) vérifie l'inégalité suivante :

$$\forall i, j \; E(U_i(C_i)) \geq E(U_i(C_j)).$$

Rothschild et Stiglitz (1976), étudiant un marché d'assurance concurrentiel pose comme condition d'équilibre la libre entrée des assureurs concurrents jusqu'à épuisement total des opportunités de profits. D'où la définition de l'équilibre RS :

Définition 1 *Un équilibre RS est une situation où :*

- *Les assureurs réalisent des profits positifs ou nuls sur chaque contrat d'assurance offert,*
- *aucun contrat offert en dehors et conjointement aux contrats d'équilibre ne serait profitable.*

Ainsi, de la maximisation du bien-être des assurés, sous les contraintes d'incitation et de profit nul, découlent les propriétés suivantes de l'équilibre RS :

Propriété 1 *L'équilibre est séparateur. Il ne peut pas être mélangeant. L'équilibre peut ne pas exister lorsque la proportion des hauts risques n'est pas suffisamment grande, c'est à dire si* $\frac{\lambda_h}{\lambda_l}$ *ne dépasse pas un certain seuil théorique* δ_{RS}.

Propriété 2 *A l'équilibre, un contrat d'assurance complète au prix actuariel est offert aux hauts risques, d'une part, et un contrat à couverture partielle au prix actuariel est offert aux bas risques, d'autre part. L'équilibre peut ne pas exister.*

1.4 L'aléa moral

En situation d'aléa moral, le principal n'observe pas les actions de l'agent après la conclusion du contrat d'assurance. La différence entre les actions et les caractéristiques inobservables est que l'agent peut modifier ses actions alors que ses caractéristiques restent inchangées, ces derniers constituant le contexte de la sélection adverse détaillée plus haut. Dans la littérature de l'assurance, il existe deux types d'aléa moral : l'aléa moral *ex ante* et l'aléa moral *ex post*. L'aléa moral *ex ante* est l'effet de l'assurance sur les activités de prévention et de protection contre la survenance d'un sinistre. L'aléa moral *ex post* est généralement référé à la fraude. Il touche les actions des assurés après la réalisation du sinistre.

Nous procédons dans ce qui suit à décrire le modèle d'aléa moral à la Arnott et Stiglitz (1988, 1991) et Arnott (1992). Des agents, identiques en ce qui concerne leurs propres caractéristiques, adoptent différemment des mesures de prévention contre la survenance d'un accident avec un niveau d'effort e. Ainsi, la probabilité d'accident est une fonction de l'effort, soit $p(e)$. Plus l'individu fournit des efforts de préventions plus la probabilité d'accident diminue.

L'effort qu'un individu peut entreprendre se résume dans les nombreuses activités et les nombreuses dépenses d'autoprotection. Donc cet effort a un coût noté $\nu(e)$.

En supposant que le coût est exprimé en terme d'utilité et non en terme monétaire[4] et en supposant que l'utilité d'un agent assuré est séparable entre la richesse et le niveau d'effort e, alors son espérance d'utilité s'écrit comme suit :

$$V(C,e) = p(e)U(W - S - \pi + R) + (1 - p(e))U(W - \pi) - \nu(e).$$

Le profit de la compagnie d'assurance qui offre le contrat d'assurance $C(\pi, R)$ est :

$$\Pi(C,e) = \pi - p(e)R.$$

Le niveau d'effort peut être soit une variable continue soit une variable discrète. Dans le dernier cas, il y aura deux niveaux niveau faible e_l et niveau élevé e_h.

Le programme de maximisation, dans un contexte d'aléa moral, se présente comme suit :

[4]Arnott et Stiglitz (1988) considèrent le coût de l'effort en terme monétaire.

$max_{C,e} V(C, e)$

sous contraintes

$\pi - p(e)R \geq 0,$

et e maximise $V(C, e)$, étant donné π, R.

Il résulte de ce programme d'optimisation que chaque agent obtient à l'équilibre le contrat d'assurance non déficitaire qui maximise son espérance d'utilité en tenant compte de son choix optimal de l'effort, et ce quelque soit la nature de la variable de l'effort, discrête ou continue.

L'agent choisit le niveau d'effort qui maximise son espérance d'utilité après avoir souscrit un contrat d'assurance. Donc e est solution de :

$$max \{V(C, e_l), V(C, e_h)\}$$

Il en ressort la courbe représentant l'ensemble des contrats laissant l'assuré indifférent entre les deux niveaux d'effort e_l ou e_h. Elle a comme équation :

$$\frac{\nu(e_l) - \nu(e_h)}{p(e_l) - p(e_h)} = U(W_1) - U(W_0)$$

Cette égalité définit la courbe qui représente l'ensemble des contrats laissant l'agent indifférent entre les deux niveaux d'effort. On en déduit que le niveau d'effort élevé e_h est fourni par l'agent à faible couverture, et un faible niveau de prévention e_l est choisi par l'agent à forte couverture d'assurance. A l'équilibre, chaque agent obtient le contrat qui maximise son espérance d'utilité, après choix optimal de l'effort, parmi les contrats réalisant une espérance de profit nulle.

1.5 Un modèle original de double asymétrie d'information : effort et aversion à l'effort

Nous considérons un modèle Principal - Agent décrivant le marché d'assurance automobile en présence de sélection adverse et d'aléa moral. Le principal représente l'assureur et l'agent appartient à un continuum d'assurés.

1.5.1 L'assuré

Chaque agent souscrit un contrat d'assurance. Il est caractérisé par son type d'aversion à l'effort que nous notons $\theta \in \Theta = [0, 1]$. Ce paramètre est une caractéristique individuelle propre à l'agent et inobservable par l'assureur. L'aversion à l'effort est considérée comme de la désutilité marginale de l'effort.

L'individu, connaissant son type θ, choisit une action $a_\theta \in A = \{\phi, E\}$ après avoir pris connaissance des contrats d'assurance offerts par l'assureur, où :

- ϕ : est l'action pour laquelle l'agent refuse l'offre de la compagnie d'assurance,
- E : est l'ensemble des niveaux d'effort après signature du contrat d'assurance. $E = \{0, e\}$. L'effort est considéré comme une variable discrète qui prend deux valeurs, soit :
  . 0 : l'assuré ne fait aucun effort de prévention. Nous le considérons comme un conducteur « *insouciant* »,
  . e : l'assuré fournit le niveau d'effort de prévention nécessaire pour se protéger contre le risque. Il est considéré comme un conducteur « *honnête* ».

L'étude porte sur un marché concurrentiel d'assurance automobile, dans lequel nous définissons un prix m comme une opportunité de marché. Nous supposons qu'il est obligatoire de s'assurer contre le risque. Nous considérons qu'aucun agent ne s'auto-assure, dans le sens où il existe une assurance minimum obligatoire pour tous les agents (le contrat responsabilité civil).

Le modèle fait intervenir une double asymétrie d'information : l'aléa moral et la sélection adverse. La sélection adverse est prise en considération dans le modèle par le fait que chaque agent connaît très bien son propre type θ et que ce dernier constitue un paramètre non observable par l'assureur. Le problème d'aléa moral a lieu dans notre modèle du moment où l'assureur n'observe pas l'effort de prévention fourni par l'agent avant l'occurrence du sinistre.

Chaque agent fait face à un risque d'accident, représenté par une perte monétaire S. Il est doté d'une richesse initiale notée W. Chaque assuré souscrit un contrat d'assurance $C = (\pi, R)$, où π représente la prime d'assurance à payer par l'assuré et R l'indemnité versée par l'assureur en cas d'accident. $S - R$ représente donc la franchise, le montant qui reste à la charge de l'assuré. Le contrat présente une clause d'exclusivité.

Les agents sont distribués en fonction de leur type θ d'aversion à l'effort. La fonction cumulative est notée $G(\theta)$. Cette distribution des agents est connue par l'assureur. Notons θ^P l'agent qui accepte de souscrire un contrat d'assurance proposé par la compagnie (P désigne la *participation*) et θ^I l'agent qui est incité à fournir de l'effort de prévention (I pour *incitation*). $G(\theta^P(\pi, R))$ est alors la proportion des agents assurés par la compagnie et $G(\theta^I(\pi, R)$ est la proportion des agents qui fournissent de l'effort de prévention e. Du moment où l'assureur connaît la distribution $G(\theta)$ des agents, alors $G(\theta^P(\pi, R))$ et $G(\theta^I(\pi, R)$ sont connus par l'assureur et par les agents.

L'ensemble des états de la nature après la conclusion d'un contrat d'assurance est $\Omega = \{z(\varepsilon), s_1, s_2\}$:

- $z(\varepsilon)$: « L'agent a un accident ». $z(\varepsilon)$ est fonction de l'effort de prévention avec $\varepsilon = 0$, ou $\varepsilon = e$. Nous interprétons $z(0)$ comme l'occurrence d'un accident responsable et $z(e)$ comme l'occurrence d'un accident non responsable.
- s_1 : « L'assuré n'a pas respecté le code de la route, a conduit en état d'ébriété, etc, et l'assureur détecte que l'assuré est responsable de l'accident ». Dans ce cas, il n'est pas indemnisé,
- s_2 : « L'assuré n'a pas respecté le code de la route, a conduit en état d'ébriété, etc, mais l'assureur ne détecte pas que l'assuré est responsable de l'accident ». Il est donc indemnisé.

Les probabilités de chaque état de la nature sont :

- $p(z(\varepsilon))$ représente la probabilité d'accident avec :
  - $p(z(0))$: la probabilité de déclarer un accident responsable sans avoir fourni l'effort de prévention, $\varepsilon = 0$,
  - $p(z(e))$: la probabilité de déclarer un accident non responsable en ayant fourni l'effort de prévention $\varepsilon = e$.
- $p(s_1) = q$ et $p(s_2) = 1 - q$ avec $0 < q < 1$. La probabilité q est supposée exogène.

Nous supposons que $p(z(0)) > p(z(e))$, dans le sens où le fait de ne pas fournir d'effort de prévention conduit à avoir une probabilité d'accident responsable supérieure à une probabilité d'accident non responsable. Dans le cas contraire, les individus n'auraient aucune incitation à faire des efforts afin de réduire la probabilité d'un accident.

L'effort de prévention peut être observable après une expertise suite à la declaration de l'accident. Quant au type θ de l'agent, il est toujours inobservable par l'assureur. Par conséquent, et d'un point de vue strictement théorique, l'assureur n'apprend rien sur le type de l'agent détecté comme « insouciant », par exemple. Par contre, il peut situer son type θ en dessus du niveau $\theta^I(\pi, R)$.

La fonction d'utilité d'un agent de type θ est définie comme suit :

$$\mathcal{U}(W, \pi, R, S, \theta, e) = U(W - \pi + R - S) - \theta e$$

L'agent choisit de fournir ou non de l'effort pour lequel son espérance d'utilité est maximum, c'est à dire :

$$E\mathcal{U}_\theta := \max\{m, E\mathcal{U}(0), E\mathcal{U}(e)\},$$

où

$$E\mathcal{U}(0) := p(z(0))\left[(1-q)U\left[W - \pi + R - S\right] + qU\left[W - \pi - S\right]\right] + (1 - p(z(0))U\left[W - \pi\right],$$

et

$$E\mathcal{U}(e) := p(z(e))U\left[W - \pi + R - S\right] + (1 - p(z(e))U\left[W - \pi\right] - \theta e.$$

$E\mathcal{U}(0)$ représente l'espérance d'utilité de l'assuré qui souscrit un contrat d'assurance et qui ne fournit aucun effort de prévention. $E\mathcal{U}(e)$ est l'espérance d'utilité de l'assuré qui souscrit un contrat d'assurance et qui fournit l'effort nécessaire de prévention. Enfin, m constitue l'espérance d'utilité de l'agent qui choisit de souscrire un contrat d'assurance offert par une autre compagnie concurrente.

Pour un niveau donné d'effort (0 ou e), les courbes d'indifférence associées chacune à un niveau d'utilité sont croissantes et concaves dans le plan (π, R).

Démonstration

- La pente d'une courbe d'indifférence d'un assuré n'étant pas incité à fournir de l'effort de prévention, dans le plan (π, R), est déterminé à partir de :

$$\frac{\delta R}{\delta \pi}|_{\overline{E\mathcal{U}(0)}} = \frac{p(z(0))\left[(1-q)U'\left[W-\pi+R-S\right]+qU'\left[W-\pi-S\right]\right]+(1-p(z(0))U'\left[W-\pi\right]}{p(z(0))(1-q)U'\left[W-\pi+R-S\right]} > 0$$

Par la suite, nous pouvons calculer $\frac{\delta^2 R}{\delta \pi^2}|_{\overline{E\mathcal{U}(0)}}$. Cette dernière est négative car $U'' < 0$.

- La pente d'une courbe d'indifférence d'un assuré qui fournit de l'effort de prévention e, dans le plan (π, R), est déterminé à partir de :

$$\frac{\delta R}{\delta \pi}|_{\overline{E\mathcal{U}(e)}} = \frac{p(z(e))U'\left[W-\pi+R-S\right]+(1-p(z(e))U'\left[W-\pi\right]}{p(z(e))U'\left[W-\pi+R-S\right]} > 0$$

et

$$\frac{\delta^2 R}{\delta \pi^2}|_{\overline{E\mathcal{U}(e)}} < 0$$

□

1.5.2 L'assureur

Le marché de l'assurance est concurrentiel. La condition du profit nul doit être alors vérifiée. Nous supposons que le comportement des agents dépend du niveau de remboursement R, et parconséquent du prix de la prime π. Les assurés qui font $\varepsilon = e$ sont au nombre de N^e. Les assurés qui ne font pas d'effort $\varepsilon = 0$ sont au nombre de N^0. Finalement, N^ϕ représente le nombre des agents qui déclinent l'offre de la compagnie d'assurance. Les agents sont hétérogènes en fonction de leur aversion à l'effort de prévention, inobservable par l'assureur.

La compagnie d'assurance propose plusieurs contrats. Ces derniers sont indexés par l'indice j par rapport aux primes d'assurance et aux indemnités qui y sont associées, avec $j = 1, \ldots, n$. Sans perte de généralité, nous classons les remboursements comme suit $R_1 < R_2 < R_3 < \ldots < R_n$.

Le profit espéré de l'assureur est donc :

$$E\Pi = \sum_{j=1}^{j=n} N^e(\pi_j - p(z(e))R_j) + \sum_{j=1}^{j=n} N^0(\pi_j - (1-q)p(z(0))R_j).$$

et la condition relative au marché concurrentiel est $E\Pi = 0$.

1.5.3 Le programme de maximisation de l'assuré

L'agent rationnel maximise son espérance d'utilité EU_θ sous les contraintes d'incitation et de participation, en tenant compte de la condition de concurrence sur le marché d'assurance. Le programme de maximisation est le suivant :

$$\max E\mathcal{U}_\theta(R)$$

$$\text{sous contrainte de :} \begin{cases} \text{Incitation} & E\mathcal{U}(e) \geq E\mathcal{U}(0), \\ \text{Participation} & E\mathcal{U}(0) \geq m, \\ \text{marché concurrentiel} & E\Pi = 0. \end{cases}$$

Nous rappelons que les agents à faible aversion à l'effort, qui fournissent l'effort de prévention nécessaire ($\varepsilon = e$) ont une espérance d'utilité :

$$E\mathcal{U}(e) := p(z(e))U\left[W - \pi + R - S\right] + (1 - p(z(e))U\left[W - \pi\right] - \theta e.$$

Cette fonction peut s'écrire sous la forme de :

$$E\mathcal{U}(e) := V(p(z(e)), R) - \theta e,$$

où

$$V(z(e), R) := p(z(e))U\left[W - \pi + R - S\right] + (1 - p(z(e))U\left[W - \pi\right].$$

L'espérance d'utilité des assurés ayant une forte aversion à l'effort et se comportant donc comme des conducteurs « *insouciants* » ($\varepsilon = 0$) est :

$$E\mathcal{U}(0) := p(z(0))\left[(1 - q)U\left[W - \pi + R - S\right] + qU\left[W - \pi - S\right]\right] + (1 - p(z(0))U\left[W - \pi\right].$$

De la même manière, cette fonction peut s'écrire somme suit :

$$E\mathcal{U}(0) := V(z(0), R)$$

Un assuré choisit rationnellement d'être vigilant et fournit de l'effort de prévention e si et seulement si :

$$\begin{cases} E\mathcal{U}(e) \geq E\mathcal{U}(0), \\ E\mathcal{U}(e) \geq m. \end{cases}$$

La première contrainte qui correspond à la contrainte d'incitation nous permet de déterminer l'ensemble des agents qui sont indifférents entre les deux niveaux d'effort de prévention 0 et e, en fonction du niveau de remboursement R offert par le contrat d'assurance. Nous notons θ_I le type de l'agent indifférent entre 0 et e. En d'autres termes, pour un remboursement R donné, l'agent ayant un niveau d'aversion à l'effort de type θ_I est indifférent entre le comportement « *insouciant* » et le comportement « *honnête* ».

À partir de la contrainte d'inictation (la première contrainte), nous déterminons ainsi l'expression de θ^I :

$$\theta^I(R) := \frac{V(z(e), R) - V(z(0), R)}{e}, \tag{1.1}$$

À partir de la deuxième contrainte qui correspond à la contrainte de participation, nous déterminons l'ensemble des agents indifférents entre partir pour une compagnie d'assurance concurrente, et souscrire un contrat et fournir de l'effort de prévention e. Nous notons le type de chacun de ces agents par θ_P qui a pour expression :

$$\theta^P(R) := \frac{V(p(z(e)), R) - m}{e}, \tag{1.2}$$

Lemme 1 *S'il existe un ensemble de θ appartenant à $[0,1]$ pour lequel $G(\theta^P(R)) \leq G(\theta^I(R))$, alors seuls les conducteurs « honnêtes » souscrivent le contrat d'assurance. Dans ce cas, le contrat est séparateur.*

Démonstration

Étant donné que G est une fonction de distribution cumulée, si $G(\theta^P(R)) \leq G(\theta^I(R))$ alors $\theta^P(R) \leq \theta^I(R)$:

$$\frac{V(p(z(e)),R) - m}{e} \leq \frac{V(p(z(e)),R) - V(p(z(0)),R,q)}{e},$$

alors :

$$V(p(z(0)),R,q) \leq m.$$

Ceci équivaut à

$$E\mathcal{U}(0) \leq m.$$

Dans ce sas, tous les conducteurs « *insouciants* » préfèrent m et souscrivent un contrat d'assurance offert par une autre compagnie concurrente. $\square$

Ainsi, nous pouvons écrire formellement :

$$\left[\theta^P(R) \leq \theta^I(R)\right] \quad \Rightarrow \quad [\varepsilon = e]\,. \tag{1.3}$$

De la même manière, les individus choisissent rationnellement de ne faire aucun effort de prévention $\varepsilon = 0$ si et seulement si :

$$\begin{cases} E\mathcal{U}(0) \geq E\mathcal{U}(e), \\ E\mathcal{U}(0) \geq m, \end{cases} \qquad \varepsilon = 0 \qquad \Longleftrightarrow \qquad \forall\theta \mid \theta^I(R) < \theta \leq \theta^P(R).$$

Lemme 2 *Pour qu'un contrat d'assurance mélangeant existe, c'est à dire un contrat choisi à la fois par les agents « insouciants » et « honnêtes », il faut que la condition suivante soit satisfaite :* $\theta^I(R) \leq \theta^P(R)$.

Démonstration

Pour tout agent ayant $\theta \leq \theta^I(R)$ alors $\varepsilon = e$. Et pour tout $\theta \in \left[\theta^I(R), \theta^I(R)\right]$ alors $\varepsilon = 0$. Par conséquent, si $\theta^I(R) \leq \theta^P(R)$ alors il exite un contrat d'assurance qui est souscrit à la fois par les « honnêtes » et par les « insouciants ». $\square$

Formellement :

$$\left[\theta^I(R) \leq \theta^P(R)\right] \quad \Rightarrow \; [\varepsilon = \{0, e\}] \,. \tag{1.4}$$

Enfin, les individus rationnels déclinent l'offre de la compagnie d'assurance si et seulement si :

$$\begin{cases} m \geq E\mathcal{U}(e), \\ m \geq E\mathcal{U}(0), \end{cases} \qquad a_\theta = \phi \qquad \Longleftrightarrow \qquad \forall \theta \mid \theta^I(R) \leq \theta^P(R) \leq \theta \leq 1.$$

1.5.4 La solution de l'assureur

Cette sous-section définit les différents niveaux qui délimitent les remboursements R proposés par l'assureur. Notons R^P le remboursement qui garantit la participation de l'agent, c'est à dire celui qui l'encourage à souscrire un contrat d'assurance. L'assureur attire des clients du moment où :

$$0 \leq \theta^P(R) < 1.$$

Ceci équivaut à

$$0 \leq V(p(z(e)), R) - m < e.$$

Comme $V(p(z(e)), R) \in \mathbb{R}$, alors le remboursement minimum de participation est :

$$\underline{R}^P = \xi(m, p(z(e))),$$

et le remboursement maximum est défini comme suit :

$$\overline{R}^P = \xi(m + e, p(z(e))).$$

avec ξ est la fonction réciproque de V.

Comme la fonction d'utilité U est croissante alors ξ l'est aussi. Nous avons donc $\underline{R}^P \leq \overline{R}^P$. Par conséquent, la contrainte de participation est croissante et convexe du moment où la fonction d'utilité est concave. Pour que les agents soient attirés par les offres de l'assureur, il faut que $\underline{R}^P \leq R^P < \overline{R}^P$.

L'assureur incite ses clients à fournir de l'effort $\varepsilon = e$ si $0 \leq \theta^I \leq 1$. En d'autres termes, il faut que

$$0 \leq V(p(z(e)), R) - V(p(z(0)), R) \leq e.$$

Notons

$$\Delta V(p(z(e)), R, q, p(z(0))) = V(p(z(e)), R) - V(p(z(0)), R).$$

Nous définissons ainsi le remboursement minimum d'incitation à l'effort $\underline{R}^I \in \mathbb{R}$ qui est égal à :

$$\underline{R}^I = \Delta\xi(p(z(e)), p(z(0))).$$

Quant au remboursement maximum $\overline{R}^I \in \mathbb{R}$, il est égal à :

$$\overline{R}^I = \Delta\xi(e, p(z(e)), p(z(0))).$$

Nous avons donc $\underline{R}^I \leq R \leq \overline{R}^I$.

L'union de ces deux ensembles $\left[\underline{R}^P, \overline{R}^P\right]$ et $\left[\underline{R}^I, \overline{R}^I\right]$ donne le meilleur ensemble d'actions possibles pour l'assureur. Il est défini par :

$$\forall \theta \in [0,1], \qquad A^* = \{R \mid \min\{\underline{R}^P, \underline{R}^I\} \leq R \leq \max\{\overline{R}^P; \overline{R}^I\}\}.$$

Nous définissons $(\tilde{\theta}, \tilde{R})$ le couple particulier d'individus $\tilde{\theta}$ et de valeur particulière du remboursement $\tilde{R}$. Il correspond à l'intersection de la contrainte de participation et de la contrainte d'incitation. Le remboursement $\tilde{R}$ est donc déterminé comme suit :

$$\theta^P(\tilde{R}) = \theta^I(\tilde{R}) \text{ ce qui équivaut à } V(p(z(0)), \tilde{R}) = m$$

alors

$$\tilde{R} = \tilde{\xi}(m, z(0))$$

En d'autres termes, $\tilde{\theta}$ correspond au niveau particulier d'aversion à l'effort au delà duquel il existe au moins un assuré qui ne serait pas incité à fournir de l'effort de prévention. $\tilde{R}$ est donc le remboursement qui laisse les « *insouciant* » indifférents entre accepter un contrat proposé par la compagnie d'assurance et ne pas fournir d'effort de prévention, ou souscrire un contrat d'assurance offert par la concurrence.

Nous représentons la contrainte d'incitation et la contrainte de participation dans les figures suivantes 1.1 et 1.2.

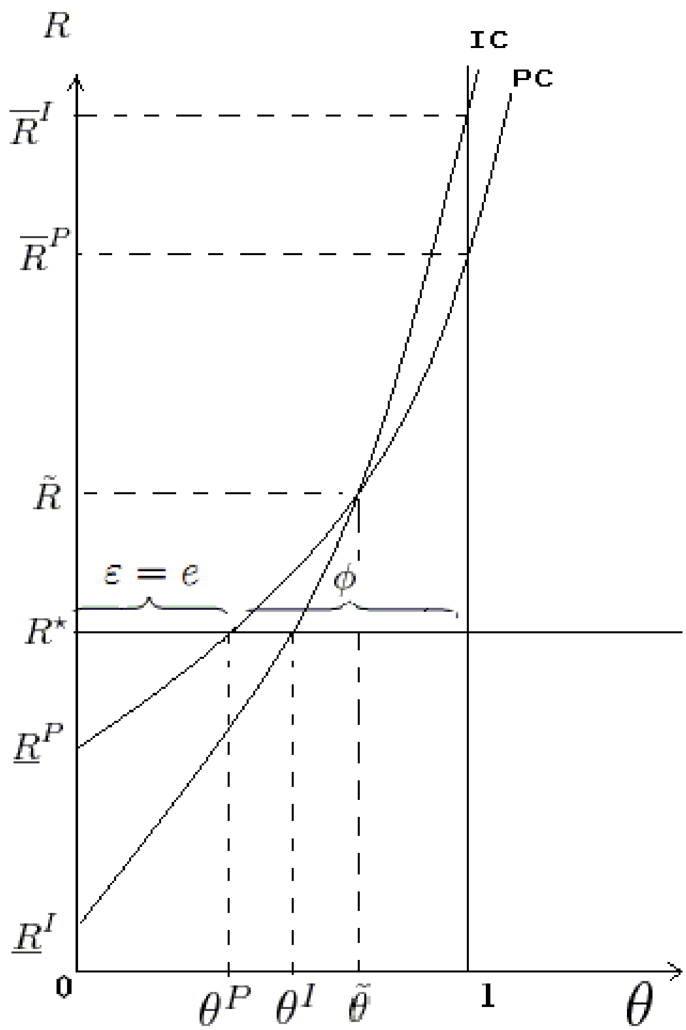


FIG. 1.1 – **Le contrat séparateur**

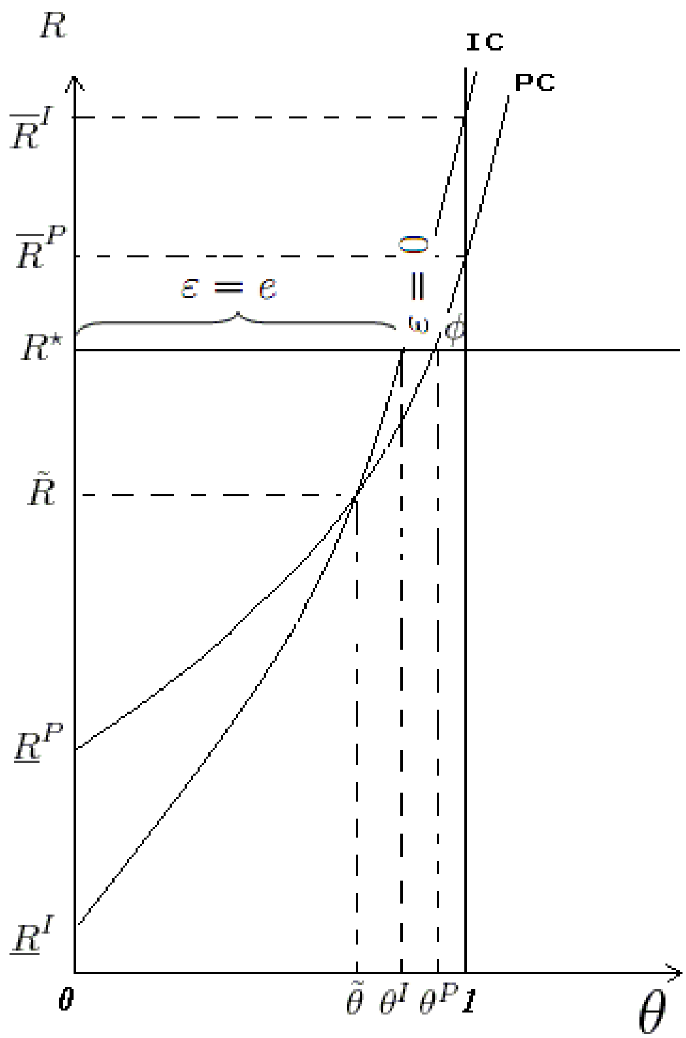


FIG. 1.2 – **Le contrat mélangeant**

Les figures 1.1 et 1.2 représentent respectivement dans le plan (θ, R), le cas d'un contrat séparateur $\boxed{\text{S}}$ et celui d'un contrat mélangeant $\boxed{\text{P}}$. Quelque soit le couple (θ, R), la pente de la contrainte de participation (1.2) et celle de la contrainte d'incitation (1.1) sont positive. Elles ne sont pas égales dans l'espace (θ, R) en raison de la présence de la probabilité q . Plus précisément, la pente de la contrainte d'incitation est supérieure à celle de la contrainte de participation.

1.5.5 Les contrats d'équilibre

La compagnie d'assurance propose plusieurs contrats. Ces derniers sont indexés par l'indice j par rapport aux primes d'assurance et aux indemnités qui y sont associées, avec $j = 1, \ldots, n$. Sans perte de généralité, nous classons les remboursements comme suit $R_1 < R_2 < R_3 < \ldots < R_n$.

L'assureur résoud l'équation suivante pour laquelle le profit espéré est égal à zéro :

$$E\Pi = 0$$

Ceci équivaut à :

$$\sum_{j=1}^{j=n} N^e(\pi_j - p(z(e))R_j) + \sum_{j=1}^{j=n} N^0(\pi_j - (1-q)p(z(0))R_j) = 0. \tag{1.5}$$

Cette condition nous permet de définir le remboursement $R_j^\star$ comme une fonction de la prime $\pi_j^\star$. Selon les valeurs appropriées de N_j^e et $N_j^0, j = 1, \ldots, n$, deux types de contrats sont possibles : les contrats séparateurs et les contrats mélangeants.

1.5.5.1 Le contrat séparateur

Nous notons dans ce qui suit R_S le remboursement défini par le contrat séparateur.

A partir de la solution de l'assuré 1.3, nous savons que si $\theta^P \leq \theta^I$, alors tous les agents qui acceptent d'être assurés par la compagnie sont incités à fournir de l'effort e de prévention. Ceci équivaut à souscrire un parmi les contrats R_S tel que $R_S < \tilde{R}$. Pour que θ^P soit inférieur à θ^I, il faut que le type de chaque assuré, donc son niveau d'aversion à l'effort, appartienne à $\left[0, \tilde{\theta}\right]$.

Si les contrats d'assurance souscrits sont des contrats séparateurs, alors tous les assurés de la compagnie d'assurance $G(\theta^P)$ correspondent à tous les agents « *honnêtes* » qui fournissent de l'effort e et qui sont au nombre de N^e, alors :

$$N^e = G(\theta^P(R_s)) = G(\frac{V(p(z(e)), R_j) - m}{e}),$$

Ainsi, le nombre N^0 des agents « insouciant » qui ne fournissent aucun effort de prévention ($\varepsilon = 0$) soit nul :

$$N^0 = G(\theta^P(R_j)) - G(\theta^I(R_j)) = 0$$

Nous pouvons donc identifier l'ensemble des contrats qui procurent une espérance de profit nul à la compagnie d'assurance. Par conséquent :

$$(\pi_s - p(z(e))R_s^\star)G(\theta^P) = 0.$$

Alors, la solution de l'assureur est définie par :

$$R_s^\star = \frac{\pi_s^\star}{p(z(e))}.$$

La valeur du remboursement est ainsi une fonction croissante de la prime d'assurance et décroissante en fonction de la probabilité d'un accident non responsable.

Pour un contrat séparateur, nous devons donc avoir

$$\underline{R}^P \leq R_s^\star \leq \tilde{R},$$

Cet ensemble de contrats d'assurance correspond à la portion de la droite délimitée par $\tilde{R}$ au delà duquel il existerait des agents « *insouciants* », et $\underline{R}^P$ en dessous duquel aucunt agent ne souscrirait de contrat d'assurance proposé par la commpagnie. Cette portion de droite a une pente $\frac{1}{p(z(e))}$ (supérieure à 1) dans le plan (π, R) . L'ensemble de ces contrats est illustré graphiquement sur la figure 1.3.

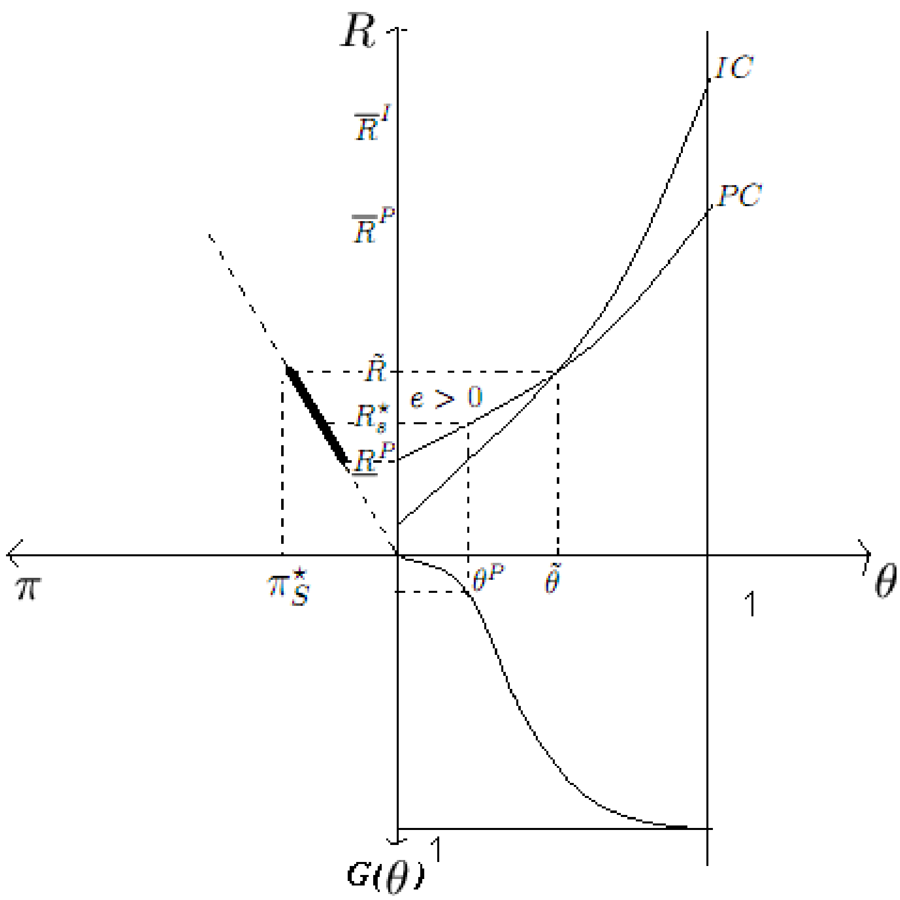


FIG. 1.3 – **L'ensemble des contrats séparateurs**

1.5.5.2 Le contrat mélangeant

Nous notons R_P le remboursement offert par un contrat mélangeant. D'après la solution de l'assuré 1.4, si $\theta^I \leq \theta^P(R)$, alors il existe au moins un contrat d'assurance offert par la compagnie qui attire à la fois des agents « *honnêtes* » et des agents « *insouciants* » qui préfèrent ne fournir aucun effort de prévention. Dans ce cas, le nombre d'agents N^e qui fournissent de l'effort correspond à $G(\theta^I)$ et le nombre d'agents N^0 qui produisent zéro effort représente le reste des assurés, soit $G(\theta^P) - G(\theta^I)$.

Nous pouvons donc identifier l'ensemble des contrats qui procurent une espérance de profit nulle en résolvant l'equation suivante :

$$[\pi - p(z(e))R_P]\, G(\theta^I) + [\pi - (1-q)p(z(0))R_P] \left[G(\theta^P) - G(\theta^I)\right] = 0.$$

Ce qui équivaut à :

$$R_p = \pi \; \frac{G(\theta^P)}{p(z(e))G(\theta^I) + (1-q)p(z(0))(G(\theta^P) - G(\theta^I))}$$

Cet ensemble de contrat d'assurance est défini par les remboursements R_p en fonction de la prime π. Il correspond à la proportion de la droite délimitée par la valeur inférieure de remboursement $\underline{R}^P$ et le remboursement maximum de participation $\overline{R}^P$. Cette droite est croissante car la pente est supérieure à zéro :

$$\frac{G(\theta^P)}{p(z(e))G(\theta^I) + (1-q)p(z(0))(G(\theta^P) - G(\theta^I))} > 0,$$

puisque $G(\theta^P) > G(\theta^I)$.

Cette pente est supérieure à celle de la droite qui représente l'ensemble des contrats séparateurs.

Démonstration

$$\frac{G(\theta^P)}{p(z(e))G(\theta^I) + (1-q)p(z(0))(G(\theta^P) - G(\theta^I))} - \frac{1}{p(z(e))}$$

$$= \frac{(p(z(e)) - (1-q)p(z(0)))(G(\theta^P) - G(\theta^I)}{p(z(e))\left[p(z(e))G(\theta^I) + (1-q)p(z(0))(G(\theta^P) - G(\theta^I))\right]} > 0$$

$\square$

La différence entre les pentes est positives car $G(\theta^P) > G(\theta^I)$.

Nous représentons sur la figure 1.4 l'ensemble des contrats mélangeants.

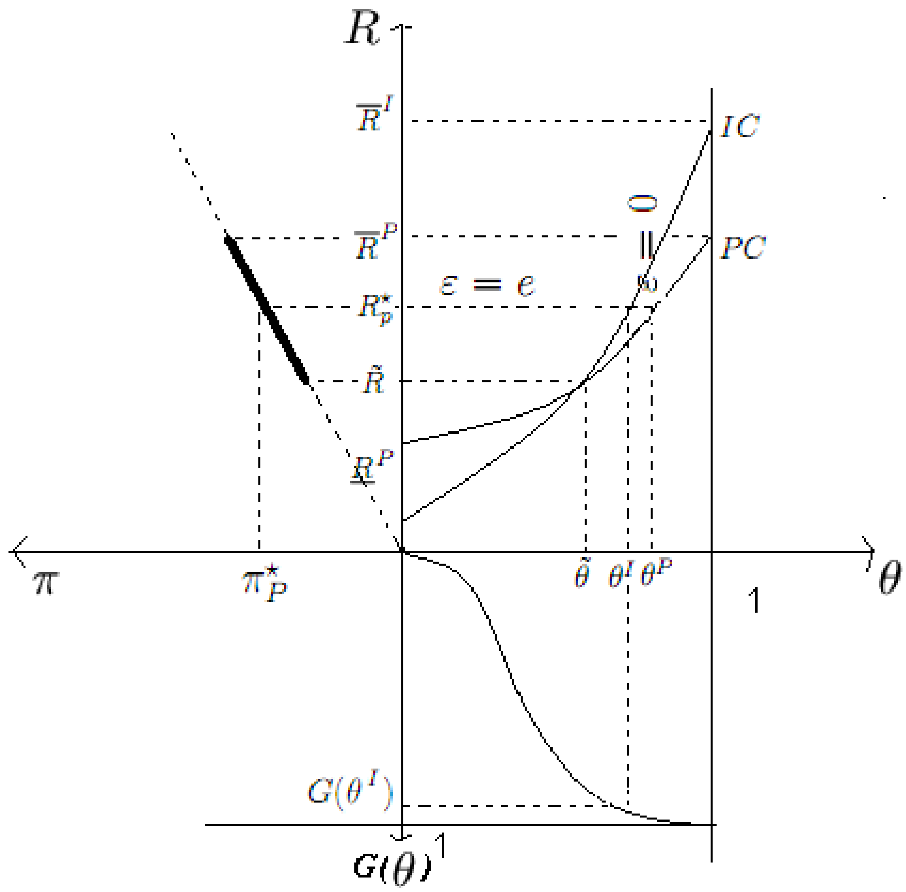


FIG. 1.4 – **L'ensemble des contrats mélangeants**

L'union de ces deux portions de droites représentant l'ensemble des contrats séparateurs et celui des contrats mélangents représentent tous les contrats réalisant une espérance de profit nulle. Chaque agent caractérisé par son type d'aversion à l'effort, à l'équilibre, obtient le contrat appartenant à cet ensemble de contrat d'assurance et maximisant son espérance d'utilité.

1.6 Conclusion

Les premières sections de ce chapitre ont constitué une partie préliminaire, où nous avons rappelé les propriétés de l'équilibre d'un marché d'assurance concurrentiel en présence, d'une part, de sélection adverse, et d'autre part, d'aléa moral.

En situation d'antisélection, le contrat d'équilibre à la Rothschild et Stiglitz proposé aux hauts risques est un contrat actuariel d'assurance complète, ce qui n'est pas le cas pour les bas risques à qui est offert un contrat à couverture partielle. L'équilibre, s'il existe, est un eqiulibre séparateur.

Dans la seconde situation où il existe de l'aléa moral, le marché propose à l'équilibre un contrat d'assurance dont le montant de la couverture dépend du niveau du coût de la prévention. Si le coût de la prévention est élevé, l'agent souscrit une assurance totale. Sinon, le marché propose un contrat d'assurance partielle. L'équilibre, dans ce cas de figure, est séparateur.

Nous avons enfin présenté un modèle où existent simultanément les deux problèmes d'asymétrie d'information. Nous avons considéré que les agents se différencient par leur niveau d'aversion à l'effort de prévention. Cette dernière définit le risque intrinsèque de l'agent. Elle constitue une caractéristique propre à chaque agent, inobservable par l'assureur. A un contrat d'assurance donné, chaque agent choisit un niveau d'effort de prévention que l'assureur n'observe pas.

Nous avons montré que si l'aversion à l'effort est suffisamment faible, alors tous les agents choisissent un contrat avec la plus faible couverture qui les incitent à fournir l'effort de prévention nécessaire pour se prémunir contre le risque. Nous pouvons penser notamment au contrat obligatoire de responsabilité civile qui offre le remboursement minimum aux assurés.

Si les agents se caractérisent par des niveaux plus élevés d'aversion à l'effort, alors ils choisissent des contrats à très forte couverture. En ce point, l'équilibre vérifie certaines propriétés obtenues par le modèle de Rothschild et Stiglitz (1976), à savoir l'équilibre est séparateur. Cependant, nous avons montré qu'un équilibre mélangeant peut exister à partir d'un certain seuil de remboursement d'assurance. Au delà de ce seuil, un même contrat d'assurance est souscrit par une proportion d'agents incités à fournir de l'effort et une autre proportion d'agents ayant un niveau d'aversion à l'effort plus élevé et ne fournissant aucun effort de prévention. Par conséquent, d'après notre modèle, il peut exister des contrats d'équilibre séparateurs mais aussi des contrats d'équilibre mélangeants.

A partir de ce résultat théorique, nous pouvons avancer que s'il existe une relation positive ou négative entre la couverture d'assurance et le risque de l'agent, l'asymétrie d'information existe sur le marché d'assurance. L'approche empirique de la corrélation positive entre le risque et la couverture se base seulement sur l'équilibre séparateur. Nous avons montré que l'équilibre mélangeant peut exister et donc une corrélation négative entre la couverture d'assurance et le risque ne signifie pas que l'hypothèse de l'asymétrie d'information est rejetée.

Ces résultats théoriques vont donc être validés empiriquement dans le chapitre suivant. Pour ce faire, nous utilisons un fichier d'assurés contenant des données individuelles d'une grande mutuelle d'assurance française. Il est nécessaire, avant de valider empiriquement nos hypothèses théoriques, de procéder à une analyse de la sinistralité des assurés en fonction de toutes les informations que l'assureur détient et qui sont aussi à notre disposition.

Chapitre 2

Etude multivariée de la relation *risque-couverture*

2.1 Introduction

Au sein d'un portefeuille d'assurance automobile, les assurés ne sont pas tous égaux devant le risque. En effet, certains présentent un profil plus dangereux que d'autres. Du fait de cette hétérogénéité, l'assureur est contraint de segmenter le portefeuille afin de constituer des classes de risques, pour ainsi proposer au souscripteur potentiel le meilleur contrat d'assurance en adéquation avec son profil. Il se fonde pour cela sur toutes les informations disponibles et observables liées au conducteur (sexe, age, ancienneté du permis de conduire...), au véhicule assuré (ancienneté du véhicule, puissance...), à la zone de circulation, à l'usage du véhicule,... etc. Malgré cela, il peut subsister une certaine hétérogénéité au sein de chaque classe, les facteurs observables n'expliquant pas complètement le risque des assurés.

L'assureur essaye donc d'appréhender, d'estimer les risques supplémentaires liés à l'asymétrie d'information à l'aide des renseignements demandés au moment de la souscription du contrat et du suivi du comportement de l'assuré au niveau de sa sinistralité. Pour faire face aux problèmes informationnels dans un environnement concurrentiel, la compagnie d'assurance adopte des stratégies lui permettant, d'une part, de garder le plus longtemps possible au sein de son portefeuille les bons risques dont la probabilité de sinistres est faible, et d'autre part, d'attirer les bons risques des autres compagnies d'assurance.

La concurrence entre les assureurs français se joue de nos jours sur le coefficient de Bonus Malus (appelé aussi Coefficient de Réduction Majoration (CRM)). Nous constatons en effet différentes approches qui vont en ce sens :

- Le *Super Bonus* : c'est une réduction supplémentaire de la prime liée au nombre d'années (3 ans minimum) où l'assuré a un bonus égal à 50. Le

Bonus à vie et le *Bonus longue durée* reposent aussi sur le même principe.

- La suppression du Malus pour les assurés qui ont un CRM égal à 50 et qui ont un seul accident responsable sur une période de 3 ans.

Il existe, bien évidemment, d'autres offres du même type que nous n'énonçons pas.

Toutes ces offres sont proposées aux sociétaires ayant atteint le bonus maximum, à savoir un coefficient de 50. Les assureurs considèrent automatiquement ces sociétaires comme les seuls bons risques. De leur point de vue, les assurés qui n'ont pas atteint le bonus maximum ne sont pas considérés comme de bons conducteurs, alors qu'ils peuvent l'être. Plus précisément, il faut 13 années de conduite sans accident responsable pour atteindre un CRM de 50. Un assuré qui a, par exemple, 10 ans de conduite sans accident responsable ne peut donc pas avoir un CRM égal à 50. Et pourtant, il est exclu par l'assureur de la catégorie des bons risques. Ce simple exemple montre bien qu'en s'appuyant uniquement sur le niveau du coefficient du bonus malus sans prendre en compte le nombre d'années d'expérience de conduite, certains bons conducteurs ne sont pas considérés par l'assureur.

Ce chapitre à trois principaux objectifs[1] :

- Le premier point consiste en la construction d'un indicateur de la sinistralité passée. A partir de données qui concernent l'année 2004, nous estimons l'effet prédictif de cette nouvelle variable sur la sinistralité future et nous évaluons sa pertinence par rapport à la variable du bonus malus.

[1]Une partie de ce chapitre a fait l'objet d'un travail en collaboration avec Mr. Michel Grun-Rehomme

- Une *proxy* de l'aversion à l'effort est créée à partir de l'indicateur de la sinistralité passée. Cette *proxy* est par la suite utilisée pour tester empiriquement les prédictions théoriques que nous avons établies dans le premier chapitre.

- Enfin, une étude empirique des problèmes informationnels est menée, en utilisant ces nouvelles variables dans l'approche de la corrélation conditionnelle entre le risque et la couverture d'assurance.

Le présent chapitre est constitué par la suite d'une section présentant les principaux travaux empiriques traitant les problèmes d'asymétrie d'information sur le marché d'assurance automobile. Ceci nous permet de situer nos résultats par rapport à la littérature. Une troisième section présente la base des données dont nous disposons, ainsi que les principales statistiques descriptives. La section suivante étudie la corrélation entre la couverture d'assurance et le risque, en utilisant des modèles univariés et bivariés. L'indicateur de la sinistralité passée est décrit dans la cinquième section. Nous y estimons un modèle probit trivarié afin d'étudier les problèmes d'endogénéité et de vérifier ensuite l'hypothèse de l'asymétrie d'information. Nous introduisons dans la sixième section, à partir de l'indicateur de la sinistralité passée, une *proxy* de l'effort qui permet de tester notre modèle théorique de principal-agent présenté dans le premier chapitre. Enfin, la dernière section présente les conclusions.

2.2 Les études empiriques de l'asymétrie d'information

L'étude empirique des problèmes d'asymétrie d'information a pris du retard par rapport aux travaux théoriques. Et ce n'est qu'à partir des années 90 qu'elle a sensiblement évolué. Les principales études empiriques des problèmes d'information sur le marché d'assurance se basent essentiellement sur l'approche de la corrélation conditionnelle. Cette approche vise à tester l'hypothèse d'une corrélation positive entre la couverture et le risque, conditionnellement à toutes les informations symétriques observables (toutes les informations qu'un assureur détient sur ses assurés) dans le contrat d'assurance. Chiappori et *al.* (2006) montrent que la propriété de la *corrélation positive* est robuste et peut être adoptée comme une procédure standard à l'étude empirique de l'asymétrie d'information.

Cette propriété est fondée sur les prédictions qui résultent des travaux théoriques sur la sélection adverse et l'aléa moral. Ces prédictions stipulent, comme nous l'avons souligné plus haut, que, d'une part, en situation de sélection adverse, un agent à haut risque choisit une couverture plus large que celle d'un agent à bas risque (Rothschild et Stiglitz (1976)). D'autre part, dans un contexte d'aléa moral, un agent qui souscrit un contrat avec une large couverture a tendance à fournir moins d'effort de prévention que celui qui a une plus faible couverture, et donc il devient plus risqué (Shavell (1979), Holmstrom (1979), Arnott and Stiglitz (1988)). De ces deux situations d'asymétrie d'information ressort clairement la correlation positive entre le risque et la couverture d'assurance. Ainsi, pour distinguer entre le problème de sélection adverse et celui d'aléa moral, il faut que le sens de causalité entre le risque et la couverture soit connu.

Parmi les premiers travaux étudiant l'asymétrie d'information sur le marché de l'assurance automobile, nous citons Dahlby (1983, 1992). Ce dernier montre l'existence de la sélection adverse sur le marché canadien d'assurance automobile. Ce résultat a été critiqué. En effet, Dahlby (1983, 1992) utilise dans sa modélisation moins d'information qu'un assureur a effectivement en disposition, ce qui biaise nécessairement son résultat.

Puelz et Snow (1994) trouvent le même résultat que Dahlby en utilisant des données provenant d'une assurance automobile américaine, à savoir l'existence d'une corrélation significativement positive entre le risque et la couverture d'assurance. Mais depuis, leur étude est remise en question. Une première critique aux interprétations de Puelz et Snow est que leur résultat ne traduit pas seulement de la sélection adverse, mais il est aussi conforme à l'hypothèse d'aléa moral qui peut exister sur ce marché d'assurance automobile (Chiappori 1999). La deuxième critique est liée aux problèmes économétriques du modèle. Dionne (1998) confirme que, malgré leurs pertinentes questions empiriques, le modèle économétrique de Puelz et Snow (1994) n'est pas bien spécifié. Les deux auteurs « n'ont pas considéré tous les instruments disponibles à l'assureur pour tenir compte de l'antisélection de façon efficace »[2]. Dionne, Gourriéroux et Vanasse (2001) mettent en évidence des effets de non-linéarités non pris en considération par Puelz et Snow dans leur équation du choix de la franchise. Ils montrent qu'utiliser le nombre d'accidents estimé[3] comme variable explicative de la franchise, pour tenir compte des effets de non-linéarité, donne des résultats différents de ceux de Puelz et Snow (1994). Ils confirment que si le modèle n'est pas correctement spécifié, de fausses conclusions peuvent être formulées sur l'hypothèse de la sélection adverse résiduelle en assurance automobile. Ils soulignent aussi que l'asymétrie d'information résiduelle peut être contrôlée par l'assureur en adoptant une procédure adéquate de classification donnant des groupes homogènes de risque.

Chiappori et Salanié (2000) proposent une alternative au test de l'asymé-

[2]Dionne, G. (1998) *Les mesures empiriques des problèmes d'information*, Cahier de recherche 98-16, HEC Montreal

[3]Les auteurs estiment une régression négative binomiale du nombre d'accidents en fonction des variables de tarifications. Cela leur permet par la suite de calculer le nombre d'accidents estimé

trie d'information. Ils montrent que l'asymétrie d'information est inexistante. Ils utilisent des données individuelles d'assurance automobile française[4]. Ils se concentrent sur les jeunes conducteurs ayant moins de trois ans d'expérience de conduite. Les auteurs considèrent deux modèles probit. Le premier estime le niveau de couverture d'assurance et le deuxième modélise l'occurrence d'un accident déclaré, conditionnellement à toutes les caractéristiques de l'assuré. Afin de conclure sur l'existence ou non de l'asymétrie d'information, ils testent, en premier lieu, l'hypothèse de la nullité de la covariance entre les termes d'erreurs, en estimant séparément les deux equations. Ensuite, ils estiment de manière simultanée les deux équations avec un modèle probit bivarié afin de tester la corrélation entre les résidus. Dans les deux cas, l'hypothèse nulle (covariance nulle ou corrélation nulle) signifie l'inexistence d'asymétrie d'information. Chiappori et Salanié (2000), en rejetant l'hypothèse alternative, indiquent qu'il n'y aucune preuve d'asymétrie d'information.

Chiappori et Salanié (2000) soulignent que ce test d'asymétrie d'information ne permet pas de distinguer entre la sélection adverse et l'aléa moral, dans le sens où il n'est pas possible de distinguer la direction de la causalité entre le risque et la couverture. En effet ceci constitue l'une des limites de l'approche de corrélation positive, et plus particulièrement quand l'étude est appliquée sur une base de données en coupe instantanée, observée sur une seule période de contrat d'assurance.

Ces auteurs concluent leur étude par un test spécifique à la détection de l'aléa moral. Ce test repose sur l'expérience passée des jeunes conducteurs assurés, en utilisant le coefficient du bonus malus. Ils exploitent l'idée que, dans le système

[4]Ces données proviennent de la Fédération Française des Sociétés d'assurance (FFSA). Elles concernent l'année 1989 contenant 1 120 000 polices avec 120 000 accidents.

français de bonus malus, les jeunes conducteurs peuvent commencer avec le bonus maximum 50, en s'inscrivant en tant que conducteurs secondaires sur le véhicule de leurs parents et ainsi bénéficier de leur coefficient de bonus malus. Chiappori et Salanié rajoutent donc une variable indicatrice décrivant si le jeune conducteur bénéficie du bonus maximum de ses parents, dans les deux modèles probit de la couverture d'assurance d'une part et de l'occurrence d'un accident d'autre part. Ils montrent, d'abord, que le fait d'avoir un coefficient égale à 50 encourage la souscription de contrats à couvertures complètes. Ils soulignent dans un second temps qu'il peut y avoir trois sortes d'impacts de la reduction maximale sur la probabilité d'accident :

- une corrélation négative entre le fait d'avoir un coefficient de 50 et la probabilité d'accident,
- une corrélation nulle et donc l'aléa moral n'existe pas,
- une corrélation positive entre le coefficient de bonus malus et la probabilité d'accident. Ceci signifie que plus le coefficient du bonus malus est faible, plus les efforts de prévention diminuent.

L'hypothèse de l'aléa moral est vérifiée dans ce troisième cas de figure. Chiappori et Salanié trouvent que la corrélation est significativement négative et donc il n'existe pas d'aléa moral chez les jeunes conducteurs.

Cependant, Richaudeau (1999) trouve des résultats partiellement différents de ceux de Dionne, Gourriéroux et Vanasse (2001) et de Chiappori et Salanié (2000). Il utilise des données individuelles de l'année 1995, issues d'une enquête faite par l'Institut National de la Recherche sur les Transports et leur Sécurité (L'INRETS)[5]. Cette base de données présente l'avantage de contenir plus d'informations que celles détenues par les assureurs, comme par exemple le nombre de

[5] Il se base sur 5703 observations.

kilomètres parcourus. Afin d'étudier la corrélation entre le nombre d'accidents et le choix du contrat d'assurance, il utilise en premier lieu une régression probit sur le choix d'un contrat *touts risques* par rapport à un contrat de responsabilité civile. Le terme d'erreur de cette équation est ensuite utilisé comme variable explicative du nombre d'accidents ajusté par un modèle binomial négatif. Si ce terme d'erreur explique significativement le nombre d'accidents, alors l'hypothèse alternative de l'asymétrie d'information est acceptée. Richedeau rejette cette hypothèse. Par contre, en excluant la variable du nombre de kilomètres et en n'introduisant que les variables tarifaires utilisées par l'assureur, Richeaudeau montre que l'hypothèse d'asymétrie d'information ne peut pas être rejetée. Il explique, à partir de ce résultat que les agents les plus couverts ont plus de mobilité que les autres. Et puisqu'ils sont plus exposés aux dangers de la circulation, ils deviennent plus risqués. Ainsi, il confirme que ce comportement n'est pas le résultat de la nature du contrat d'assurance et il ne traduit donc pas de l'aléa moral mais plutôt de la sélection adverse.

D'autres auteurs reprennent par la suite le même modèle de détection de l'asymétrie d'information que Chiappori et Salanié (2000), à savoir le modèle probit bivarié s'appuyant sur l'approche de corrélation positive. Nous citons dans ce cas l'étude de Cohen (2005). L'auteur confirme le résultat de Chiappori et Salanié (2000) en ce qui concerne les jeunes conducteurs. Cohen (2005) utilise des données israéliennes d'assurance automobile[6]. La base de données contient des informations différentes par rapport à celles qu'un assureur français peut détenir, en particulier celles concernant l'historique de la sinistralité. Ce sont les assurés qui rapportent eux même l'information des accidents déclarés dans le passé, ce qui peut diminuer la crédibilité de cette variable. La qualité d'une telle variable

[6]La base de données concerne 104.639 assurés.

peut introduire des biais dans les résultats. En France, le système réglementé du bonus malus impose la transparence entre les assureurs concernant le passé des déclarations des assurés ce qui garantit sa fiabilité. L'auteur étudie non seulement la population des jeunes conducteurs mais il l'étend aux conducteurs expérimentés qui ont plus de trois ans d'expérience de conduite. En utilisant le modèle probit bivarié à la Chiappori et Salanié (2000) et en testant la corrélation entre couverture et sinistres, Cohen (2005) trouve que contrairement aux jeunes conducteurs, l'hypothèse de l'asymétrie d'information n'est pas rejetée chez les conducteurs expérimentés. L'auteur explique ce résultat par l'effet de l'apprentissage que les conducteurs expérimentés acquièrent tout au long de leur expérience de conduite, leur permettant de développer de l'information privée sur leur propre type de risque.

Nous remarquons ainsi que la plupart des études sus-cités, exceptée celle de Cohen (2005) ne confirment pas l'existence de l'asymétrie d'information sur le marché d'assurance. L'un des points communs que nous pouvons observer dans ces travaux est le choix du contrat d'assurance. En effet, tous ces auteurs utilisent deux types de contrat d'assurance : faible couverture (i.e. assurance obligatoire) contre forte couverture (i.e. assurance *tous risques*). En réalité, la compagnie d'assurance propose plusieurs contrats avec plusieurs couvertures optionnelles et différentes franchises. Considérer cette multitude d'offres dans l'approche de la corrélation positive permet de se rapprocher de la réalité. C'est ainsi que nous procédons dans ce chapitre. Une minorité de travaux ont fait ainsi, nous citons l'étude de Grun-Rehomme et Benlagha (2007) et Kim et *al.* (2009).

Grun-Rehomme et Benlagha (2007) ont appliqué le même modèle que Chiappori et Salanié (2000) en modélisant, dans l'équation du choix de la couverture, quatre différents contrats d'assurance (un contrat de responsabilité civile et trois contrats *tous risques* proposant différents niveaux de franchises). En utilisant des données françaises de l'année 2004 qui concernent les jeunes conducteurs[7], ils montrent que la corrélation entre les résidus de l'équation des sinistres et ceux de l'équation des couvertures d'assurance est significativement positive. Ils confirment que certes il n'existe pas de l'asymétrie d'information chez les jeunes conducteurs choisissant entre une couverture de responsabilité civile et une couverture *tous risques*. Cependant, ils ont montré que des problèmes informationnels peuvent subsister entre l'assureur et les assurés qui choisissent entre plusieurs couvertures *tous risques* à différents niveaux de franchise, ce qui peut se traduire par la sélection adverse et/ou l'aléa moral.

Kim et *al.* (2009), en reprenant le modèle de Richaudeau (1999), considèrent trois niveaux de couvertures d'assurance. Les données qu'ils utilisent concernent une seule année de contrat d'assurance. Elles proviennent d'une compagnie d'assurance coréenne[8]. Ils appliquent en premier lieu une régression sur la couverture d'assurance, mais au lieu d'appliquer un probit binomial comme dans le modèle de Richaudeau, ils appliquent un probit ordonné vu la nature de la variable de couverture qu'ils utilisent. Ils introduisent ensuite les résidus de cette équations comme variable explicative dans la régression binomiale négative du nombre de sinistres. Il en résulte une corrélation significative entre le terme d'erreurs et la sinistralité. Ainsi, Kim et *al.* (2009) ne rejettent pas l'hypothèse de l'asymétrie d'information sur le marché d'assurance coréen.

[7]Le fichier contient 28.148 observations concernant les jeunes conducteurs qui ont moins de trois ans d'expérience de conduite.

[8]Les assureurs coréens utilisent les mêmes variables de tarification qu'un assureur français détient. Comme en France, les pouvoirs publiques coréens imposent le système du bonus malus aux sociétés d'assurance.

Un deuxième point commun à ces travaux peut être souligné. Toutes les études que nous avons exposées plus haut utilisent des données en coupe instantanées qui ne décrivent q'une seule période de contrat d'assurance. De nombreux auteurs (Chiappori (2000), Chiappori et Salanié (2000), Chaippori et al. (2004)) ont souligné, qu'en se basant sur de telles données, l'approche de la corrélation positive entre la couverture d'assurance et les sinistres déclarés permet seulement de conclure sur la présence de l'asymétrie d'information sans pouvoir faire la distinction entre la sélection adverse et l'aléa moral. Avec des données statiques, il n'est pas facile d'identifier le sens de causalité entre le risque et la couverture. Ce n'est qu'avec des données multidimensionnelles, décrivant plusieurs années de contrat d'assurance, que les deux formes d'asymétrie d'information peuvent être identifiées (Chiappori, 2000).

Dionne (1998) suggère que la sélection adverse concerne les caractéristiques propres aux individus, donc exogènes. Tandis qu'en situation d'aléa moral, les actions sont endogènes, elles peuvent être modifiées en tout temps. Ainsi, s'appuyer sur des données dynamiques permet de séparer l'antisélection de l'aléa moral .

Chiappori et al. (2006) montrent qu'en situation d'aléa moral, l'intensité des déclarations individuelles des sinistres décroit avec le nombre des accidents passés. L'idée est qu'en présence d'aléa moral, un accident responsable déclaré à l'année t augmente la prime à l'année t+1, en raison du coefficient du bonus malus. Il s'ensuit que l'occurrence d'un accident a un effet positif sur les incitations du conducteur à la prudence, et par conséquent la probabilité d'un accident futur diminue. En utilisant des données d'assurance française et en se basant sur la relation négative de l'occurrence des sinistres, ils montrent que l'hypothèse de présence d'aléa moral n'est pas vérifiée.

Abbring et *al.* (2003(a)) utilisent des données longitudinales qui concernent

79 684 polices de l'année 1987 à l'année 1989 et qui proviennent d'une compagnie d'assurance française. Ils adoptent le modèle de Heckman et Borjas (1980) pour tester l'hypothèse de la sélection adverse. Ils se basent sur la même idée de corrélation négative énoncée par Chiappori et *al.* (2006) et ils montrent qu'il n'existe aucune preuve d'aléa moral.

Dionne, Michaud et Dahchour (2006) soulignent que la corrélation négative détaillée par Chiappori et al. (2006) n'est valide que si l'effet du bonus-malus sur les incitations est de court terme. Dionne, Michaud et Dahchour (2006) font intervenir dans leur étude dynamique non seulement les sinistres mais aussi le choix du contrat. Ils utilisent des données de panel françaises sur trois ans. Ils analysent les prédictions théoriques suggérant que, dans une régression dynamique des sinistres :

- en situation d'aléa moral, la corrélation entre la couverture en t-1 et les accidents déclarés en t est positive,
- en situation de sélection adverse, la corrélation entre la couverture en t-1 et les accidents déclarés en t est nulle.

Dionne, Michaud et Dahchour (2006) valident empiriquement l'hypothèse d'aléa moral et distinguent cette asymétrie d'information de la sélection adverse en utilisant un test de causalité de Granger (Granger, 1969).

Cohen et Einav (2007) utilisent une méthode différente de celle de Chiappori et Salanié (2006). En se basant sur des données de panel israélen, Cohen et Einav étudient la relation entre le risque et l'aversion au risque. Ils utilisent l'espérance d'utilité théorique d'un individu assuré, exprimée en fonction de la prime, du niveau de la franchise et de la probabilité d'accident. Ils déterminent à partir de cette espérance d'utilité un paramètre de l'aversion absolue au risque en fonction des sinistres. En estimant un modèle structurel conditionnel à toutes les caractéristiques observables des individus, Cohen et Einav trouvent une corrélation positive entre l'aversion au risque et la sinistralité. Ils montrent que cette corrélation positive confirme l'existence de la sélection adverse et/ou de l'aléa moral sur le marché israélien d'assurance automobile.

2.3 Les données

Toute compagnie d'assurance dispose d'un fichier clients qui enregistre des quantités importantes d'informations. Avant d'opter pour une modélisation économétrique explicative, il est souvent utile d'analyser les données sans formuler d'hypothèses à leur égard. Toute étude sophistiquée d'une base de données doit être précédée d'une étude exploratoire à l'aide de plusieurs outils, certes rudimentaires mais robustes. C'est la meilleure façon de se familiariser avec les données.

La première étape de toute investigation dans les données est l'examen des statistiques univariées des variables afin de détecter d'éventuelles anomalies dans leur distribution (valeurs manquantes, erronées ou atypiques). Dans un second temps, les statistiques bivariées permettent de repérer les incohérences entre les variables. Ces deux étapes permettent l'apurement et le redressement des données (imputation ou non des données manquantes, transformation logarithmique, discrétisation des variables continues, suppression de quelques observations le cas échéant).

2.3.1 Présentation des données

La base de donnée provient d'une grande mutuelle d'assurance française et contient des informations sur 50 000 observations pour des véhicules 4 roues de tourisme durant l'année 2004. Un enregistrement correspond à un couple (assuré, véhicule). Si un assuré a deux véhicules, il a deux enregistrements dans le fichier. En raison d'une charte de confidentialité, nous avons l'obligation de ne pas révéler le nom de la compagnie d'assurance.

Chaque assuré dispose dans le fichier de plusieurs informations qui concernent ses propres caractéristiques, les caractéristiques de son véhicule, du contrat d'assurance qu'il a choisi et des sinistres qu'il a déclarés. Ainsi, les variables se répartissent en quatre catégories :

- les caractéristiques propres au conducteur assuré,
- les caractéristiques du véhicule assuré,
- les contrats d'assurance souscrits
- et les variables décrivant la sinistralité déclarée, en termes de fréquences et de coûts, durant l'année 2004.

Le fichier comprend des variables catégorielles comme le sexe du conducteur, le type du conducteur, la formule d'assurance, la catégorie socioprofessionnelle et l'usage du véhicule. La base contient aussi des variables quantitatives : l'âge du conducteur, l'ancienneté du permis de conduire, l'ancienneté du véhicule, la puissance du véhicule, le coefficient de réduction majoration, le nombre de sinistre, les coûts des sinistres. (Les modalités des différentes variables qualitatives se trouvent dans l'annexe 1).

2.3.1.1 Les caractéristiques du conducteur

L'âge du conducteur : exprimé en années.

Le sexe du conducteur : variable binaire.

La catégorie socioprofessionnelle du conducteur : codée en 17 modalités (cf. annexe 1). A propos de la profession, il faut signaler d'une part que cette variable n'est pas toujours de bonne qualité dans un portefeuille puisque souvent l'assuré ne signale pas à son assureur un changement de profession. D'autre part, cette mutuelle d'assurance a la particularité d'avoir un nombre important d'assurés qui travaillent dans le secteur de l'éducation et en particulier des enseignants. Ces derniers représentent presque 60% du portefeuille. Nous codons donc cette variable en une variable indicatrice, prenant la modalité 1 pour « *enseignant* » et 0 pour les autres catégories socioprofessionnelles.

Le numéro de département : numéro de département du domicile de l'assuré. Elle est codée comme variable indicatrice prenant la modalité 1 pour « *la zone rurale* » et 0 pour « *la zone urbaine* ».

Le type du conducteur : variable binaire. Le sociétaire se déclare à la compagnie d'assurance, soit en tant que conducteur principal, soit en tant que conducteur secondaire. Quand une personne est amenée à conduire occasionnellement un véhicule assuré, elle se déclare en tant que conducteur secondaire. Souvent, les jeunes conducteurs, afin de réduire le coût de leur premier contrat d'assurance, se font déclarer en tant que conducteurs secondaires sur le véhicule de leurs parents. Ces derniers payent un supplément de cotisation pour que les jeunes puissent acquérir des antécédents d'assurance.

L'ancienneté du permis de conduire : exprimée en années. Les compagnies d'assurance en France distinguent, en fonction de cette caractéristique, les jeunes conducteurs des conducteurs expérimentés. Sont considérés comme jeunes conducteurs ou conducteurs novices, les conducteurs titulaires d'un permis de moins de trois ans. Les conducteurs expérimentés ont donc plus de trois ans d'expérience de conduite. Sont désignés aussi par conducteurs novices, les titulaires d'un permis de trois ans et plus, ne pouvant pas justifier d'une assurance effective au cours des trois dernières années précédant la souscription d'un nouveau contrat.

La variable Bonus - Malus : le système de bonus malus est réglementé et imposé à toutes les sociétés d'assurances par les pouvoirs publics. Il est exprimé par des coefficients de réduction ou de majoration (CRM), compris entre 0,50 et 3,50. Ce coefficient prend en considération l'expérience de conduite et le nombre d'accidents responsables enregistrés. Le conducteur qui n'enregistre pas d'accidents responsables bénéficie d'un bonus, et celui qui est responsable d'un accident est pénalisé d'un malus et voit sa prime augmenter l'année suivante. Précisément, un nouvel assuré commence par un CRM égal à 100%. S'il enregistre un accident responsable, son coefficient augmente de 25%. Par contre, quand aucun sinistre responsable n'est enregistré durant l'année, le CRM diminue de 5%. Ainsi, l'assuré atteint le coefficient minimal de 50% après 13 années de conduite sans aucun accident responsable. Le CRM représente donc un élément d'information important sur la sinistralité antérieure de l'assuré.

Une des conséquences de l'instauration du système de bonus malus est l'apparition du phénomène de *bonus hunger* ou *la soif du bonus*. Puisque le bonus malus augmente avec le nombre des sinistres responsables et indépendamment de leur montant, les assurés ont intérêt à dédommager eux même les petits sinistres dont ils sont responsables et ne déclarer que les sinistres à montants élevés, bien sûr supérieurs à la franchise fixée dans le contrat. Ils évitent ainsi les pénalités imposées par la compagnie d'assurance au niveau de leur coefficient de réduction majoration. Le nombre de sinistres enregistrés par l'assureur peut donc ne pas correspondre au nombre réel d'accidents causés par l'assuré durant l'année.

2.3.1.2 Les caractéristiques du véhicule

L'ancienneté du véhicule : exprimée en années.

La puissance réelle du véhicule : elle décrit la puissance du moteur en chevaux Din (Deutsch Industrie Normen). Cette mesure donne une vision plus réaliste de la puissance effective au niveau des roues (1 ch. Din = 0,735 Watt).

2.3.1.3 Les caractéristiques des contrats d'assurance

Cette compagnie d'assurance propose quatre formules de garantie pour l'assurance d'un véhicule quatre roues de tourisme : un contrat Responsabilité Civil et 3 formules Dommage au Véhicule (tous risques) avec différents niveaux de franchises.

La formule Responsabilité Civile (RC) : c'est la garantie minimum d'un contrat d'assurance automobile. Elle est obligatoire et couvre le conducteur responsable d'un accident. Dans cette formule, sont garanties les dommages corporels, l'assistance touristique, domestique et en cas de panne, les dommages matériels (suite à des catastrophes naturelles, à des évènements climatiques, à des attentats), la protection juridique en cas d'évènements garanties.

La formule Dommage au Véhicule (Formule 1) : elle comprend la garantie RC + la protection du véhicule, quelle qu'en soit la cause, avec un niveau élevé de franchise.

La formule Dommage au Véhicule (Formule 2) : elle inclut les garanties des formules précédentes + des garanties supplémentaires, comme par exemple, la proposition d'un véhicule de remplacement en cas de réparation suite à un vol ou un accident, le remboursement à valeur d'achat jusqu'à 12 mois, etc. La franchise fixée dans ce type de formule est moyenne.

La formule Dommage au Véhicule (Formule 3) : elle inclut les garanties des formules précédentes, avec d'autres options avec un plus haut niveau de prise en charge et d'assistance (par exemple assistance panne 0 km, remboursement à valeur d'achat jusqu'à 24 mois), avec une faible franchise.

Tout conducteur doit au minimum être assuré pour les dommages corporels et matériels causés aux tiers par le véhicule assuré (art. 211-1 et 211-5 du Code des assurances). Avec un contrat d'assurance contenant une franchise (*Formule 1*, *Formule 2*, *Formule 3*), aucune indemnité n'est versée lorsque la franchise est plus élevée que la perte. Dans le cas contraire, l'indemnité est égale à la perte dont on déduit la franchise.

2.3.1.4 Les caractéristiques des sinistres

Le nombre des sinistres : c'est le nombre d'accidents déclarés par l'assuré à la compagnie d'assurance pour l'année 2004. Comme nous l'avons souligné précédemment dans le paragraphe décrivant le coefficient du bonus malus, il ne s'agit pas du nombre réel des sinistres causés par l'assuré au cours de l'année. L'assuré peut décider de ne pas déclarer tous les sinistres à la compagnie d'assurance. Sa décision est influencée par la nature du contrat d'assurance.

Le montant des dépenses cumulées : lorsque ce montant figure pour une police, il représente les dépenses cumulées depuis l'enregistrement d'un ou plusieurs sinistres. Il est réinitialisé lors de chaque remise en cours.

Le montant des dépenses évaluées : ce montant correspond au montant évalué à priori par l'assureur des dégâts engendrés par les sinistres. Généralement, pour les dommages corporels, l'assureur ne peut définir le montant total de ses dépenses qu'après plusieurs mois, voir plusieurs années, d'où cette nécessité de faire une prévision de ses dépenses. Le coût des sinistres corporels comprend plusieurs composantes : indemnités des personnes physiques, soins, tierce personne, préjudices personnels et économiques. En, générale, les sinistres corporels représentent environ 3% du nombre total de sinistres indemnisés, mais plus du quart du coût total des sinistres.

2.3.1.5 Apurement de la base donnée

L'apurement du fichier est réalisé en corrigeant, d'une part, les erreurs de saisie (en se basant, bien évidemment, sur les conseils donnés par la compagnie d'assurance), et d'autre part, en écartant les incohérences entre les variables, qui sont difficiles à modifier. Tout au long de la thèse, les études statistiques et économétriques sont, en conséquence, basées sur un échantillon composé de 41 375 observations.

2.3.2 Les statistiques descriptives

Dans ce qui suit, nous exposons les principales statistiques descriptives. L'analyse exploratoire des données est détaillée dans l'annexe 2. L'échantillon fait apparaître environ autant d'hommes que de femmes assurés. Les conducteurs secondaires représentent 36% de la population, contre 64% pour les sociétaires déclarés comme conducteurs principaux. Presque la moitié des sociétaires réside en zone urbaine. La mutuelle d'assurance qui nous a fourni les données est initialement réservée aux enseignants. Ce n'est qu'à la fin des années 1990 qu'elle a été ouverte aux autres catégories socioprofessionnelles. Ceci explique le pourcentage élevé des enseignants représentant environ 60% des sociétaires. Vu l'importance de cette catégorie dans l'échantillon, nous transformons la variable ayant à l'origine 17 modalités en une variable indicatrice : enseignant et non enseignant.

TAB. 2.1: **La répartition des sociétaires selon le sexe, la profession et le type du conducteur**

Variable	**sexe**		**Type du conducteur**		**Profession**		**Zone de résidence**	
	Homme	Femme	Principal	Secondaire	Enseignant	Autre	Zone urbaine	Zone rurale
%	49.22	50.78	64.32	35.68	59.45	40.55	45.62	54.38

La moyenne d'âge des sociétaires est de 47 ans et l'ancienneté du permis de conduire est d'en moyenne 25 ans. 50% des véhicules des sociétaires ont, au maximum, 6 ans d'ancienneté, et des puissances assez faibles en chevaux DIN. Nous remarquons que l'age maximum des voitures assurés est de 83 ans. Nous pouvons penser qu'il y a un certain pourcentage de voiture de collection dans notre échantillon. Nous avons, d'ailleurs, 46 véhicules assurés ayant plus de 30 ans d'ancienneté. Cependant, nous n'avons aucune information sur le fait qu'ils soient des voitures de collection ou non. La distribution du Bonus Malus montre que plus de la moitié des assurés ont le bonus maximal 50 et seuls 25% de l'échantillon a un coefficient supérieur à 68. Cet échantillon présente donc une majorité d'assurés à bas risque. Ceci est représentatif de la population française des assurés automobiles. En effet, selon les chiffres publiés en l'année 2008 par la Fédération Française des Sociétés d'Assurance (FFSA), 64% des assurés automobiles français ont atteint le bonus maximal, soit un coefficient de réduction - majoration égal à 50.

TAB. 2.2: **Les statistiques descriptives des caractéristiques du sociétaire et de son véhicule**

Variable	Moyenne	Ecart type	Min	Max	Médiane	Q1	Q3
Ancienneté du permis	25.2699	12.990	0	78	26	15	35
Age du conducteur	46.6332	14.337	18	98	47	35	57
Ancienneté du véhicule	7.28817	5.055	0	83	6	3	10
Puissance du véhicule	83.9253	27.163	40	485	78	60	100
Bonus Malus	61.2703	17.582	50	232	50	50	68

La sinistralité est décrite dans le fichier par le nombre des sinistres déclarés par l'assuré, les coûts individuels cumulés et les coûts estimés des sinistres. Environ 11,8% de la population ont déclaré au moins un accident responsable pendant l'année, et 4% ont enregistré au moins 3 accidents dans l'année engendrant des dépenses à la compagnie d'assurance. Le nombre maximum d'accidents responsables déclarés est égal à 10, pour seulement deux sociétaires.

Remarque 1 *Evaluer le risque de l'assuré fait partie des objectifs de l'étude . Il est alors évident de s'intéresser aux accidents responsables déclarés à l'assureur et non pas aux accidents non responsables qui n'engendre aucun coût à la compagnie d'assurance. Alors tout au long de l'étude, la sinistralité se réfère aux accidents responsables.*

Dans l'échantillon, 10.6% des sociétaires ont déclaré des sinistres responsables dont les coûts sont fixés. Cependant, il existe dans le fichier 1.23% des sinistres responsables dont la détermination des coûts n'est pas encore cloturée. Ces accidents présentent donc des coûts évalués, pour lesquels la compagnie ne peut pas définir le montant exact au cours de l'année.

TAB. 2.3: **L'analyse descriptive des coûts cumulés et des coûts estimés des sinistres**

Variable	**Moyenne**	**Ecart type**	**Min**	**Max**	**Médiane**	**Skewness**	**Kurtosis**
Coût>0	1822.5	2012.8	500	45971.98	1221.6	6.74	92.6
Coût estimé>0	559	964.2	100.5	9298.7	225.1	7.2	60.8

La distribution des coûts des sinistres est fortement asymétrique à droite, présentant quelques valeurs extrêmes. En effet, un faible pourcentage de sociétaires cause des accidents responsables générant des coûts très élevés que la compagnie d'assurance doit supporter et auxquels elle applique un traitement particulier. Il est donc indispensable de se rappeler de ces *outliers* dans les modélisations ultérieures.

La répartition des conducteurs selon le contrat d'assurance choisi est illustrée dans le tableau 2.4.

TAB. 2.4: **La répartition des sociétaires selon le type du contrat d'assurance**

Contrat d'assurance	**RC**	**Formule 1**	**Formule 2**	**Formule 3**
%	42.5	17.02	8.99	31.74

La *Formule 2*, intermédiaire entre une assurance *tous risques* avec une franchise faible et celle avec la franchise la plus élevée, n'est souscrite que par 9% des assurés. Environ 42% des sociétaires ont choisi la couverture minimum obligatoire. D'après la Fédération Française des Sociétés d'Assurance (FFSA), la couverture obligatoire de la responsabilité civile représente 38% des cotisations sur le marché de l'assurance automobile en France. Notre échantillon est ainsi représentatif des assurés automobiles français du point de vue des souscriptions des contrats d'assurance.

Les sinistres responsables et les contrats d'assurance Au niveau de l'influence de l'étendue des couvertures d'assurance sur la fréquence des sinistres, nous pouvons nous attendre à deux types de résultats :

1) Soit nous observons, d'une part, une faible fréquence de sinistralité de la part des assurés ayant souscrit uniquement la garantie de la responsabilité civile et se sachant les moins risqués. Et d'autre part, plus d'accidents de la part des plus couverts. Cet éventuel résultat justifie la présence d'un avantage informationnel de la part des assurés, dans le sens où, en présence d'asymétrie d'information, les prédictions théoriques stipulent que les agents les plus risqués sont plus couverts que les moins risqués (Rothschild et Stiglitz, 1976).

2) Soit nous notons une faible sinistralité de la part des plus couverts. Ce qui signifie qu'une forte aversion au risque est interprétée par la souscription des garanties tous risques à très faible franchise. Et dans ce cas, la sinistralité est corrélée négativement avec les garanties à franchises élevées.

En croisant les variables décrivant l'occurrence des sinistres et les contrats d'assurance, nous remarquons selon les résultats illustrés dans le tableau 2.5 que la sinistralité est plus importante chez les sociétaires ayant souscrit le contrat obligatoire par rapport au reste de l'échantillon (15.5%).

Tab. 2.5: **La répartition du portefeuille selon l'occurrence des sinistres responsables et les contrats d'assurance**

%	**Sinistres = 0**	**Sinistres > 0**
RC	84.54	15.46
Formule 1	96.02	3.98
Formule 2	92.18	7.82
Formule 3	87.65	12.35
test de Kruskal-Wallis : $\chi^2 = 697.12$, $Pr < .0001$		

Cependant, 12% des assurés qui ont déclaré au moins un sinistre responsable pendant l'année ont choisi la plus forte couverture (contre 4% et 8% des sociétaires pour les deux autres garanties tous risques). Et ce sont ces sociétaires qui ont engendré, en moyenne, la charge la plus importante que supporte la compagnie d'assurance (voir tableau 2.6).

Tab. 2.6: **La distribution des coûts des sinistres par contrat d'assurance**

	coûts > 0		
	Moyenne	**Médiane**	**Ecart type**
RC	1706.5	1167	1728
Formule 1	1898	1319	2505
Formule 2	1913.9	1278	2038.7
Formule 3	1993	1310	2341
test de Kruskal-Wallis : $\chi^2 = 24.62$, $Pr < .0001$			

Tous ces résultats descriptifs doivent nécessairement être complétés par une analyse predictive prenant en considération toutes les informations dont nous disposons. Ainsi, nous devons isoler les effets propres de toutes les variables, nous permettant d'avoir une idée plus approfondie sur l'intensité de la liaison entre le choix de la garantie d'assurance et le risque de l'assuré mesuré en termes de fréquences et de coûts de sinistres, conditionnellement aux informations observables par l'assureur.

L'analyse que nous appliquons par la suite est réalisée sur le sous-échantillon des jeunes conducteurs d'une part et celui des conducteurs expérimentés d'autre part. Les jeunes conducteurs (resp. conducteurs expérimentés) sont de l'ordre de 1 149 assurés (resp. 40 226 assurés), soit 3% (resp. 97%) de tout l'échantillon. Le choix de former deux sous-groupes d'assurés vient du fait que l'échantillon contient des informations concernant des catégories hétérogènes de conducteurs ayant des historiques de conduite différents. Homogénéiser les caractéristiques des assurés permet d'isoler les effets des observables propres aux jeunes conducteurs sur leur sinistralité déclarée à la compagnie d'assurance, et d'en faire la comparaison par rapport au reste de la population qui compose les conducteurs expérimentés.

Nous résumons dans le tableau 2.7 suivant les statistiques descriptives des jeunes conducteurs et des conducteurs expérimentés.

Tab. 2.7: **Résumé des statistiques descriptives des jeunes conducteurs et des conducteurs expérimentés**

Variables		**Jeunes conducteurs**	**Conducteurs expérimentés**
Sexe	**Homme**	54.57 %	49.07 %
	Femme	45.43 %	50.93 %
Age du conducteur (années)	**Moyenne**	21	47
	Ecart-type	4	14
	Min	18	21
	Max	67	98
Profession	**Enseignant**	54.22 %	59.60 %
	Autres	45.78 %	40.40 %
Lieu de résidence	**Urbain**	45.34 %	45.62 %
	Rural	45.62 %	54.38 %
Ancienneté du permis (années)	**Moyenne**	1	26
	Ecart-type	1	13
	Min	0	78
	Max	2	3
le type du conducteur	**Principal**	13.40 %	65.77 %
	Secondaire	86.60 %	34.23 %
Ancienneté du véhicule (années)	**Moyenne**	9	7
	Ecart-type	5	5
	Min	0	0
	Max	39	83
Puissance du véhicule (Din)	**Moyenne**	66.5	84.4
	Ecart-type	15.5	27.25
	Min	40	40
	Max	140	485
Bonus Malus (coefficient)	**Moyenne**	96	60
	Ecart-type	11	17
	Min	50	50
	Max	156	232
Contract	**RC**	11.40 %	43.14 %
	Formule 1	37.16 %	16.44 %
	Formule 2	18.89 %	8.71 %
	Formule 3	32.55 %	31.72 %
Accidents responsables déclarés	**Non accident**	89.56 %	88.13 %
	1 et plus	10.44 %	11.87 %
Coûts cumulé des sinistres (euros)	**Moyenne**	201	217.95
	Ecart-type	872	860.66
	Min	0	0
	Max	17753.75	45971.97
Coûts évalués des sinistres (euros)	**Moyenne**	103.4	103.37
	Ecart-type	365.35	400.62
	Min	0	0
	Max	8913.57	9298.71

2.4 Etude de la corrélation entre la couverture d'assurance et le risque

Nous rappelons qu'en utilisant des données en coupe instantanée, l'approche de la corrélation positive ne permet pas de distinguer le sens de causalité entre le risque et la couverture et de différencier ainsi la sélection adverse de l'aléa moral. Alors que des données multidimensionnelles permettent de distinguer les deux problèmes d'asymétrie d'information (Dionne et *al.* (2004), Abbring et *al.* (2003)).

Une autre limite dans l'analyse des données décrivant une seule période de contrat d'assurance concerne l'information relative à l'historique des assurés en matières de sinistralité. Ne pas contrôler l'hétérogénéité des individus peut causer des problèmes d'estimation dans l'étude de l'asymétrie d'information. Plusieurs travaux utilisent le coefficient du bonus malus comme variable décrivant l'historique des agents. Nous étudions ce point avec plus de détail dans la section 2.5.

Dans ce qui suit, nous effectuons des estimations séparées du choix du contrat d'assurance d'une part et de la sinistralité d'autre part, en tenant compte des différentes caractéristiques de l'assuré et de son véhicule. Ensuite nous reprenons l'idée de Chiappori et Salanié (2000) afin de tester la présence de l'asymétrie d'information.

2.4.1 Le choix du contrat d'assurance et les caractéristiques observables

Nous estimons dans cette partie le choix des contrats d'assurance en fonction des différentes caractéristiques du conducteur et de son véhicule. Nous choisissons d'appliquer le modèle logit univarié ordonnée vu le caractère cumulatif des garanties associées aux différents contrats d'assurance.

Notons C_i la variable ordonnée décrivant le choix du contrat d'assurance par un assuré i.

$$C_i = \begin{cases} 0 & \text{si l'assuré a souscrit la RC} \\ 1 & \text{si l'assuré a souscrit la Formule 1,} \\ 2 & \text{si l'assuré a souscrit la Formule 2,} \\ 3 & \text{si l'assuré a souscrit la Formule 3,} \end{cases}$$

La variable latente continue $C_i^\star$ associée à la variable observée C_i est définie comme suit :

$$C_i^\star = \alpha + \beta X_i + \epsilon_i$$

avec

$$\begin{cases} C_i = 0 & \text{si} \quad C_i^\star < l_1 \\ C_i = 1 & \text{si} \quad l_1 < C_i^\star < l_2 \\ C_i = 2 & \text{si} \quad l_2 < C_i^\star < l_3 \\ C_i = 3 & \text{si} \quad l_3 < C_i^\star \end{cases}$$

Le vecteur X_i inclut toutes les variables observables décrivant les caractéristiques de l'assuré et de son véhicule. Les limites et les paramètres sont inconnus et doivent être estimés. Pour assurer l'identification des paramètres, il est nécessaire d'imposer la nullité de la première limite l_1, et dans ce cas, nous conservons la constante dans le modèle.

Par exemple, la probabilité que $C_i = 0$ est

$$\begin{aligned} Pr(C_i = 0) &= Pr(C_i^\star < l_1) = Pr(\beta X_i + \epsilon_i < l_1) \\ &= Pr(\epsilon_i < l_1 - \beta X_i) = F(l_1 - \beta X_i). \end{aligned}$$

Et la probabilité que $C_i = 1$ est

$$\begin{aligned} Pr(C_i = 1) &= Pr(l_1 < C_i^\star < l_2) = Pr(l_1 \leq \beta X_i + \epsilon_i < l_1) \\ &= Pr(\epsilon_i < l_2 - \beta X_i) - Pr(\epsilon_i < l_1 - \beta X_i) \\ &= F(l_2 - \beta X_i) - F(l_1 - \beta X_i), \end{aligned}$$

avec $F(.)$ la fonction de répartition de la loi logit. Les deux dernières probabilités sont déterminées de façon similaire.

La valeur des coefficients obtenus dans l'équation importe peu. C'est seulement leur signe et leur significativité qui nous intéressent le plus. Nous pouvons, toutefois, étudier la contribution de chaque variable explicative sur la probabilité de choisir un contrat avec franchise par rapport au contrat RC en ayant recours au calcul des effets marginaux $\partial C/\partial X_k$ des k variables explicatives incluses dans le modèle. Pour les variables muettes, $\partial C/\partial X_k$ représente une variation discrète de la variable de 0 à 1.

Les résultats de l'estimation

Les résultats des estimations du modèle logit ordonné sur le sous-échantillon des jeunes conducteurs et celui des conducteurs expérimentés sont présentés dans le tableau 2.8. Les effets marginaux sont décrits dans les tableaux 2.9 et 2.10

Tab. 2.8: **Les contrats d'assurance pour les jeunes conducteurs. Résultats de l'estimation du modèle logit ordonné**

	Les jeunes conducteurs			**Les conducteurs expérimentés**		
Paramètres	**Estimations**	**Erreur type**		**Estimations**	**Erreur type**	
Constante	2.967243	0.650069	***	0.451569	0.081813	***
Homme	-0.018775	0.111660		0.139143	0.019601	***
Age du conducteur	-0.017055	0.015749		-0.022285	0.002000	***
Enseignant	0.229758	0.110647	**	-0.153840	0.019185	***
Zone urbaine	0.031959	0.109337		-0.040418	0.018873	**
Ancienneté du permis	0.097565	0.074328		-0.009461	0.002283	***
Conducteur principal	0.112594	0.184892		0.005327	0.020064	
Bonus malus	0.003323	0.005297		0.006858	0.000647	***
Ancienneté du véhicule	-0.118007	0.012106	***	0.032525	0.001881	***
Puissance du véhicule	0.003557	0.003644		0.005731	0.000380	***
Limit2	2.106607	0.094748	***	0.717500	0.008238	***
Limit3	3.024283	0.107639	***	1.105187	0.009963	***

La significativité au seuil de 10%, 5% et 1% est noté respectivement par $*$, $**$ et $***$

Tab. 2.9: **Les effets marginaux, jeunes conducteurs**

Variables	**RC**	**Formule 1**	**Formule 2**	**Formule 3**
Enseignant	-0.0205187	-0.0316039	0.0043270	0.0477956
Ancienneté du véhicule	0.0105427	0.0162384	-0.0022233	-0.0245579

Tab. 2.10: **Les effets marginaux, conducteurs expérimentés**

Variables	**RC**	**Formule 1**	**Formule 2**	**Formule 3**
Homme	-0.0324053	0.0011272	0.0031424	0.0281358
Age du conducteur	0.0051475	-0.000179045	-0.000499162	-0.0044693
Enseignant	0.0357725	-0.0012443	-0.0034689	-0.0310593
Zone urbaine	0.0093790	-0.00032647	-0.00090942	-0.0081431
Ancienneté du permis	0.0022168	-0.000077107	-0.000214968	-0.0019247
Bonus malus	-0.0015995	0.000055637	0.000155111	0.0013888
Ancienneté du véhicule	-0.0075434	0.000262381	0.000731498	0.0065495
Puissance du véhicule	-0.0013302	0.000046267	0.000128989	0.0011549

Nous constatons à partir des résultats de la régression du logit ordonné que les caractéristiques de l'assuré et de son véhicule ont des effets différents sur le choix de la garantie selon l'expérience de conduite (le fait d'être jeune conducteur ou conducteur expérimenté). Pour les jeunes conducteurs, seules la catégorie socio-professionnelle et l'ancienneté du véhicule sont significatives. Le coefficient estimé de l'ancienneté du véhicule est négatif. L'assuré se voit généralement proposer des couvertures dégressives[9] avec l'ancienneté de la voiture. Ceci est confirmé par le résultat des effets marginaux : la probabilité de souscrire des contrats d'assurance à très faible ou moyenne franchise diminue avec l'ancienneté du véhicule, toutes choses égales par ailleurs. Par contre plus le véhicule est ancien plus la probabilité de signer un contrat RC par rapport aux contrats *tous risques* augmente. Le coefficient du bonus malus ne présente aucun effet significatif. Ceci s'explique par le fait que le jeune conducteur n'a pas encore acquis des antécédents d'assurance. Ainsi, les informations que l'assureur détient sur le jeune conducteur n'interviennent pas toutes dans la probabilité de souscrire un contrat "tous risques".

[9] Comme par exemple les contrats proposés par la Crédit Mutuelle

Par contre, les variables observées par l'assureur décrivant les caractéristiques du conducteur expérimenté et de son véhicule sont toutes significatives dans l'explication du choix de la garantie, à part le type du conducteur qui ne joue aucun effet sur le choix du contrat d'assurance.

2.4.2 L'occurrence des sinistres

Notons Y_i une variable binaire décrivant l'occurrence des sinistre pour un assuré i.

$$Y_i = \begin{cases} 0 & \text{si l'assuré n'a déclaré aucun sinistre responsable,} \\ 1 & \text{si l'assuré a déclaré au moins un sinistre responsable au cours de l'année} \end{cases}$$

La variable latente $Y_i^\star$ associée à la variable observée Y_i est expliquée en fonction du choix du contrat d'assurance et des caractéristiques observables de l'assuré i et de sa voiture. Dans un premier temps, nous appliquons un modèle logit binomial où nous introduisons le choix du contrat d'assurance comme une variable binaire Z qui prend la valeur 0 si le contrat est de type responsabilité civil et 1 si le contrat est un contrat *tous risque*, (i.e. toutes les formules confondues 1, 2 et 3 ont dans ce cas la même modalité). Plus formellement :

$$Y_i^\star = \alpha + \beta X_i + \gamma Z_i + \epsilon_i$$

En second lieu, la variable C_i décrivant le choix de contrat à 4 modalités remplace la variable Z_i, alors :

$$Y_i^\star = \alpha + \beta X_i + \gamma C_i + \epsilon_i$$

Dans les deux cas,

$$\begin{cases} Y_i = 0 \text{ si } Y_i^\star \leq 0 \\ Y_i = 1 \text{ si } Y_i^\star > 1 \end{cases}$$

Appliquer ces deux équations nous permet de comparer l'effet de la variable du contrat d'assurance sur la sinistralité dans le cas où elle prend seulement deux niveaux de couverture comme dans les travaux de Chiappori et Salanié (2000), Cohen(2005), etc, et dans le cas où les couvertures avec différentes franchises sont prises en considération. Ceci permet d'avoir une aide à l'interprétation des résultats des modèles bivariés appliqués dans la section suivante.

Les coefficients des deux logits binomiaux sont reportés aux tableaux 2.11 et 2.12 pour lesquels correspondent respectivement les tableaux des effets marginaux 2.16 et 3.29. Dans le premier tableau 2.11, apparaît la variable binomiale Z du contrat d'assurance. Alors que la variable C multinomiale ressort dans le tableau 2.12. Ces modélisations sont appliquées pour les jeunes conducteurs et les conducteurs expérimentés séparément.

TAB. 2.11: **L'occurrence des sinistres. Le modèle logit binomial**

$$Y_i^\star = \alpha + \beta X_i + \gamma Z_i + \epsilon_i$$

Paramètres	**Jeunes conducteurs**		**Conducteurs expérimentés**	
Constante	-10.319553	***	-3.741526	***
	(1.175321)		(0.130707)	
Homme	0.314163		-0.129537	***
	(0.227280)		(0.033387)	
Age du conducteur	-0.000791		0.000563	
	(0.031240)		(0.003304)	
Enseignant	0.063823		-0.013268	
	(0.220810)		(0.032597)	
Zone urbaine	-0.220388		-0.102047	***
	(0.219115)		(0.031876)	
Ancienneté du permis	0.789228	***	0.015156	***
	(0.175691)		(0.003765)	
Conducteur principal	-0.412430		0.068296	**
	(0.354891)		(0.034301)	
Bonus malus	0.075250	***	0.027235	***
	(0.008312)		(0.000939)	
Ancienneté du véhicule	-0.010999		-0.057008	***
	(0.023535)		(0.004007)	
Puissance du véhicule	-0.000301		0.003988	***
	(0.007301)		(0.000617)	
Formule tous risques	-0.378657		-0.547268	***
	(0.327153)		(0.035461)	

La significativité au seuil de 10%, 5% et 1% est noté respectivement par $*$, $**$ and $***$

Les valeurs entre parenthèses représentent les écarts-types des paramètres estimés

Tab. 2.12: **L'occurrence des sinistres. Le modèle logit binomial**

$$Y_i^\star = \alpha + \beta X_i + \gamma C_i + \epsilon_i$$

Paramètres	Jeunes conducteurs		Conducteurs expérimentés	
Constante	-10.841549	***	-3.929808	***
	(1.198334)		(0.132005)	
Homme	0.300660		-0.109413	***
	(0.230204)		(0.033515)	
Age du conducteur	0.008686		0.000284	
	(0.030834)		(0.003307)	
Enseignant	-0.001646		-0.024875	
	(0.223922)		(0.032659)	
Zone urbaine	-0.240695		-0.115846	***
	(0.221382)		(0.031949)	
Ancienneté du permis	0.755312	***	0.014656	***
	(0.177527)		(0.003774)	
Conducteur principal	-0.452277		0.072716	**
	(0.355778)		(0.034391)	
Bonus malus	0.075633	***	0.028228	***
	(0.008311)		(0.000951)	
Ancienneté du véhicule	0.052751	**	-0.025828	***
	(0.028203)		(0.004429)	
Puissance du véhicule	0.000172		0.004095	***
	(0.007529)		(0.000621)	
Formule 1	-1.330205	***	-1.467841	***
	(0.427666)		(0.077520)	
Formule 2	-0.789098	*	-0.835355	***
	(0.406015)		(0.072232)	
Formule 3	-0.037929		-0.375423	***
	(0.335574)		(0.036580)	

La significativité au seuil de 10%, 5% et 1% est noté respectivement par $*$, $**$ and $***$

Les valeurs entre parenthèses représentent les écarts-types des paramètres estimés

TAB. 2.13: **Les effets marginaux**

$$Y_i^\star = \alpha + \beta X_i + \gamma Z_i + \epsilon_i$$

Variables	**Jeunes conducteurs**	**Conducteurs expérimentés**
Homme	n.s.	-0.0129493
Age du conducteur	n.s.	n.s.
Enseignant	n.s.	n.s.
Zone urbaine	n.s.	-0.0102012
Ancienneté du permis	0.0597361	0.0015151
Conducteur principal	n.s.	0.0068272
Bonus malus	0.0056956	0.0027225 ***
Ancienneté du véhicule	n.s.	-0.0056989
Puissance du véhicule	n.s.	0.000398616
Formule tous risque	n.s.	-0.0547078

TAB. 2.14: **Les effets marginaux**

$$Y_i^\star = \alpha + \beta X_i + \gamma C_i + \epsilon_i$$

Variables	**Jeunes conducteurs**	**Conducteurs expérimentés**
Homme	n.s.	-0.0108873
Age du conducteur	n.s.	n.s.
Enseignant	n.s.	n.s.
Zone urbaine	n.s.	-0.0115275
Ancienneté du permis	0.0558448	0.0014584
Conducteur principal	n.s.	0.0072357
Bonus malus	0.0055920	0.0028088
Ancienneté du véhicule	0.0039002	-0.0025701
Puissance du véhicule	n.s.	0.000407478
Formule 1	-0.0983501	-0.1460596
Formule 2	-0.0583428	-0.0831232
Formule 3	n.s.	-0.0373571

Chaque valeur représente l'effet marginal de la variable X_k sur la probabilité que $Y = 1$

ns signifie que la variable est non significative

L'hypothèse de nullité globale des coefficients des deux estimations, d'après les tests de vraisemblance, de Score et de Wald, est rejetée. Les critères d'Akaïké, de Schwarz et de la log-vraisemblance montrent que le modèle ajustent correctement les données. Ceci est d'autant plus justifié par le fait que les paires concordantes entre les probabilités prédites et les valeurs observées sont respectivement de 76% et 66% pour les jeunes conducteurs et les conducteurs expérimentés dans la première estimation, et de 80.5% et 66% pour la deuxième modélisation.

En ce qui concerne les jeunes conducteurs, les estimations des deux modèles logits affichent les résultats suivants.

Les effets des variables explicatives différent du premier modèle où nous avons introduit la variable binaire du choix de contrat, au deuxième modèle où les différents contrats proposés sont pris en considérations. En effet, parmi les caractéristiques du conducteur et de son véhicule, l'ancienneté du véhicule n'a aucun effet statistiquement significatif. Alors qu'il s'avère dans le deuxième modèle que l'ancienneté du véhicule contribue significativement, au seuil de 5%, à la hausse de la probabilité de déclarer au moins un accident responsable (même si cette contribution est de l'ordre de 0.04%, ce qui est très faible). Les deux autres variables qui sont l'ancienneté du permis de conduire et le bonus malus contribuent presque de la même manière dans la probabilité de déclarer au moins un sinistre responsable.

Dans le premier modèle, l'ancienneté du permis contribue à augmenter d'environ 6% la probabilité de déclarer au moins un accident responsable à l'assureur au cours d'une année. Un jeune conducteur, comme il a été supposé, a un maximum de 2 ans d'expérience de conduite. Le conducteur au bout de 2 ans d'expérience commence à avoir une certaine aisance au volant qui le laisse surestimer son mode de conduite. Il en résulte une mauvaise évaluation de son propre risque et tend à avoir un effet significatif sur la sinistralité.

Le bonus malus apparaît dans les résultats de l'estimation, toutes choses égales par ailleurs, comme variable affectant la probabilité de la sinistralité au cours d'une année, même si l'effet n'apparaît pas comme relativement élevé. En effet, un accroissement du coefficient de réduction majoration du jeune conducteur assuré augmente de 0.6% la probabilité de déclarer au moins un accident responsable. Le système bonus malus pénalise les assurés responsables dans un ou plusieurs accidents en augmentant leur coefficient de réduction majoration, et ainsi leur prime d'assurance, l'année suivant la déclaration des sinistres. Le coefficient estimé positif associé à la variable de bonus malus peut être alors interprété comme suit : un jeune assuré qui a déclaré au moins un sinistre responsable dans le passé aura une forte probabilité d'en déclarer d'autres. Cependant, il est à noter que le passé d'un jeune conducteur ne constitue que deux années de conduite au maximum (se référer à la définition du jeune conducteur).

Le point principal qui ressort de ces deux estimations est que l'effet de la variable du choix du contrat n'est pas le même d'une équation à l'autre. Aucune contribution significative de la couverture d'assurance n'est détectée dans le cas où on estime l'effet d'un contrat *tous risques* par rapport à un contrat de responsabilité civile, indépendamment des niveaux de franchises proposés. Par contre, lorsque les différentes formules de garanties sont prises en considération, l'effet de la couverture sur la probabilité des sinistres devient significatif pour les jeunes conducteurs. En effet, l'estimation fait ressortir des coefficients significatifs associés aux formules de dommage au véhicule. L'effet est essentiellement significatif pour les contrats à forte et moyenne franchise (*Formule 1* et *Formule 2*), ce qui n'est pas le cas pour le contrat avec des garanties plus étendues (*Formule 3*). Toutes choses égales par ailleurs, le fait que l'assuré ait souscrit la *Formule 1* (resp. la *Formule 2*) diminue de 10% (resp. 6%) sa probabilité d'accident, que s'il était couvert exclusivement par une garantie de responsabilité civile. Cependant, la souscription de contrat *Formule 3* présentant les garanties les plus étendues (et donc la franchise la plus faible) ne joue aucun rôle statistiquement significatif sur la sinistralité des jeunes conducteurs. Une remarque peut être faite à ce niveau : la contribution en valeur absolue diminue de la *Formule 1* à la *Formule 2* jusqu'à ce qu'elle s'annulle pour la *Formule 3*, ce qui peut expliquer d'une manière la non significativité de la variable binaire Z du choix du contrat dans la premier modèle. Il est donc plus pertinent de prendre en considération tous les niveaux de garanties pour les contrats de dommage au véhicule.

L'estimation des deux modèles pour la population des conducteurs expérimentés présente quasiment les mêmes résultats quant aux effets des variables explicatives sur l'occurrence des sinistres responsables, quelque soit la variable du choix de contrat d'assurance introduite (binaire Z ou multinomiale C).

L'âge et la catégorie socioprofessionnelle n'ont aucun effet significatif sur la probabilité de déclarer au moins un sinistre responsable. Toutes les autres variables décrivant les caractéristiques du conducteur expérimenté et de son véhicule ont des pouvoirs statistiquement significatifs sur la sinistralité, à savoir la zone géographique du domicile, le sexe du conducteur, l'ancienneté du permis de conduire, le type de l'assuré (qu'il soit conducteur principal ou non), le coefficient du bonus malus, l'ancienneté du véhicule et sa puissance. Le paramètre estimé associé à la variable de la puissance du véhicule présente le signe attendu. Plus le véhicule est puissant plus le risque afférent peut être non négligeable. Cependant, cette variable ne contribue qu'à un niveau assez faible (0.04%) dans l'augmentation de la probabilité de déclarer un sinistre responsable. L'effet de la variable sexe qui n'est pas significative pour les jeunes conducteurs fait apparaître une sursinistralité statistiquement significative pour les conductrices expérimentées femmes par rapport aux hommes (1.1%).

Tous les coefficients estimés associés aux différentes formules d'assurance *tous risques* apparaissent significativement différentes de zéro au seuil de 1%. L'impact des contrats d'assurance *tous risques* relativement au contrat RC sur la probabilité de déclarer des sinistres responsables est le même que pour les jeunes conducteurs, à savoir une relation négative entre les différentes couvertures d'assurance *tous risques* et la sinistralité.

Le terme de perturbation ϵ inclue une hétérogénéité inobservée qui correspond à des caractéristiques individuelles non observées (ou non observables) par l'assureur, comme par exemples les variables comportementales liées au mode de conduite de l'assuré (conduire en état d'ébriété, ne pas mettre la ceinture de sécurité...). Dans une spécification simple, tel que le modèle logit univarié, une hétérogénéité non observée n'introduit pas de biais dans l'estimation des paramètres, à condition que les caractéristiques individuelles inobservables ne soient pas corrélées avec les variables explicatives. Cependant, il existe plusieurs facteurs cachés, inobservables par l'assureur au moment du choix du niveau de la couverture d'assurance et qui peuvent ainsi influencer indirectement la probabilité de la déclaration des sinistres. Nous pouvons donc soupçonner une corrélation entre ces variables inobservables et le choix du contrat d'assurance, ce dernier étant choisi et souscrit en fonction des caractéristiques propres au conducteur et à son véhicule. Des biais de sélection peuvent ainsi surgir. La relation de causalité entre le risque et la couverture illustrée par ces modèles logits univariés ne répond donc pas à la question de l'asymétrie d'information. Par conséquent, il est indispensable de prendre en compte l'endogénéité de la variable du choix du contrat d'assurance en appliquant des modèles bivariés récursifs.

2.4.3 Le modèle probit bivarié récursif

La modélisation économétrique que nous appliquons dans cette partie est différente de celle de Chiappori et Salanié (2000). Ces auteurs ont utilisé un modèle probit bivarié où la sinistralité et la couverture d'assurance sont deux variables endogènes estimées d'une manière simultanée en fonction des caractéristiques observables de l'assuré et de son véhicule.

Nous choisissons d'utiliser un modèle probit bivarié récursif qui estime aussi de façon simultanée la couverture et la sinistralité conditionnellement à toutes les observables, mais qui retient comme élément prépondérant de l'équation de la sinistralité le choix de la couverture d'assurance.

Le modèle probit bivarié récursif se présente comme suit :

$$\begin{cases} Z^\star = \alpha X_1 + \varepsilon_1 \\ Y^\star = \beta X_2 + \gamma Z + \varepsilon_2 \end{cases}$$

où $Z^\star$ est une variable latente continue associée à la variable observée Z telle que :

$$Z = \begin{cases} 1 \quad si\ Z^\star > 0, \text{ l'assuré a choisi un contrat } \textit{tous risques}, \\ 0 \quad si\ Z^\star \leq\ 0, \text{ l'assuré a choisi le contrat obligatoire RC.} \end{cases}$$

et $Y^\star$ est la variable latente décrivant la déclaration d'au moins un sinistre responsable au cours de l'année, associée à la variable observée Y telle que :

$$Y = \begin{cases} 1 & si\ Y^\star > 0,\ \text{l'assuré a déclaré au moins un sinistre responsable,} \\ 0 & si\ Y^\star \leq\ 0,\ \text{l'assuré n'a déclaré aucun sinistre responsable.} \end{cases}$$

Nous supposons que les erreurs ε_1 et ε_2 sont normalement et conjointement distribuées telles que $E(\varepsilon_1) = E(\varepsilon_2) = 0$, $Var(\varepsilon_1) = Var(\varepsilon_2) = 1$ et :

$$\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix} \approx N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right]$$

où ρ est le coefficient de corrélation entre les termes résiduels, qui prend la valeur zéro si les deux équations sont indépendantes.

$Y^\star$ peut être écrite en fonction de $Z^\star$ comme suit :

$$Y^\star = \beta X_2 + \gamma Z + \rho \left(Z^\star - \alpha X_1\right) + u$$

avec u un terme d'erreur suivant une loi normale d'espérance nulle et de variance $(1 - \rho^2)$.

Ce modèle estime d'une manière simultanée deux équations, expliquant respectivement la couverture d'assurance et la sinistralité. L'équation de la couverture d'assurance fait intervenir comme variables explicatives les caractéristiques du conducteur et de son véhicule. Quant à l'équation de la sinistralité, la variable observée du choix de la couverture Z y est introduite en tant que variable explicative en plus des caractéristiques du conducteur et de son véhicule.

Le choix de ce modèle récursif à la place d'un modèle à la Chiappori et Salanié (2000) est motivé par le fait qu'il permet d'analyser l'ineterdépendance entre la sinistralité et le choix de la garantie, conditionnellement à toutes les variables observables, en prenant en considération l'endogénéité de la couverture d'assurance dans l'équation de la sinistralité. En d'autres termes, le modèle corrige à la fois l'endogénéité potentielle de la variable de couverture d'assurance et l'hétérogénéité inobservée et/ou inobservable par la corrélation entre les termes d'erreurs des deux équations de la sinistralité et de la couverture.

Le modèle décrit, d'une part, l'effet indirect conditionnel entre la couverture et le risque à travers les caractéristiques individuelles inobservables (captées par le coefficient de corrélation ρ entre les termes d'erreurs), et d'autre part, la relation directe entre le risque et la couverture (si, dans l'équation de la sinistralité, le paramètre γ associé à la couverture d'assurance est différente de zéro).

D'une part, cette modélisation permet de déterminer ce qui de la couverture d'assurance et des caractéristiques de l'assuré et du véhicule est le plus déterminant dans l'équation de l'occurrence du sinistre. Le paramètre γ associé à la variable de la couverture d'assurance, s'il est différent de zéro, signifie qu'il existe un lien direct significatif entre le choix de la couverture et la probabilité de sinistre responsable. Certes, le signe de ce paramètre γ décrit une certaine relation de causalité entre la couverture et les sinistres. Et on pourrait dire que l'existence d'aléa moral est testée en examinant si le paramètre γ est significativement positif. Mais on sait qu'une telle procédure est particulièrement vulnérable à un biais lié aux phénomènes d'anti-sélection. En effet, les conducteurs qui décident de souscrire une assurance *tous risques* se caractérisent peut être par des comportements plus risqués que d'autres. Dans ce cas, loin de mesurer l'aléa moral, les estimations mettrait en évidence une causalité inverse à celle de l'aléa moral : on observerait seulement le fait que les plus exposés au risque d'accident s'assurent mieux. Ce biais d'anti-selection est lié à l'hétérogénéité non observée qui figure à la fois dans les perturbations ε_1 et ε_2 qui ne seront donc pas indépendantes.

Identification du modèle En raison de la récursivité du modèle, les paramètres du modèle ne peuvent pas être identifiés si les variables explicatives des deux équations sont identiques (Maddala, 1983, pp. 117-147). Afin d'assurer l'identification du modèle, nous écartons une variable explicative de l'équation de l'occurrence du sinistre. Nous l'introduisons, ensuite, dans la première équation du choix du contrat d'assurance en tant que variable instrumentale. Il n'est pas aisé de choisir la bonne variable instrumentale. Cette dernière doit être, bien évidemment, exogène et corrélée avec la variable décrivant le type du contrat. Pour cela, nous testons deux premiers candidats : la profession et le type du conducteur.

La raison pour laquelle nous pensons en premier lieu à ces deux variables pour en choisir une variable instrumentale est que, à notre avis, la profession ou le type du conducteur n'affectent que d'une façon indirecte la sinistralité à travers la variable explicative endogène du choix du contrat d'assurance : le fait d'appartenir à une catégorie socio-professionnelle peut ne pas jouer directement sur le fait de déclarer un accident responsable, par contre la souscription d'un contrat d'assurance en dépend directement. De la même manière pour le type du conducteur, si l'assuré est le conducteur principal, le contrat d'assurance est attribué en fonction des caractéristiques liées au conducteur. Cependant, si l'assuré est le conducteur secondaire, le contrat d'assurance dépend de ses caractéristiques ainsi que des caractéristiques propres au conducteur principal[10].

Il en ressort que le choix d'une variable instrumentale parmi les deux candidats

[10]La variable qui décrit si l'assuré est le conducteur principal ou secondaire a été choisie par Grun-Rehomme et Benlagha (2007) comme variable instrumentale dans leur étude sur l'asymétrie d'information chez les jeunes conducteurs.

n'est valide que pour les conducteurs expérimentés : Seule la profession du conducteur a un impact significativement positif sur le choix du contrat d'assurance. Pour les jeunes conducteurs, un problème du choix de la variable instrumentale subsiste : ni la profession, ni le type du conducteur ne sont corrélés significativement avec la couverture d'assurance. Il est donc nécessaire d'opter pour un autre candidat.

Nous choisissons de tester l'ancienneté du véhicule comme variable instrumentale. L'ancienneté du véhicule a un effet direct sur la souscription du contrat d'assurance : plus la voiture est ancienne, plus sa valeur est faible et plus son propriétaire est incité à ne pas faire des dépenses élevées pour l'assurer. D'ailleurs, les véhicules les plus âgés sont moins souvent garantis en formules *tous risques*. Le fait que plus la voiture est ancienne plus la couverture est faible, incite l'assuré à ne pas déclarer tous les sinistres responsables qu'il peut dédommager lui même, et cela à cause de la faible couverture d'assurance (ou respectivement de la forte franchise). La variable de l'ancienneté du véhicule est supposée influencer le choix de la couverture d'assurance mais pas directement la déclaration des sinistres responsables. Suite à ce raisonnement, nous introduisons la variable de l'ancienneté du véhicule dans l'équation du choix du contrat et nous l'écartons de l'équation de la sinistralité.

Les résultats montrent que l'ancienneté du véhicule est significativement corrélée au type du contrat d'assurance. Ceci confirme sa validité en tant que variable instrumentale. Les coefficients estimés du modèle probit bivarié sont reportés au tableau 2.15 où les jeunes conducteurs et les conducteurs expérimentés sont présentés séparément.

Tab. 2.15: **le model probit bivarié**

	Les jeunes conducteurs			**Les conducteurs expérimentés**		
Variables	**Paramètres**	**Erreur type**		**Paramètres**	**Erreur type**	
	Equation de la couverture d'assurance					
Constante	-1.247670	0.632389	**	-0.459533	0.059945	***
Homme	-0.089760	0.110755		0.171433	0.014275	***
Age du conducteur	0.016211	0.018376		-0.017120	0.001407	***
Enseignant	0.080669	0.110277		-0.157704	0.013948	***
Zone urbaine	-0.077095	0.168721		-0.044736	0.020773	**
Ancienneté du permis	-0.098158	0.076261		-0.008473	0.001612	***
Conducteur principal	0.019323	0.188699		0.002233	0.014548	
Bonus malus	0.006020	0.004835		0.009672	0.000497	***
Ancienneté du véhicule	0.142644	0.014815	***	0.105343	0.001850	***
Puissance du véhicule	0.008580	0.003994	**	0.004576	0.000282	***
	Equation de l'occurrence d'au moins un sinistre responsable					
Constante	-5.681965	0.685093	***	-1.964406	0.079771	***
Homme	0.166087	0.115496		-0.042434	0.017632	**
Age du conducteur	0.001476	0.014121		-0.003237	0.001711	*
Enseignant	0.046583	0.112667		-0.033896	0.017064	**
Zone urbaine	-0.234092	0.186785		-0.025113	0.025322	
Ancienneté du permis	0.407441	0.085475	***	0.006392	0.001955	***
Conducteur principal	-0.215001	0.178660		0.041492	0.017670	**
Bonus malus	0.040939	0.004403	***	0.016192	0.000514	***
Puissance du véhicule	-0.000323	0.003671		0.003513	0.000310	***
Formule tous risques	-0.283383	0.477063		-0.880578	0.044432	***
Rho	0.038755	0.268827		0.342960	0.027917	***

La significativité au seuil de 10%, 5% et 1% est noté respectivement par $*$, $**$ et $***$

TAB. 2.16: **Les effets marginaux**

Variables	**Jeunes conducteurs**	**Conducteurs expérimentés**
Equation de la couverture		
Homme	n.s.	0.0558628
Age du conducteur	n.s.	-0.0055786
Enseignant	n.s.	-0.0513891
Zone urbaine	n.s.	-0.0145776
Ancienneté du permis	n.s.	-0.0027609
Conducteur principal	n.s.	n.s.
Bonus malus	n.s.	0.0031516
Ancienneté du véhicule	0.0233424	0.0343267
Puissance du véhicule	0.0014040	0.0014912
Equation du sinistre		
Homme	n.s.	-0.0085083
Age du conducteur	n.s.	-0.000649017
Enseignant	n.s.	-0.0067965
Zone urbaine	n.s.	n.s.
Ancienneté du permis	0.0603743	0.0012817
Conducteur principal	n.s.	0.0083196
Bonus malus	0.0060663	0.0032466
Puissance du véhicule	n.s.	0.000704348
Formule tous risques	n.s.	-0.1765644

Chaque valeur représente l'effet marginal de la variable X_k sur la probabilité que $Y = 1$

ns signifie que la variable est non significative

Le premier résultat à souligner est que l'effet direct du choix du contrat d'assurance sur la probabilité de déclarer au moins un sinistre responsable n'est pas significatif pour les jeunes conducteurs (γ_{jeunes} est égal à zéro). La corrélation entre les termes des résidus des deux équations n'est pas non plus significative. Ils n'existe pas d'interdépendance entre la couverture et la déclaration des accidents responsables à travers les caractéristiques individuelles inobservables. Ceci confirme les résultats de Chiappori et Salanié (2000), à savoir l'hypothèse d'asymétrie d'information est rejetée pour les jeunes conducteurs. Ces résultats peuvent être interprétés par le fait que l'assureur détient toutes les informations nécessaires sur les jeunes conducteurs pour contrôler les problèmes d'asymétrie d'information.

Par contre, pour les conducteurs expérimentés, les résidus de l'équation de la couverture d'assurance et ceux de l'équation de l'occurrence des sinistres sont significativement et positivement corrélés. Nous ne rejetons donc pas le fait qu'il existe des facteurs inobservables par l'assureur qui affectent en même temps le choix du contrat d'assurance et la sinistralité. L'hypothèse d'asymétrie d'information est ainsi vérifiée chez les conducteurs expérimentés. Nous retrouvons ici les mêmes résultats que ceux de Cohen (2005).

Ce modèle probit bivarié modélise la sinistralité et le choix du contrat d'assurance, en condensant les différentes garanties tous risques en une seule modalité. Notre fichier est plus riche en information sur les couvertures d'assurance que celle de Chiappori et Salanié (2000). Elle comprend des variables sur plusieurs contrats tous risques avec différents niveaux de franchises. Nous exploitons ces informations et nous soulignons qu'en intégrant les différents contrats d'assurance, le modèle se rapproche plus de la réalité. En effet, les compagnies d'assurance proposent plusieurs offres aux agents en diversifiant les garanties tous risques optionnelles.

2.4.4 Le modèle bivarié ordonné récursif

Nous avons montré à partir de la modélisation logit univarié qu'il ressort une contribution significative de la couverture d'assurance sur la sinistralité si toutes les formules d'assurance sont prises en considération, et ce principalement pour les jeunes conducteurs. C'est la raison pour laquelle nous appliquons un modèle bivarié ordonnée récursif. Le caractère ordonnée du modèle vient du fait que nous introduisons une variable ordonnée vu la nature cumulative des quatre contrats d'assurance.

$$\begin{cases} C^\star = \alpha X_1 + \varepsilon_1 \\ Y^\star = \beta X_2 + \gamma C + \varepsilon_2 \end{cases}$$

$Y^\star$ est la même variable latente continue que le modèle précédent associée à la variable observée Y ($Y = 1$ *si* $Y^\star > 0$ *et* $Y = 0$ *si* $Y^\star \leq 0$) décrivant l'occurrence des sinistres au cours de l'année telle que :

$$Y = \begin{cases} 1 & \text{si l'assuré a déclaré au moins un sinistre responsable,} \\ 0 & \text{si l'assuré n'a enregistré aucun sinistre responsable pendant l'année.} \end{cases}$$

$C^\star$ est la variable latente continue associée à la variable observée C telle que :

$$\begin{cases} C_i = 0 & \text{si} \quad C_i^\star < l_1 \\ C_i = 1 & \text{si} \quad l_1 < C_i^\star < l_2 \\ C_i = 2 & \text{si} \quad l_2 < C_i^\star < l_3 \\ C_i = 3 & \text{si} \quad l_3 < C_i^\star \end{cases}$$

La variable observée du choix de contrat d'assurance est définie comme suit :

$$C = \begin{cases} 0 & \text{si l'assuré a souscrit le contrat RC,} \\ 1 & \text{si l'assuré a souscrit la Formule 1,} \\ 2 & \text{si l'assuré a souscrit la Formule 2,} \\ 3 & \text{si l'assuré a souscrit la Formule 3.} \end{cases}$$

Les limites l_i sont inconnues et doivent être estimées. Pour conserver la constante dans le modèle, nous imposons la nullité de la première limite l_1. Afin d'assurer l'identification du modèle, nous utilisons la même variable instrumentale que celle du modèle précédent, soit l'ancienneté du véhicule.

Les termes d'erreurs sont supposés corrélés et suivant la loi normale :

$$\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix} \approx N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right]$$

Comme dans le modèle bivarié dichotomique, si la corrélation ρ est nulle alors les deux équations peuvent être estimées séparément et le caractère endogène de la variable décrivant le type du contrat est rejeté. Si ρ est significativement différent de zéro, il existe une interdépendance conditionnelle entre la sinistralité et les couvertures d'assurance, et l'hypothèse d'asymétrie d'information est ainsi vérifiée.

Les résultats de l'estimation sont présentés dans le tableau 2.17. Les effets marginaux sont présentés en annexe 4.

TAB. 2.17: **Le modèle probit ordonné bivarié**

	Les jeunes conducteurs			**Les conducteurs expérimentés**		
Variables	**Paramètres**	**Erreur type**		**Paramètres**	**Erreur type**	
	Equation de la couverture d'assurance					
Constante	1.420036	0.373577	***	0.251873	0.049264	***
Homme	-0.019296	0.066405		0.085795	0.012002	***
Age du conducteur	-0.009603	0.009545		-0.013261	0.001216	***
Enseignant	0.125699	0.065966	*	-0.097294	0.011750	***
Zone urbaine	0.028391	0.100867		-0.020046	0.017667	
Ancienneté du permis	0.043790	0.044472		-0.005975	0.001390	***
Conducteur principal	0.048984	0.111113		0.002417	0.012266	
Bonus malus	0.002407	0.003088		0.004515	0.000397	***
Ancienneté du véhicule	-0.051645	0.006756	***	0.019505	0.001132	***
Puissance du véhicule	0.002437	0.002172		0.003567	0.000230	***
Limite 1	1.171767	0.048652	***	0.441549	0.005027	***
Limite 2	1.708783	0.054715	***	0.678268	0.006019	***
	Equation de l'occurrence d'au moins un sinistre responsable					
Constante	-4.348312	1.142567	***	-1.228825	0.171339	***
Homme	0.121086	0.113361		-0.005293	0.017142	
Age du conducteur	0.001096	0.013399		-0.006243	0.001658	***
Enseignant	0.053357	0.109492		-0.061754	0.015899	***
Zone urbaine	-0.224642	0.177264		-0.024587	0.022916	
Ancienneté du permis	0.364932	0.089082	***	0.003190	0.001871	*
Conducteur principal	-0.189183	0.172122		0.034662	0.016094	**
Bonus malus	0.038088	0.005734	***	0.015026	0.000642	***
Puissance du véhicule	0.000618	0.003776		0.003715	0.000286	***
Formule 1	-1.173750	0.353281	***	-1.083692	0.028794	***
Formule 2	-1.288521	0.517942	**	-1.096826	0.038352	***
Formule 3	-1.486673	0.797197	*	-1.553581	0.119435	***
Rho	0.507792	0.257349	**	0.674391	0.062056	***

La significativité au seuil de 10%, 5% et 1% est notée respectivement par $*$, $**$ et $***$

Le coefficient de corrélation entre les résidus des deux équations est significatif et positif quelque soit l'expérience de conduite de l'assuré. Ceci montre que l'offre de la couverture d'assurance est endogène et que son effet sur la sinistralité est sous-estimé dans un modèle logit simple.

Le modèle ordonné bivarié permet ainsi de réévaluer l'effet de la couverture d'assurance sur la sinistralité. Certes, les coefficients estimés des différentes formules d'assurance dans l'équation de la sinistralité sont négatifs, c'est-à-dire la couverture d'assurance affecte à la baisse la probabilité d'enregistrer des sinistres. Cependant, le coefficient de corrélation entre les termes d'erreurs des deux équations est significativement positif. Il en ressort un biais lié à l'hétérogénéité inobservée qui figure à la fois dans les perturbations de l'équation de l'occurrence du sinistre et ceux de l'équation du choix du contrat.

Cette hétérogénéité est interprétée par des caractéristiques individuelles inobservables qui peuvent varier d'un conducteur novice à un conducteur expérimenté. Les conducteurs expérimentés peuvent détenir des informations privées significatives sur leur type de risque en faisant une autocritique sur eux même ou en observant simplement leur comportent dans d'autres dimensions de leur vie courante. Ce qui peut être différent pour les jeunes conducteurs. En effet, une mauvaise évaluation de leur propre risque et une surestimation de leur habileté dans la conduite peuvent avoir des effets sur la sinistralité. Ces informations résumées par des variables comportementales qui concernent le mode de conduite sont difficilement observables (ou totalement inobservables) par l'assureur. Alors qu'il existe d'autres variables que l'assureur peut observer comme le nombre de kilomètres parcourus, mais qu'une multitude de compagnies françaises d'assurance ne demandent pas à ses assurés.

Nous avons montré que l'asymétrie d'information entre assureur et assuré peut subsister dans la sélection adverse parmi les deux sous populations de conducteurs. Cependant l'aléa moral peut expliquer en partie cette asymétrie d'information. C'est surtout pour les conducteurs expérimentés que ce phénomène peut être le plus significatif puisqu'ils ont suffisamment d'expérience pour savoir quel type de précaution il faut prendre et donc influencer leur niveau d'incitation à se prémunir contre le risque. Il est alors difficile de distinguer entre les deux types d'asymétrie d'information qu'il soit de la sélection adverse ou de l'aléa morale.

Ces conclusions diffèrent des résultats trouvés dans l'étude de Chiappori et Salanié (2000) où l'hypothèse de l'asymétrie d'information n'est pas acceptée. En effet, ils ont trouvé que l'asymétrie d'information est très négligeable chez les jeunes conducteurs ainsi que chez les conducteurs expérimentés. Cohen (2005) aussi a trouvé des résultats différents : l'asymétrie d'information est seulement vérifiée chez les conducteurs expérimentés. Chiappori et Salanié (2000) et Cohen (2005) n'ont pas pris en considération les différents niveaux de couvertures *tous risques*, c'est la raison pour laquelle nos résultats concernant l'asymétrie d'information sont différents. Cependant nous retrouvons les mêmes résultats que ceux des travaux de Grun-Rehomme et Benlagha (2007) en ce qui concerne les jeunes conducteurs. Les mêmes conclusions de l'existence de l'asymétrie d'information ont été soulignées dans l'étude récente de Kim et *al.* (2009) appliquée sur le marché coréen, où les auteurs ne rejettent pas l'hypothèse de l'asymétrie d'information.

Nous avons évoqué plus haut une limite à la modélisation appliquée sur des bases de données en coupe transversale. Cette limite concerne l'historique de la sinistralité des assurés. Chiappori et Salanié (2000) ont souligné que le fait de ne pas contrôler l'historique des assurés dans les estimations peut être une source d'hétérogénéité des individus. Des problèmes de mauvaise spécification peuvent ainsi apparaître, ce qui tend à sur-estimer le degré d'asymétrie d'information.

Certes, les informations décrites sur une seule période de contrat d'assurance comprennent le coefficient du bonus malus qui est calculé sur la base du nombre passé d'accidents responsables. Plusieurs études l'utilisent comme un indicateur de la sinistralité passée. D'ailleurs, c'est ce que nous venons d'appliquer dans cette section. Cependant, nous avons donné dans l'introduction un exemple qui montre que, toutes choses égales par ailleurs, le bonus malus peut ne pas être une mesure précise du type de risque *ex ante* de l'assuré, et crée par la suite des biais d'estimation. Nous proposons dans ce qui suit une méthode d'évaluation de la sinistralité passée des assurés.

2.5 La sinistralité passée et l'asymétrie d'information : une étude originale

Une première sous-section décrit la sinistralité passée. Ensuite, sa *proxy* est utilisée dans des modèles bivariés et trivariés afin d'en conclure de la présence de l'asymétrie d'information.

2.5.1 La sinistralité passée

Du fait de l'inversion du cycle de production en assurance, l'assureur se doit d'anticiper le montant des dommages qu'il devra rembourser. La charge financière occasionnée par une police du portefeuille est inconnue au début de la période d'assurance, alors que la prime est déjà encaissée. La prime proposée aux assurés est basée sur des études économétriques et actuarielles des données stockées les années précédentes.

La sinistralité passée constitue, parmi ces données, un élément important dans la mesure où elle peut apporter de l'information sur le niveau de risque de l'assuré. Ne pas utiliser de telle information dans l'approche de la corrélation conditionnelle peut générer des surestimations de l'asymétrie d'information et implique que les déclarations passées constituent une information privée de l'assuré par rapport à l'assureur, ce qui n'est sûrement pas le cas dans le système français d'assurance.

Nous essayons donc de mesurer le niveau de risque de chaque conducteur assuré en fonction de l'observation de sa sinistralité passée. Pour cette étude, il nous faudrait des données longitudinales, afin d'évaluer les changements de comportement liés à la sinistralité, toutes choses égales par ailleurs (Abbring et al., 2003a, Abbring et al., 2003b). La prise en compte de données historiques peut permettre éventuellement de préciser le sens de causalité entre le choix du contrat et la sinistralité. Le fait est que nous détenons des données en coupe instantanée décrivant une seule période. Comment alors observer (ou estimer) la sinistralité passée en absence de données longitudinales ? Comment peut-on contourner l'obstacle du caractère unidimentionnel des données ?

Nous pouvons aborder ce problème à partir du Coefficient de Réduction Majoration (CRM ou coefficient bonus malus) de chaque assuré, qui fait intervenir le nombre passé d'accidents responsables. Le CRM, seul, ne permet pas d'avoir une bonne représentation de la sinistralité passée pour les souscripteurs d'un contrat de moins de 13 ans. En effet, un jeune assuré depuis 5 ans, sans accident a un CRM égal à 77.5, alors qu'un assuré qui avait un CRM à 50 et qui vient d'avoir un sinistre responsable aura un CRM égale à 62.5. En comparant seulement les coefficients associés à ces deux assurés, on dira que le second conducteur est mieux que le premier en termes de risque. Ceci n'est pas vrai. L'information est alors corrigée en observant le nombre d'année de conduite de chaque assuré. Le CRM permet, en tenant compte de l'ancienneté du permis de conduire, d'avoir une bonne idée sur l'occurrence des sinistres, c'est-à-dire si l'assuré a déclaré au moins un accident responsable dans le passé ou non. Cependant, nous ne pouvons pas avoir la fréquence exacte des sinistres passés. Nous construisons dans ce qui suit cette variable d'occurrence des déclarations antérieures.

Notons S_p une variable binaire décrivant le passé des sinistres des assurés, et plus précisément l'occurrence des accidents responsables antérieurs. Elle est définie comme suit :

$$S_p = \begin{cases} 1 & \text{si l'assuré a déclaré au moins un accident responsable dans le passé,} \\ & \text{Nous le considérons comme un } \textit{mauvais conducteur}. \\ 0 & \text{si l'assuré n'a déclaré aucun accident responsable dans le passé,} \\ & \text{Nous le considérons comme un } \textit{bon conducteur}. \end{cases}$$

Plus formellement, si nous notons n_i l'ancienneté du permis de conduire qu'un assuré i possède et BM_i son coefficient du bonus malus, alors :

$S_{p_i} = 0$, si l'on est dans l'un des cas suivants
- $n_i \geq 13$ et $BM_i = 50$
- $n_i < 13$ et $BM_i = E^{n_i}(95)$

$S_{p_i} = 1$, si l'on est dans l'un des cas suivants
- $n_i \geq 13$ and $BM_i > 50$
- $n_i < 13$ and $BM_i > E^{n_i}(95)$

où $E^{n_i}(95)$ désigne la partie entière de la valeur obtenue à la $n_i^{\text{ème}}$ itération des puissances de 0,95, sachant qu'à chaque étape, on considère également la partie entière, notée E. La relation de récurrence s'écrit :

$$E^{n_i}(95) = E[E^{n_i-1}(95) \times (0.95)]$$

Le choix d'un seuil de 13 ans d'ancienneté de conduire est justifié par le fait qu'au bout de 13 années de conduite sans accidents responsables, l'assuré enregistre son coefficient maximum de bonus malus, soit 50.

Souvent les jeunes conducteurs se font déclarer en tant que conducteurs secondaires sur le véhicule de leurs parents. Ils peuvent ainsi acquérir leurs antécédents d'assurance en payant un supplément de cotisation. Ces antécédents d'assurance concerne le coefficient du bonus malus. En effet, dans notre fichier, 37% des jeunes conducteurs affichent des coefficients de bonus malus inférieurs aux coefficients qui leur deverait être accordés s'ils étaient déclarés comme conducteurs principaux. Lorsque la variable S_p est égale à zéro , ces 37% de jeunes conducteurs sont écartés de l'analyse. Par conséquent, nous évitons une surévaluation de la variable S_p des conducteurs que nous avons considéré comme bons et principalement les jeunes. Le nombre de jeunes conducteurs passe alors de 1149 assurés à 720 assurés sur les quels nous appliquons par la suite notre étude.

Le tableau 2.19 présente la distribution des assurés en fonction de la sinistralité passée S_p.

TAB. 2.18: **La distribution des sociétaires en fonction de la sinistralité passée**

%	**Jeunes conducteurs**	**Conducteurs expérimentés**
Bon conducteur	33.75	58.03
Mauvais conducteur	66.25	41.97

Nous devons souligner que cet indicateur ne constitue qu'une *proxy*. Il est clair qu'il ne s'agit que d'une estimation de la sinistralité passée. En effet, on pourrait déclarer comme mauvais conducteur, un bon conducteur. La définition retenue ici de *mauvais* conducteur peut être discutable. Par exemple, un conducteur qui n'a jamais eu d'accidents responsables (pendant 5 à 10 ans) et qui n'a pas contracté d'assurance pendant plusieurs années (son CRM n'a pas bougé durant cette interruption) sera considéré comme un mauvais conducteur dans notre approche, alors que l'on préférerait le considérer comme un bon conducteur. On peut donc avoir une légère surestimation des mauvais conducteurs, qu'il nous est impossible de corriger.

Tab. 2.19: **La distribution des sociétaires en fonction de la sinistralité passée et du choix du contrat**

%	**RC**	**Formule 1**	**Formule 2**	**Formule 3**
Jeunes conducteurs				
Bon conducteur	9.47	38.27	19.75	32.51
Mauvais conducteur	10.27	38.36	21.17	30.19
Conducteurs expérimentés				
Bon conducteur	51.85	12.93	7.19	28.03
Mauvais conducteur	31.09	21.29	10.81	36.81

Ces premiers résultats croisés entre la sinistralité passé et le type de contrat d'assurance ne nous permettent pas d'avoir une première idée sur une potentielle sélection adverse pour les jeunes conducteurs. Nous remarquons que les *bons* et les *mauvais* conducteurs sont distribués de façons presque similaire en fonction du type de contrat d'assurance. Ceci n'est pas le cas pour les conducteurs expérimentés, où 52% de *bons conducteurs* sont assurés en responsabilité civil contre 31% de *mauvais conducteurs*.

2.5.2 Endogénéité de la sinistralité passée et le modèle probit trivarié récursif

Le fait d'être bon ou mauvais conducteur (variable décrite par la sinistralité passée S_p) est sans doute lié aux différentes caractéristiques du conducteur et du véhicule. Ceci fait de la variable sinistralité passée, en plus du choix du contrat d'assurance et de la sinistralité présente, une variable endogène. Un modèle trivarié récursif (équations emboîtées) permet de prendre en compte l'interdépendance entre la sinistralité antérieure, le choix de contrat et la sinistralité déclarée l'année en cours. Nous proposons de lever l'hypothèse d'exogénéité de la sinistralité passée et d'analyser l'efficience comparée du choix de contrat et de la sinistralité passée.

Nous nous proposons donc de tester un modèle probit trivarié (dans ce cas, préférable au modèle Logit, pour éviter une sur pondération des cas extrêmes). Cette modélisation permet de tester à la fois le caractère explicatif de la sinistralité passée et du choix de contrat sur la sinistralité actuelle et leur endogénétité.

Plus précisément, le cadre formel de notre spécification est le suivant :

$$\begin{cases} S_p^\star = k_1 + \alpha_1 X_1 + \varepsilon_1 \\ Z^\star = k_2 + \beta_1 X_2 + \beta_2 S_p + \varepsilon_2 \\ Y^\star = k_3 + \gamma_1 X_3 + \gamma_2 Z + \gamma_3 S_p + \varepsilon_3 \end{cases}$$

Avec

- X_1 (resp. X_2 et X_3) est le vecteur des variables exogènes caractéristiques du conducteur et du véhicule expliquant la sinistralité passée (resp. le type du contrat et la sinistralité actuelle),
- Z le type de contrat souscrit est une variable binaire, contrat RC ou contrat *tous risques* (ce dernier regroupe les différents contrats *dommage au véhicule* en une seule modalité),
- S_p la variable binaire de la sinistralité passée définie plus haut,
- Y la variable de sinistralité (actuelle) égale à 0 si l'assuré n'a pas eu d'accident responsable durant l'année et égale à 1 sinon.

$Y^\star$, $Z^\star$ et $S_p^\star$ sont les variables latentes continues associées respectivement aux variables observées Y, Z et S_p (($S_p = 1$ *si* $S_p^\star > 0$ *et* $S_p = 0$ *si* $S_p^\star \leq 0$). Les termes d'erreurs ε_1, ε_2 et ε_3 des trois différentes équations sont supposés corrélés :

$$\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix} \approx N \left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{12} & 1 & \rho_{23} \\ \rho_{13} & \rho_{23} & 1 \end{pmatrix} \right]$$

Si ces coefficients sont significativement nuls, les différentes équations peuvent être estimées chacune indépendamment de l'autre.

Comme dans le modèle bivarié, certaines contraintes d'identification du modèle doivent être imposées afin d'estimer tous les paramètres. Du fait de la récursivité du modèle, les résidus des trois équations latentes n'étant pas indépendants, les paramètres du modèle ne peuvent pas être identifiés si les variables explicatives des différentes équations sont identiques (Maddala, 1983, pp. 117-147). Les vecteurs des différentes caractéristiques de l'assuré et du véhicule, X_1, X_2 et X_3, doivent ainsi être différents d'une équation à une autre. Ceci est bien évidemment justifié par le choix de la variable instrumentale, comme il a été décrit dans le modèle bivarié. Nous procédons comme suit pour le choix des variables instrumentales : Nous utilisons comme dans le modèle bivarié précédent la variable de l'ancienneté du véhicule comme un instrument dans l'équation du choix du contrat d'assurance. C'est à dire nous l'écartons de l'équation de la sinistralité et nous le gardons dans l'équation de la couverture. Les raisons du choix de cette variable restent les mêmes que dans le modèle bivarié. En ce qui concerne l'instrument de la sinistralité passée, nous testons plusieurs candidats parmi les variables exogènes (sauf bien sûre l'ancienneté du véhicule). Nous validons le type du conducteur comme variable instrumentale. La contribution de cette variable sur le choix du contrat d'assurance transite par son effet direct sur la sinistralité passée qui constitue une variable explicative de la couverture.

Nous présentons le modèle en considérant le vecteur (Y, Z, S_p). Pour chaque observation i (un couple (assuré, véhicule)), (Y_i, Z_i, S_{pi}) peut prendre huit triplets binaires différents. La modalité de référence est $(0, 0, 0)$, à savoir les bons conducteurs qui ont souscrit un contrat DV et avec une sinistralité actuelle nulle.

La difficulté réside dans l'estimation des paramètres du modèle. Les techniques d'estimation des modèles multivariés à variables discrètes se sont développées seulement dans les années 90, raison pour laquelle ces modèles n'ont été appliqués que récemment dans différents domaines tels que : le marché de travail (Havet, 2006), l'assurance santé (Gibbons et al. 1998, Zhao et al., 2008), les sciences politiques (Lawrence, 1997), le marketing (Smith et Danaher, 2009). Sur le marché d'assurance, il n'existe que quelques travaux empiriques très récents s'appuyant sur ce type de modèle. Nous citons l'étude originale de Young et *al.* (2009). Ces auteurs utilisent un portefeuille provenant d'une compagnie d'assurance automobile de Singapour. Ils modélisent les différents types d'accidents réclamés à l'assureur (dommages corporels, dommages matériels, vol,...), en s'appuyant sur un modèle probit multivarié.

Pour estimer les intégrales triples dans la fonction de la log vraisemblance, une méthode du maximum de vraisemblance simulée (SML) a été développée dans les années 80 par Lerman et Manski (1981). Hajivassiliou, McFadden et Ruud (1996) sont parvenus à montrer, dans leur étude approfondie sur les probabilités simulées, que la manière la plus fiable et la plus précise pour simuler les probabilités issues de la loi normale multivariée est le simulateur GHK. Ce dernier a été suggéré de façon indépendante par Geweke (1991), Hajivassiliou (1990) et Keane (1994), d'où GHK. Il s'agit d'une méthode d'approximation de la valeur des probabilités issues de la loi normale trivariée . Les estimateurs sont donc les solutions de la maximisation de la log-vraisemblance simulée. Les résultats de l'estimation[11] du probit trivarié sont présentés dans le tableau 2.20 suivant.

[11] L'estimation a été réalisée avec STATA par la commande *triprobit* (voir Terracol, A. (2002), *triprobit and the GHK simulator : a short note*, annexe à la commande Stata triprobit)

Tab. 2.20: **Le modèle probit trivarié**

	Les jeunes conducteurs			**Les conducteurs expérimentés**		
Variables	**Paramètres**	**Erreur type**		**Paramètres**	**Erreur type**	
	Equation de la sinistralité passée					
Homme	-.1178012	.1008155		.056132	.0141795	***
Age du conducteur	.0458181	.0167754	**	-.0483056	.0005185	***
Enseignant	.0649594	.0993508		-.0410787	.0138607	**
Zone rurale	-.0707622	.1544051		-.0826979	.0136093	***
Conducteur principal	.4400459	.1397856	**	.20658	.014602	***
Ancienneté du véhicule	.0061595	.0096397		-.0009167	.0014543	
Puissance du véhicule	-.000153	.0032415		-.0027981	.000279	***
Constante	-.6219816	.4172047		2.187976	.0385154	***
	Equation de la couverture d'assurance					
Homme	.1155049	.1220986		.1658623	.0138572	*
Age du conducteur	-.0194789	.0122861		-.0206628	.0005983	***
Enseignant	-.0412597	.119855		-.1610393	.0137798	***
Zone rurale	.0615026	.1793348		-.0586215	.013491	***
Ancienneté du véhicule	.1103306	.0163044	***	.1027508	.0015911	***
Puissance du véhicule	.0072914	.0041408	*	.0042651	.0002738	***
Sinistralité passée	1.373085	.1652415	***	.5014691	.0216091	***
Constante	-.8830826	.4627834	*	-.0721895	.0451707	
	Equation de la sinistralité actuelle					
Homme	.1438004	.1164947		-.0605599	.0169848	**
Age du conducteur	.0252188	.0113748	**	-.0763325	.0165651	**
Enseignant	.0271783	.1105018		-.0019205	.0010382	***
Zone rurale	.1049082	.180478		-.0230261	.01699	**
Puissance du véhicule	.0030218	.0035312		.0030114	.0003114	
Sinistralité passée	-.7258632	.3436088	**	.148588	.0453421	
Contrat tous risques	.3130471	.217915		-.5187596	.0251578	***
Constante	-1.706576	.4379503	***	-1.060746	.0833884	***
$\rho_{\textbf{sinistralié passée - couverture}}$	-.80497639	.05727343	***	-.12374225	.01109357	***
$\rho_{\textbf{sinistralié passée - sinistralité actuelle}}$	.77379352	.1344513	***	.14536595	.02506752	***
$\rho_{\textbf{sinistralié actuelle - couverture}}$	-.4974583	.11159986	***	.10010997	.01133842	***

La significativité au seuil de 10%, 5% et 1% est notée respectivement par $*$, $**$ et $***$

Notre modèle trivarié peut être assimilé à un modèle dynamique décrivant deux périodes de sinistralités : la sinistralité au temps $t-1$ et la sinistralité au temps t. Chiappori et *al.* (2006) montrent qu'en présence d'aléa moral, les accidents de l'année t et ceux de l'année $t-1$ sont négativement corrélés. Abbring et *al.* (2003) se fondent sur ces prédictions théoriques pour pouvoir distinguer entre l'aléa moral et la sélection adverse. En utilisant des données françaises multidimentionnelles, où ils n'avaient qu'une dynamique de la sinistralité, ils rejettent l'hypothèse de l'aléa moral. Nous rappelons que Dionne et *al.* (2006), en faisant intervenir quant à eux la dynamique des sinistres et celle des contrats d'assurance, montrent que :

- si la couverture en $t-1$ et les accidents en t sont positivement corrélés alors il existe de l'aléa moral,
- si la corrélation entre les accidents en $t-1$ et la couverture en t est positive alors il existe de l'apprentissage asymétrique (*learning*),
- en situation de sélection adverse, la couverture en $t-1$ et les accidents en t présentent une corrélation nulle,
- et si la corrélation entre les accidents en t et la couverture en t est positive alors il existe de l'asymétrie d'information résiduelle (aléa moral et/ou sélection adverse).

En s'appuyant sur ces différentes prédictions et ne détenant pas de l'information sur les types de contrats souscrits avant l'année 2004, nous ne pouvons répondre qu'aux questions de l'apprentissage et de l'asymétrie d'information résiduelle. Quant au test de l'aléa moral, nous l'appliquons au sens de Chiappori et *al.* (2006).

En ce qui concerne l'apprentissage asymétrique, Dionne et al. (2006) avancent que les assurés qui causent plus d'accidents[12] dans le passé apprennent, plus vite que les assureurs (d'où le terme apprentissage asymétrique), qu'ils sont plus susceptibles d'être des mauvais risques et choisiraient par la suite de souscrire des contrats à couvertures plus larges.

Le tableau de l'estimation du modèle probit trivarié fait ressortir les principaux résultats suivants. Pour les jeunes conducteurs ainsi que les conducteurs expérimentés, les termes d'erreurs des différentes équations sont deux à deux significativement corrélés. Plus précisément, la corrélation entre les perturbations de la sinistralité passée et celles du choix de la couverture d'assurance est négative. Plus le conducteur est risqué dans le passé, plus il souscrit des couvertures au tiers. L'asymétrie en matière d'apprentissage entre assureur et assuré n'existe donc pas quelque soit l'expérience de conduite des conducteurs.

En ce qui concerne la corrélation entre les résidus de la sinistralité passée et ceux des déclarations actuelles, le coefficient est significativement positif. D'après les prédictions théoriques de Chiappori et *al.* (2006), le problème d'aléa moral est inexistant. Les comportements des conducteurs face au risque routier, quelque soit jeunes ou expérimentés, ne change pas au cours du temps. Les mauvais conducteurs restent mauvais et de la même manière pour les bons conducteurs. Nous en concluons que la sinistralité peut être considérée comme un risque intrinsèque et donc une caractéristique propre à l'assuré qui ne change pas et ne s'influence pas par la nature du contrat. Ainsi les conclusions faites dans les précédentes sections concernant la nature de l'asymétrie d'information (où les modèles probits ordonnés bivariés où sont modélisés simultanément la sinistralité et le choix du contrat d'assurance) peuvent être réduites aux seuls problèmes d'antisélection et non plus d'un mélange des deux phénomènes d'asymétrie d'information. Ce ré-

[12] les accidents effectivement causés par les assurés et non pas les accidents déclarés à l'assureur.

sultat est intéressant dans le sens où cette modélisation trivariée nous a permis de distinguer l'aléa moral de l'antisélection et affiner les conclusions des analyses basées sur l'approche de la corrélation positive.

Le dernier coefficient décrit la corrélation entre la sinistralité actuelle et le choix de la couverture d'assurance. Le signe de ce coefficient est différent entre les jeunes conducteurs et les conducteurs expérimentés. Pour les conducteurs expérimentés, le signe de la corrélation entre les termes d'erreur est significativement positif. Il existe alors de l'asymétrie d'information résiduelle chez les conducteurs expérimentés. Cependant, les jeunes conducteurs présentent une corrélation significativement négative. Ceci signifie que l'hétérogénéité inobservée relative à l'équation de la sinistralité actuelle et celle qui figure dans les perturbations de la couverture d'assurance n'interagissent pas dans le même sens. Dans d'autres termes, les jeunes conducteurs les plus exposés au risque d'accident optent pour une faible couverture et vice versa. Les caractéristiques individuelles inobservées représentant cette hétérogénéité reflètent dans ce cas de l'aversion au risque de la par des jeunes conducteurs. Si nous observons avec plus de détail l'équation de la sinistralité actuelle, nous remarquons que l'effet direct du choix du contrat sur l'occurrence des sinistres n'est pas significatif. Ceci n'est pas le cas dans la modélisation bivariée de la sinistralité et la couverture d'assurance, ce qui peut expliquer en partie ce signe négatif du coefficient.

Plusieurs critiques peuvent être faite suite à cette modélisation. Elles concernent les variables explicatives de l'occurrence de sinistres antérieurs. En effet, nous supposons que les informations qui existent dans le fichier sont fixées dans le temps. En d'autre termes, nous supposons que l'assuré ne change pas de voiture, ni de profession, ou de lieu d'habitation. Ces hypothèses sont ainsi très restrictives. C'est la raison pour laquelle nous utilisons dans la partie suivante l'indicateur de la sinistralité passée non pas comme une variable endogéne mais comme un moyen de construire une deuxième *proxy* de l'effort. Cette dernière nous permet de valider empiriquement les prédictions théoriques que nous avons déjà établies dans le chapitre 1.

2.6 L'aversion à l'effort de prévention

Le modèle théorique que nous avons établit, décrit dans le chapitre 1, introduit l'aversion à l'effort comme variable décrivant le type de risque de l'agent. Elle constitue le paramètre de la sélection adverse.

La variable que nous construisons est donc celle qui décrit l'aversion à l'effort de l'assuré. La difficulté réside dans le fait que cette variable constitue une information non observable et non quantifiable par l'assureur, donc difficilement mesurable.

Nous mesurons le niveau d'aversion à l'effort de chaque assuré en fonction de l'occurrence des sinistres passés et les accidents responsables actuels. L'idée est d'observer l'évolution entre la sinistralité passée et la sinistralité actuelle. Notre légitimité de construire une variable décrivant une caractéristique propre à l'assuré, en fonction des accidents, est fondée sur le fait que la sinistralité représente un risque intrinsèque au conducteur, chose que nous avons montré dans la section précédente. Ceci nous permet d'avoir une nouvelle mesure plus fine du risque intrinsèque de l'assuré qui prend en compte à la fois le bonus malus, l'ancienneté du permis de conduire et les accidents responsables déclarés. En plus cette variable est indépendante des potentiels changements d'informations de l'assuré ou de son véhicule dans le temps, chose que nous avons soulevé dans la section précédente.

Ainsi, nous nous n'appuyons sur aucune hypothèse à l'égard des caractéristiques observables de l'assuré. Ce qui donnera des résultats plus fiables en ce qui concerne le test de l'hypothèse de l'anti-sélection.

Cette variable décrivant l'aversion à l'effort est construite plus formellement comme suit :

- Si $S_p = 1$ et $Y = 1$, c'est à dire l'assuré a enregistré au moins un accident responsable avant 2004 et continue d'enregistrer des accidents responsables au cours de l'année 2004, alors nous le considérons comme un agent qui a une très « *forte aversion à l'effort* ».

- Si $S_p = 0$ et $Y = 0$, c'est à dire l'assuré n'a pas déclaré d'accidents responsable avant 2004, ni au cours de l'année 2004, alors nous le considérons comme un agent qui a une très « *faible aversion à l'effort* » de prévention.

- Dans les autres cas, c'est à dire $S_p = 0$ et $Y = 1$ ou $S_p = 1$ et $Y = 0$, nous considérons l'assuré comme un agent ayant une « *moyenne aversion à l'effort* ».

La variable de l'aversion à l'effort θ se présente comme suit :

$$\theta = \begin{cases} 0 & \textit{Faible aversion à l'effort} \\ 1 & \textit{Moyenne aversion à l'effort} \\ 2 & \textit{Forte aversion à l'effort} \end{cases}$$

TAB. 2.21: **Répartition de l'échantillon en % selon θ**

θ	**Jeunes**	**Expérimentés**
0	31.94	52.32
1	56.81	41.52
2	11.25	6.16

TAB. 2.22: **L'aversion à l'effort et le type du contrat d'assurance**

θ	**Jeunes**				**Expérimentés**			
%	**RC**	**Formule 1**	**Formule 2**	**Formule 3**	**RC**	**Formule 1**	**Formule 2**	**Formule 3**
0	9.13	40.00	19.57	31.30	50.18	14.04	7.58	28.20
1	9.54	39.85	21.76	28.85	33.83	20.76	10.29	35.11
2	14.81	25.93	18.52	40.74	46.05	7.67	7.59	38.70

Remarque 2 *Une limite à ce schéma peut être soulevée. Les assurés que nous avons désignés comme des conducteurs moyens peuvent être des mauvais comme ils peuvent être de bons conducteurs. Mais le fait de ne pas observer les fréquences exactes des sinistres survenus avant l'année 2004 nous laisse incapable de faire cette distinction et d'être précis sur le niveau exact d'aversion à l'effort qui décrit les conducteurs moyens.*

Nous appliquons un modèle probit bivarié récursif qui estime de manière simultannée l'aversion à l'effort et le choix du contrat d'assurance. La variable de l'aversion à l'effort qui décrit le risque intrinsèque de l'assuré dépend des caractéristiques observables propres à l'assuré et à son véhicule. Elle est introduite comme élément prépondérant de l'équation du choix de la couverture d'assurance, conditionnellement à toutes les observables. Les résidus de l'équation de l'aversion à l'effort et celle des contrats d'assurance sont supposés corrélés. Si la corrélation est significativement positive, l'hypothèse de la sélection adverse est vérifiée. Ce modèle nous permet donc d'examiner l'existence de biais liés aux phénomènes d'anti-sélection et non pas de l'aléa moral.

Le modèle probit ordonné bivarié se présente formellement comme comme suit :

$$\begin{cases} \theta^\star = \beta X_2 + \varepsilon_1 \\ C^\star = \alpha X_1 + \gamma\theta + \varepsilon_2 \end{cases}$$

où X_1 et X_2 représentent les vecteurs des caractéristiques observables de l'assuré et du véhicule.

$\theta^\star$ est la variable latente continue associée à la variable observée de l'aversion à l'effort θ telle que :

$$\begin{cases} \theta = 0 \quad \text{if} \quad \theta^\star < l_1 \\ \theta = 1 \quad \text{if} \quad l_1 < \theta^\star < l_2 \\ \theta = 2 \quad \text{if} \quad l_2 < \theta^\star \end{cases}$$

Nous rappelons $C^\star$ la variable latente continue associée à la variable du type de contrat observée C telle que :

$$\begin{cases} C_i = 0 & \text{if} \quad C_i^\star < l_1 \\ C_i = 1 & \text{if} \quad l_1 < C_i^\star < l_2 \\ C_i = 2 & \text{if} \quad l_2 < C_i^\star < l_3 \\ C_i = 3 & \text{if} \quad l_3 < C_i^\star \end{cases}$$

La variable observée du choix de contrat d'assurance est définie comme suit :

$$C = \begin{cases} 0 & \text{si l'assuré a souscrit le contrat RC,} \\ 1 & \text{si l'assuré a souscrit la Formule 1,} \\ 2 & \text{si l'assuré a souscrit la Formule 2,} \\ 3 & \text{si l'assuré a souscrit la Formule 3.} \end{cases}$$

Les termes d'erreurs sont supposés corrélés comme suit :

$$\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix} \approx N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right]$$

Nous imposons la nullité des premières limites afin de garder les constantes dans le modèle. Les résultats sont représentés dans le tableau 2.23.

TAB. 2.23: **Le modèle ordonné bivarié**

	Les jeunes conducteurs			Les conducteurs expérimentés		
Variables	**Paramètres**	**Erreur type**		**Paramètres**	**Erreur type**	
	Equation de l'aversion à effort					
Constante	-1.382310	0.257991	***	1.701870	0.034142	***
Homme	0.047443	0.071511		0.015036	0.012435	
Age du conducteur	0.057514	0.009469	***	-0.033510	0.000462	***
Enseignant	-0.110374	0.054914	**	-0.071944	0.010203	***
Zone urbaine	-0.027285	0.069973		-0.081392	0.011980	***
Conducteur principal	0.535535	0.111953	***	0.157979	0.012882	***
Ancienneté du véhicule	0.005899	0.007174		-0.013934	0.001292	***
Puissance du véhicule	-0.000694	0.002370		-0.001410	0.000244	***
Limite 2	1.434272	0.060039	***	1.613766	0.010918	***
	Equation de la couverture d'assurance					
Constante	-0.053522	0.289116		0.154298	0.037146	***
Homme	0.022970	0.063925		0.057196	0.011528	***
Age du conducteur	0.030971	0.009447	***	0.001456	0.000896	
Zone urbaine	-0.006067	0.062827		0.016189	0.011149	
Conducteur principal	0.359554	0.107982	***	-0.071140	0.011961	***
Ancienneté du véhicule	-0.039247	0.006865	***	0.023145	0.001133	***
Puissance du véhicule	0.001614	0.002102		0.003463	0.000221	***
Faible aversion à l'effort	1.107469	0.086434	***	-1.137777	0.031013	***
Forte aversion à l'effort	-0.953651	0.119663	***	0.979654	0.036245	***
Limite 2	0.898586	0.057437	***	0.353555	0.007185	***
Limite 3	1.301862	0.077168	***	0.543795	0.010320	***
Rho	0.780063	0.049759	***	-0.720939	0.021609	***

La significativité au seuil de 10%, 5% et 1% est notée respectivement par $*$, $**$ et $***$

Le principal résultat qui ressort de ces estimations est que les coefficients de corrélations entre les perturbations des deux équations de l'aversion à l'effort et du choix du contrat d'assurance sont significativement différents de zéro. Cependant, le signe de la corrélation n'est pas le même entre les jeunes conducteurs et les conducteurs expérimentés. Nous rappelons d'après nos résultats théoriques qu'il existe en situation d'asymétrie d'information des contrats d'équilibres séparateurs et des contrats mélangeants. Ces derniers types de contrats attirent les agents de différents niveaux d'aversion à l'effort : les agents qui ont de faibles aversions à l'effort et d'autres qui ont de fortes aversion à l'effort. Nous vérifions ainsi notre prédiction théorique en montrant suite à ce résultat empirique qu'il existe de l'hétérogénéité non observée avec des contrats attirant les différents types d'aversion à l'effort.

En effet, pour les jeunes conducteurs, la corrélation est positive ce qui signifie que l'hypothèse de l'anti-sélection est vérifiée. Nous confirmons le fait qu'il existerait des variables comportementales non observées par l'assureur décrivant le mode de conduite comme par exemple conduire en état d'ébriété, ne pas porter la ceinture de sécurité qui sont des comportements plus ou moins répandus chez les jeunes conducteurs. Les conducteurs expérimentés quant à eux présentent une corrélation négative qui traduit une hétérogénéité non observée reflétant plus de l'aversion au risque : les moins averses à l'effort, donc les moins exposés au risque d'accident choisissent la couverture la plus forte.

2.7 Conclusion

Ce chapitre a soulevé les problèmes liés à l'asymétrie d'information en utilisant des extensions aux modèles de Chiappori et Salanié (2000) et en utilisant de nouvelles variables décrivant le risque lié à l'assuré.

Le premier résultat principal concerne le risque intrinsèque à l'assuré. Nous avons montré que les mauvais conducteurs restent de mauvais conducteurs, et les bons conducteurs restent des bons conducteurs. Ce résultat va à l'encontre de l'idée qu'avec l'âge, l'expérience de la sinistralité, le changement de situation familiale et professionnelle, les conducteurs deviennent plus responsables, plus respectueux des codes de la société et des autres. Ceci est d'autant plus confirmé avec le rejet de l'hypothèse de l'aléa moral. Les comportements des conducteurs face au risque routier ne changent pas. La corrélation entre les résidus de la sinistralité passée et ceux des sinistres actuels est significativement positive.

C'est à partir de cette constatation qu'il nous a été possible de créer une variable décrivant l'aversion à l'effort. Elle correspond à son risque intrinsèque. Ceci nous a permis de tester l'hypothèse de la pure sélection adverse sans que ce problème d'asymétrie d'information soit mélangé a de l'aléa moral. La modèlisation bivariée du choix de la couverture d'assurance et de l'aversion à l'effort a ciblé le problème de l'anti-sélection, au moyen de données en coupe transversale.

La construction de la variable aversion à l'effort comme caractéristique intrinsèque liée au risque de l'assuré, qui fait intervenir à la fois l'ancienneté de permis, le coefficient du bonus malus et la déclaration des sinistres actuels, a permis d'élargir la classe des bons conducteurs à des jeunes conducteurs souvent considérés comme des mauvais risques.

En effet, nous avons montré que parmi ces jeunes conducteurs, il existe environ 20% qui sont des bons conducteurs, non averse à fournir de l'effort de prévention contre le risque d'accident. Et puisqu'il a été montré que les bons conducteurs restent des bons conducteurs, alors cette proportion des jeunes conducteurs peut être considérée par l'assureur comme des bons risques et ainsi leur proposer les mêmes tarifs que les conducteurs expérimentés. Alors, dans ce cas, il n'est pas indispensable d'appliquer une surprime à ces jeunes conducteurs et le fait de majorer leur prime ne serait pas justifié. Dans le chapitre suivant, nous répondons à cette question de surprime appliquée au jeunes conducteurs.

Chapitre 3

Etudes de la sinistralité et de la prime pure chez les conducteurs novices et les conducteurs expérimentés

3.1 Introduction

L'objectif de ce chapitre est de comparer la sinistralité entre deux catégories de conducteurs : les jeunes conducteurs et les conducteurs expérimentés. Pour ce faire, plusieurs méthodes d'estimation sont appliquées sur le nombre de sinistres déclarés, d'une part, et leur coûts, d'autre part.

L'intérêt de cette comparaison, et plus particulièrement, celle appliquée sur les coûts des sinistres, est d'étudier la suppression de la majoration appliquée quasi systématiquement par les assureurs sur les primes des conducteurs novices par rapport aux conducteurs expérimentés. En effet, Les compagnies d'assurance en France ont la possibilité de rajouter une surprime pour les jeunes conducteurs sur leur contrat d'assurance qui peut atteindre un maximum de 100% de la cotisation de base. Pour inciter les jeunes conducteurs à la prudence, la surprime est réduite de moitié par année sans accident ayant engagé la responsabilité du conducteur. Elle est supprimée après deux années. La surprime est appliquée avant la clause de bonus-malus, de sorte que la réduction serait plus forte pour le conducteur novice prudent, mais la majoration est plus élevée pour l'imprudent. Par contre, des avantages tarifaires sont appliqués pour les conducteurs qui ont bénéficié d'un apprentissage anticipé de la conduite. La surprime passe à 50% au moment de la souscription.

Certes, l'application de la surprime constitue, de la part des compagnies d'assurance, une manière de lutter contre l'insécurité routière, en ciblant les conducteurs novices considérés comme les plus risqués. Selon la Fédération Française des Sociétés d'Assurance (FFSA), "les personnes dont le permis est récent, en particulier les jeunes, provoquent plus d'accidents que la moyenne des conducteurs. Un conducteur âgé de 18 à 20 ans provoque environ deux fois et demi plus d'accidents qu'un conducteur âgé de 30 ans ou plus". Mais, au lieu de sanctionner les jeunes conducteurs, l'assureur aurait-il d'autres alternatives à l'application de la surprime, lui serait-il plus profitable (à long terme) de fidéliser les jeunes conducteurs en leur proposant des prix similaires aux conducteurs expérimentés ? On dira dans ce cas que les bons risques quitteront la compagnie d'assurance si les risques des jeunes conducteurs sont mutualisés avec les conducteurs expérimentés considérés comme les moins risqués.

Cependant nous avons pu montré dans le chapitre précédent, que les jeunes conducteurs ne sont pas tous de mauvais risques (comme le suggère la majorité des compagnies d'assurance). D'après les résultats du chapitre précédent, il existe environ 20% des jeunes conducteurs que l'assureur pourra considérer comme des bons risques. Parmi les objectifs de l'assureur, c'est garder le plus longtemps possible les bons risque dans son portefeuille. Alors est-il intéressant pour l'assureur de ne pas appliquer une surprime aux jeunes conducteurs afin de fidéliser les bons risques parmi eux ? Ceci mettrait-il en cause la pérennité de la compagnie d'assurance ? Nous essaierons de répondre à ces questions tout au long de ce chapitre.

Ce chapitre est organisé comme suit. La seconde section présente les principales statistiques descriptives des jeunes conducteurs, d'une part et des conducteurs expérimentés, d'autre part. Dans la section suivante, la sinistralité est modélisée en termes de fréquence et de coût pour les deux sous-groupes de conducteurs. Les réponses aux questions de la surprime appliquée aux jeunes conducteurs sont données dans la section 4. Enfin, la dernière section présente les conclusions.

3.2 Les statistiques descriptives des jeunes conducteurs et des conducteurs expérimentés

Nous présentons dans ce qui suit les principaux résultats comparant les statistiques descriptives entre les jeunes conducteurs et les conducteurs expérimentés (le détail de l'analyse exploratoire se trouve en annexe 3).

L'échantillon fait apparaître, pour les jeunes conducteurs, une préférence pour les contrats tous risques par rapport au contrat d'assurance de responsabilité civile, et essentiellement les formules à faible et à très forte franchise (respectivement 37% et 31.5%). Contrairement aux conducteurs expérimentés, 47% des assurés ont souscrit exclusivement la garantie RC.

TAB. 3.1: **La répartition en pourcentage des jeunes conducteurs et des conducteurs expérimentés selon le choix du contrat d'assurance**

	RC	**Formule 1**	**Formule 2**	**Formule 3**
Jeunes conducteurs	11.40	37.16	18.89	32.55
Conducteurs expérimentés	43.14	16.44	8.71	31.72
Test du $\chi^2 = 677.14$, $Pr < .0001$				

Ceci laisse à penser qu'un jeune conducteur, n'ayant pas assez d'expérience au volant, préfère la formule "tous risques" à la formule RC pour se prémunir d'avantage contre le risque. Les conducteurs expérimentés, du fait qu'ils ont plus d'aisance au volant que les jeunes conducteurs, voient leur choix orienté vers des couvertures faibles. Nous pouvons aussi évoquer, qu'au moment de la souscription d'un contrat d'assurance, le choix d'un jeune conducteur peut être influencé par les conseils de ses parents ou de son assureur. Ce dernier préfère souvent proposer des contrats tous risques.

Notons Y la variable décrivant l'occurrence des sinistres. Elle est égale à 1 si l'assuré a déclaré au moins un accident responsable, et elle est égale à 0 sinon.

TAB. 3.2: **La répartition en pourcentage des jeunes conducteurs et des conducteurs expérimentés selon le choix du contrat d'assurance en fonction de l'occurrence des sinistres**

		RC	**Formule 1**	**Formule 2**	**Formule 3**
Jeunes conducteurs	$Y = 0$	10.98	38.97	19.05	31
	$Y = 1$	15	21.67	17.50	45.83
Test du $\chi^2 = 17.58,\ Pr = 0.0005$					
Conducteurs expérimentés	$Y = 0$	41.38	17.94	9.12	31.57
	$Y = 1$	56.21	5.32	5.65	32.82
Test du $\chi^2 = 682.32,\ Pr < .0001$					

D'après le tableau 3.2 ci dessus, la part des assurés sinistrés, qu'ils soient jeunes conducteurs ou expérimentés, ayant choisi le contrat *Formule 3* avec la plus faible franchise est plus importante que celle des assurés qui ont souscrit les autres types de contrats tous risques avec des franchises plus élevées (46% pour les jeunes et 33% pour les expérimentés). La principale différence entre les jeunes conducteurs et les conducteurs expérimentés est que le pourcentage le plus élévé des conducteurs expérimentés sinistrés est celui des assurés au tiers. Contrairement aux jeunes conducteurs, la part des assurés ayant déclaré au moins un sinistre responsable augmente avec la couverture d'assurance. Les test du χ^2 montrent l'existence d'une relation statistiquement significative entre l'occurrence des sinistres et le type du contrat d'assurance souscrit. Ceci doit être vérifiée par une analyse prenant en considération toutes les variables dont nous disposons.

Remarque 3 *Nous avons mentionné plus haut que les coûts des accidents peuvent être soit des coûts réels soit des coûts évalués. Lorsque la valeur de la variable coût réel est égale à zéro, le coût de l'accident n'est pas forcément nul. Il peut correspondre à un coût évalué par l'assureur. C'est la raison pour laquelle nous procédons au changement de la variable des coûts des sinistres qui sera définie comme suit :*

$$\textit{Le coût des sinistres} = \begin{cases} \textit{coût réel} & \textit{si le coût réel} \neq 0, \\ \textit{coût évalué} & \textit{si le coût réel} = 0 \ \textit{et le coût estimé} \neq 0, \\ 0 \ \textit{sinon.} & \end{cases}$$

3.3 La sinistralité en fonction de la couverture d'assurance et des caractéristiques observables

Dans cette section, nous analysons l'effet le plus déterminant entre les caractéristiques observables sur la sinistralité, en termes de coût et de fréquence pour les jeunes conducteurs et les conducteurs expérimentés.

3.3.1 Le nombre des sinistres

Nous avons estimé dans le chapitre précédent moyennant un modèle logit binomial l'occurrence des sinistres responsables (voir les tableaux 2.11 et 2.12). Il reste donc à savoir l'effet des variables observables sur le nombre des sinistres en utilisant des modèles de comptages, à savoir le modèle de Poisson et le modèle binomial négatif[1].

3.3.1.1 Le modèle de régression de Poisson

Le nombre de sinistres responsables observés par l'assureur peut être appréhendé comme des réalisations de variables aléatoires discrètes que nous notons $S = s_i$ associées à l'assuré i, avec $i = 1, ..., n$. On admet habituellement qu'ils peuvent être considérés comme la réalisation d'une variable aléatoire de loi de Poisson, à n composantes indépendamment distribuées, de paramètre $\Lambda = (\lambda_1, ..., \lambda_n)$, conditionnée par les variables explicatives $X_i = (x_{i1}, ..., x_{ip})$ et dont la probabilité est définie par :

$$p(s_i) = Pr(S = s_i/X_i) = e^{-\lambda_i}\frac{\lambda_i^{s_i}}{s_i!},\ avec\ ln(\lambda_i) = X_i'\beta.$$

Les estimations de la régression du modèle de poisson sont représentées dans le tableau 3.3.

[1]Les techniques d'analyse de la fréquence des sinistres font l'objet de différents travaux actuariels. Denuit et *al.* (2007) fournissent une étude approfondie sur les modèles de comptages appliqués aux sinistres automobiles. Boucher et *al.* (2008) utilisent des données de panel pour analyser les modèles classiques de comptage, à savoir le modèle de poisson et le modèle binomial négatif.

Tab. 3.3: **Ajustement du modèle de régression de Poisson pour les jeunes conducteurs et les conducteurs expérimentés**

Paramètres	**Jeunes conducteurs**		**Conducteurs expérimentés**	
Constante	-6.7270	***	-2.7725	***
	(0.8186)		(0.1138)	
Homme	0.1247		-0.1081	***
	(0.1992)		(0.0322)	
Age du conducteur	-0.0205		-0.0046	
	(0.0267)		(0.0032)	
Enseignant	0.1742		-0.0349	
	(0.1961)		(0.0314)	
Zone rurale	0.1808		0.1386	***
	(0.1903)		(0.0308)	
Ancienneté du permis	0.6042	***	0.0146	***
	(0.1657)		(0.0036)	
Conducteur principal	-0.6421	*	0.0156	
	(0.3451)		(0.0330)	
Bonus malus	0.0476	***	0.0228	***
	(0.0046)		(0.0007)	
Ancienneté du véhicule	0.0378		-0.0246	***
	(0.0241)		(0.0043)	
Puissance du véhicule	0.0033		0.0036	***
	(0.0064)		(0.0006)	
Formule 1	-1.3415	***	-1.5123	***
	(0.3625)		(0.0836)	
Formule 2	-0.7549	**	-0.6994	***
	(0.3345)		(0.0694)	
Formule 3	-0.1655		-0.3236	***
	(0.2593)		(0.0348)	

La significativité au seuil de 10%, 5% et 1% est noté respectivement par $*$, $**$ and $***$

Les valeurs entre parenthèses représentent les écarts-types des paramètres estimés

Le tableau 3.4 illustre l'analyse de *Type 3*. Cette dernière permet d'examiner la contribution de chacune des variables par rapport à un modèle ne la contenant pas.

Tab. 3.4: **Statistiques du rapport de vraisemblance pour analyse de type 3**

Variables	**Jeunes conducteurs**	**Conducteurs expérimentés**
	Pr > ChiSq	
Homme	0.5298	0.0008
Age du conducteur	0.4217	0.1405
Enseignant	0.3718	0.2668
Zone rurale	0.3400	<.0001
Ancienneté du permis	0.0001	<.0001
Conducteur principal	0.0515	0.6350
Bonus malus	<.0001	<.0001
Ancienneté du véhicule	0.1225	<.0001
Puissance du véhicule	0.6114	<.0001
Formule 1	0.0003	<.0001
Formule 2	0.0248	<.0001
Formule 3	0.5284	<.0001

Les estimations des paramètres par les régression de Poisson font ressortir de nouveaux effets, par rapport au modèle logit, contribuant à l'augmentation du nombre de sinistres déclarés à la compagnie d'assurance par les jeunes conducteurs et les conducteurs expérimentés. En effet, pour les jeunes conducteurs, le faite d'être conducteur principal ou conducteur secondaire apparaît comme variable statistiquement significative au seuil de 10%. Alors qu'elles n'ont aucun impact sur la déclaration d'au moins un sinistre responsable dans la regression logistique. Contrairement aux conducteurs expérimentés, le type du conducteur n'a acun effet significatif sur le nombre des sinistres.

Une hypothèse du modèle de poisson suggère que le nombre moyen des sinistres soit égale à la variance : $E(s_i/X_i) = V(s_i/X_i) = \lambda_i$. Ceci constitue une contrainte assez forte dans la mesure où il y a une plus forte variabilité du nombre d'accidents que le prévoit la loi de Poisson. Nous parlons donc de sur-dispersion. Ne pas prendre en considération cette sur-dispersion dans l'analyse produit des sous-estimations des écarts types des estimateurs, générant une significativité surestimée de certaines variables explicatives. Une variable jugée pertinente dans le modèle de Poisson pourrait ainsi ne plus l'être après une prise en compte de la sur-dispersion.

3.3.1.2 Le modèle binomial négatif

La principale raison d'une éventuelle sur-dispersion dans les données est que le paramètre λ_i ne varie pas seulement en fonction des variables explicatives observables, mais aussi des caractéristiques inobservables qui ne sont pas contrôlées. Par conséquent, le paramètre estimé ne représente pas seulement l'effet de la variable associée, mais il inclue aussi l'effet des facteurs cachés. Il est alors plus adéquat d'avoir :

$$s_i \approx Poi(exp\,(X_i'\beta + \epsilon_i))$$

où la variable aléatoire ϵ_i représente l'effet résiduel des caractéristiques inobservables.

En notant $\Sigma_i = exp(\epsilon_i)$,

$$E(s_i) = E(\Sigma_i)exp(X_i'\beta) = \lambda_i E(\Sigma_i)$$

et

$$V(s_i) = \lambda_i E(\Sigma_i) + \lambda_i^2 V(\Sigma_i)$$

où la variance est bien supérieure à l'espérence. La loi binomiale négative permet de tenir compte de la sur-dispersion et où les effets aléatoires Σ_i sont indépendants et de même loi Gamma de moyenne 1 et de variance $1/a$. Dans ce cas,

$$E(s_i) = \lambda_i \text{ et } V(s_i) = \lambda_i + \lambda_i^2/a$$

Le modèle de régression binomial négatif n'impose donc pas l'égalité de l'espérance et de la variance des sinistres tout en prenant en compte l'hétérogénéité résiduelle de l'échantillon des assurés. Les résultats de l'estimation de la régression binomiale négative pour les jeunes conducteurs et les conducteurs expérimentés sont illustrés dans le tableau 3.5

TAB. 3.5: **Ajustement du modèle de régression binomiale négative**

Paramètres	**Jeunes conducteurs**		**Conducteurs expérimentés**	
Constante	-8.4927	***	-3.2211	***
	(1.2145)		(0.1588)	
Homme	0.3825		-0.1204	***
	(0.2415)		(0.0384)	
Age du conducteur	0.0205		-0.0018	
	(0.0255)		(0.0038)	
Enseignant	0.0704		-0.0452	
	(0.2330)		(0.0373)	
Zone rurale	0.2167		0.1416	***
	(0.2268)		(0.0365)	
Ancienneté du permis	0.6545	***	0.0143	***
	(0.1753)		(0.0043)	
Conducteur principal	-0.8629	**	0.0334	
	(0.3961)		(0.0389)	
Bonus malus	0.0562	***	0.0272	***
	(0.0087)		(0.0012)	
Ancienneté du véhicule	0.0452		-0.0295	***
	(0.0324)		(0.0048)	
Puissance du véhicule	-0.0028		0.0046	***
	(0.0079)		(0.0008)	
Formule 1	-1.4244	***	-1.5456	***
	(0.4464)		(0.0763)	
Formule 2	-0.5604		-0.7656	***
	(0.4307)		(0.0744)	
Formule 3	0.2215		-0.3477	***
	(0.3556)		(0.0421)	
Dipersion	7.0193		8.2493	
	(1.1139)		(0.1964)	

La significativité au seuil de 10%, 5% et 1% est noté respectivement par $*$, $**$ and $***$

Les valeurs entre parenthèses représentent les écarts-types des paramètres estimés

Les estimations présentées dans la tableau ci dessus sont corrigées de la sur-dispersion. Le paramètre estimé de la dispersion pour les jeunes conducteurs (resp. les conducteurs expérimentés) est égal à 7.3. Le test du rapport de vraisemblance rejette l'hypothèse nulle de la nullité du paramètre de dispersion pour les deux catégories des conducteurs (Pour les jeunes conducteurs : $Pr \geq \chi^2 = 0.000$. Pour les conducteurs expérimentés : $Pr \geq \chi^2 = 0.000$). Le paramètre de dispersion est donc significativement différent de zero, ce qui confirme que le modèle binomial négatif est meilleur que le modèle Poisson. La figure 3.1 comparant la distribution des nombres de sinistres observés et prédits par les deux modèles illustre bien le bon ajustement binomial négatif.

Les résultats montrent bien des différences dans la significativité de certaines variables explicatives par rapport au modèle de Poisson. Pour les jeunes conducteurs, parmi les variables qui s'avèrent n'avoir aucun effet significatif sur le nombre de sinistres responsables, nous trouvons la *Formule 2* du contrat d'assurance. Cependant, les mêmes variables contribuent significativement dans les deux modèles de poisson et binomial négatif.

3.3.2 La sinistralité en termes de coût

Cette sous-section étudie la corrélation entre les coûts des accidents déclarés et les différentes variables observables.

L'analyse des coûts des sinistres est sensiblement plus compliquée que celle des fréquences. Ceci est en raison du nombre assez limité des observations qui concernent les polices sinistrées utilisées en vue d'étudier la loi des montants des sinistres. Ces derniers sont beaucoup plus difficiles à modéliser que les nombres. Certains sinistres graves nécessitent des délais assez longs pour être clôturés. Par

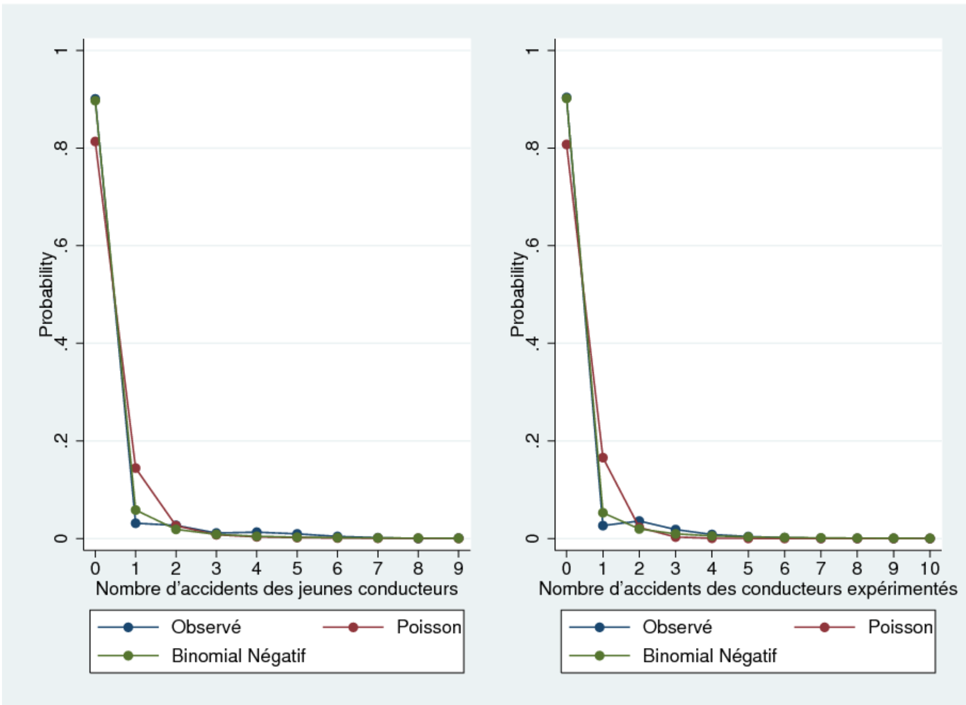


FIG. 3.1 – **Comparaison entre le nombre des accidents observés et les probabilités prédites par les modèles de Poisson et Binomial Négatif**

exemple, dans le cas d'un sinistre RC avec dégâts corporels, il faut attendre que l'état de la victime soit stabilisé avant de pouvoir fixer le montant exact des indemnités à payer. C'est la raison pour laquelle l'assureur dispose dans ses fichiers des estimations des coûts des sinistres non réglés l'année en cours.

Un point important doit être soulevé à ce niveau d'analyse. Les coûts des sinistres, comme nous l'avons montré dans la partie descriptive des données, sont très dispersés et leur distribution présente une queue très étalée à droite. Cette allure fortement asymétrique est expliquée par la détection de quelques sinistres graves dans l'échantillon (pour les jeunes conducteurs et les conducteurs expérimentés) causés par un nombre très faible d'assurés. Il importe donc d'apporter une attention particulière à ce type de sinistres graves.

Les lois continues les plus utilisées pour l'étude des coûts de sinistres en assurance non vie sont la loi Gamma et la loi Log-normale. Pinquet (1996), par exemple, dans son étude de la "prise en compte du coût des sinistres dans les systèmes bonus malus", a appliqué la loi Gamma et la loi Log-normale sur l'espérance du coût des sinistre. Cependant l'utilisation de la loi Gamma, la loi Log-normale ou n'importe quel autre modèle paramétrique procurent certes un bon ajustement global de la distribution des coûts, mais l'ajustement des queues peut être particulièrement mauvais. C'est la raison pour laquelle la théorie des valeurs extrêmes s'impose dans ce cas de figure.

Le coût total des sinistres supporté par la compagnie d'assurance est ainsi défini comme la somme des coûts standards ou normaux des sinistres et des excessifs engendrés par des sinistres graves. Plus précisément, le coût total peut s'écrire comme suit :

$$C_{i\ total} = \sum_{k=1}^{N_i^{petit}} C_{ik} + \sum_{k=0}^{N_i^{grand}} L_{ik}$$

où

N_i^{petit} est le nombre de sinistres ayant des coûts standards, déclaré par l'assuré i,

C_{ik} représente le coût standard du $k^{\text{ème}}$ sinistre déclaré par l'assuré i,

N_i^{grand} représente le nombre de sinistres graves ayant des coûts très élevés, déclaré par l'assuré i (en général, les sinistres graves déclarés par assuré sont au nombre de 0 ou de 1),

et L_{ik} est le coût du $k^{\text{ème}}$ sinistre grave déclaré par l'assuré i.

Cette représentation de la charge des sinistres exige une analyse séparée entre les coûts standards et les coûts extrêmes. Ceci n'est possible à appliquer que si un seuil optimal séparant les deux niveaux des coûts des sinistres est déterminé. Pour cela, nous appliquons une approximation de la distribution de Pareto Généralisée afin d'estimer un seuil au-delà duquel les coûts des sinistres peuvent être qualifiés d'extrêmes.

Pour un seuil u suffisamment élevé, la fonction de répartition de la loi Pareto Généralisée associée à la variable aléatoire X sachant $X > u$ est définie par :

$$G(x) = \begin{cases} 1 - \left(1 + \xi \frac{x}{\beta}\right)^{-\frac{1}{\xi}} & \text{si } \xi \neq 0 \\ 1 - exp(-\frac{x}{\beta}) & \text{si } \xi = 0 \end{cases}$$

La valeur du seuil u peut être fixée dès que l'approximation de la loi Paréto Généralisée à la queue de la distribution des coûts devient adéquate. Sans perte de généralité, nous choisissons comme seuil à partir duquel les coûts des sinistres peuvent être supposés comme extrêmes le quantile extrême estimé à 99.9%, et ce pour les jeunes conducteurs comme pour les conducteurs expérimentés.

Les figures 3.2 et 3.3 représentent la fonction de distribution cumulée des observations se situant au niveau de la queue. Elles illustrent, respectivement pour les jeunes conducteurs et pour les conducteurs expérimentés, les quantiles extrêmes estimés par l'approximation de la loi Paréto Généralisée. Ces quantiles sont estimés avec une probabilité de 0.999 d'appartenir à la queue de la distribution des coûts avec un intervalle de confiance de 95%. Le tableau 3.6 suivant présente la valeur du seuil ainsi que les bornes supérieures et inferieures de l'intervalle de confiance.

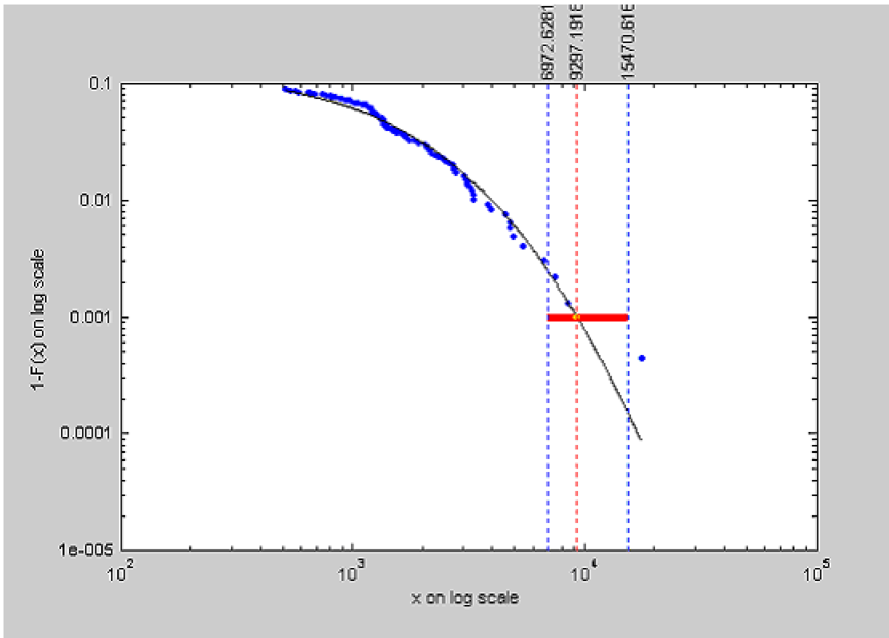


FIG. 3.2 – **Estimation du quantile extrême à 99,9% par l'approximation GPD, les jeunes conducteurs**

TAB. 3.6: **Le choix du seuil et approximation GPD**

euros	**Jeunes conducteurs**	**Conducteurs expérimentés**
Limite inférieure	6973	9243
Quantile extrême estimé	9297	9914
Limite supérieure	15471	10756

Pour les jeunes conducteurs, par exemple, le seuil à partir duquel les coûts sont considérés comme extrêmes est égale à 9297 euros.

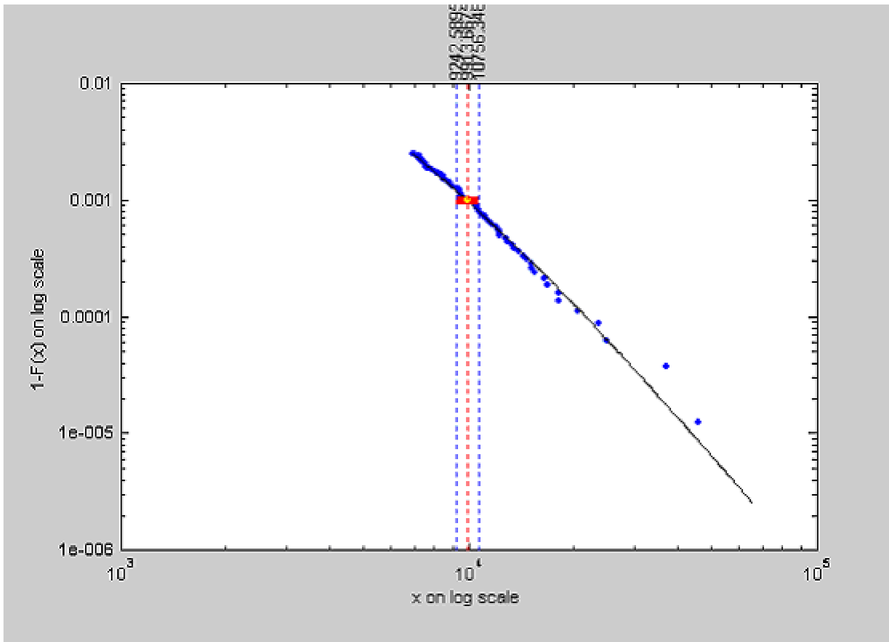


FIG. 3.3 – **Estimation du quantile extrême à 99,9% par l'approximation GPD, les conducteurs experimentés**

Après avoir choisi le seuil à partir duquel les coûts des sinistres sont supposés comme extrêmes, les coûts standards C_i sont par la suite estimés en fonction des différentes caractéristiques observables par le modèle de régression Gamma. D'autres modèles de regression peuvent être aussi appliqués comme le modèle log-normal ou la distribution gaussienne inverse.

Soit C_{ik} le coût du $k\grave{e}me$ sinistre standard causé par l'assuré i. Nous supposons que les coûts individuels de chaque sinistre $C_{i1},\ C_{i2},\ C_{i3},...$ sont indépendants et identiquement distribués, de même loi Gamma ayant la fonction de densité de probabilités suivante :

$$f(c_{ik}/\mu_i, \upsilon) = \frac{1}{\Gamma(\upsilon)} \left(\frac{\upsilon c_{ik}}{\mu_i} \right)^{\upsilon} exp \left(-\frac{\upsilon c_{ik}}{\mu_i} \right) \frac{1}{c_{ik}}$$

La moyenne et la variance sont définies respectivement comme suit :

$$\mu_i = E(C_{ik}/x_i) = exp(\beta^{'} x_i) \ \ et \ \ V(C_{ik}/x_i) = \frac{[exp(\beta^{'} x_i)]^2}{\upsilon}$$

Remarque 4 *La base de données dont nous disposons ne comprend que les coûts cumulés individuels, c'est à dire le total des coûts de tous les sinistres enregistrés par chaque assuré, et non pas le détail des coûts par sinistre déclaré. Il est plus pratique dans ce cas de travailler avec les coûts moyens C_i/s_i, où s_i est le nombre de sinistres déclarés par l'assuré i. Nous spécifions donc un poids pour chaque assuré correspondant au nombre de sinistres déclarés, en multipliant le paramètre υ par le nombre de sinistres s_i (le paramètre υ contrôle la forme de la fonction de densité de probabilités).*

Le tableau 3.7 présente les résultats de la modélisation des coûts des sinistres standards pour les jeunes conducteurs et les conducteurs expérimentés.

TAB. 3.7: **Les coûts des sinistres par la regression Gamma**

Paramètres	**Jeunes conducteurs**			**Conducteurs expérimentés**		
Constante	4.8340	(0.9497)	***	6.2826	0.1014	***
Homme	-0.1107	(0.1947)		-0.0379	0.0246	
Age du conducteur	0.0158	(0.0204)		0.0045	0.0025	*
Enseignant	0.1851	(0.1761)		-0.0750	0.0241	***
Zone rurale	0.0486	(0.3059)		0.0304	0.0363	
Ancienneté du permis	-0.2251	(0.1788)		0.0006	0.0028	
Conducteur principal	-0.5262	(0.3447)		0.0352	0.0252	
Bonus malus	0.0149	(0.0047)	***	0.0071	0.0006	***
Ancienneté du véhicule	0.0162	(0.0252)		-0.0230	0.0028	***
Puissance du véhicule	0.0048	(0.0059)		0.0003	0.0005	
Formule 1	0.0461	(0.3677)		-0.0962	0.0619	
Formule 2	-0.2686	(0.2941)		-0.0703	0.0506	
Formule 3	0.3834	(0.2279)	*	-0.0078	0.0269	
Scale	0.3051	(0.0258)		0.3730	0.0047	

La significativité au seuil de 10%, 5% et 1% est noté respectivement par $*$, $**$ and $* * *$

Les valeurs entre parenthèses représentent les écarts-types des paramètres estimés

Les résultats fournis par le tableau 3.7 montrent que l'effet du coefficient de réduction-majoration, toutes choses égales par ailleurs, contribue significativement, au seuil de 1%, dans l'alourdissement des charges des sinistres déclarés, et ce pour les jeunes conducteurs comme pour les conducteurs expérimentés. Cet effet posistif du bonus malus sur les coûts des sinistres va dans le même sens du résultat trouvé à partir de la regression des fréquences des sinistres. Ceci se traduit par le fait que plus le nombre d'accidents déclarés dans le passé est élevé, plus l'étendu et la fréquence des dommages qui surviennent par la suite sont importants. En d'autres termes, le mauvais conducteur, jeune ou expérimenté, reste mauvais conducteur. Nous confirmons nos résultats du premier chapitre, à savoir la sinistralité est une caractéristique du risque intrinsèque. Ceci va dans le même sens des résultats de Dahchour et Dionne (2002) qui ont trouvé une corrélation positive entre le coefficient du bonus-malus et le nombre d'accident.

Pour les jeunes conducteurs, contrairement aux conducteurs expérimentés, la couverture d'assurance joue un rôle statistiquement significatif, au seuil de 10%, dans l'ampleur des dégats causés par l'assuré. Plus précisément, le contrat d'assurance proposant la franchise la plus faible, relativement au contrat *RC*, contribue à l'augementation des coûts des sinistres, donc à l'aggravation des accidents. Ici, nous pouvons detecter une certaine corrélation positive entre le risque et la couverture d'assurance, mais sans pour autant confirmer l'hypothèse de la présence d'asymétrie d'information ; le risque d'erreur de rejeter l'abscence de corrélation entre la *Formule 4* et les coûts des sinistres n'est pas négligeable (10%) et les paramètres estimés des autres couvertures d'assurance ne sont pas significatifs.

En ce qui concerne les autres variables explicatives des coûts de sinistres, l'ancienneté du véhicule, le fait d'être enseignant et l'age du conducteur contribuent significativement dans la variation des coûts des sinistres, seulement pour les conducteurs expérimentés. Les coûts des sinistres diminuent avec le fait d'être enseignant, contrairement aux estimations de la fréquence des sinistres où cette variable n'a aucun effet significatif. Cependant, toutes les caractéristiques propres au jeune conducteur n'ont aucun effet significatif sur la gravité des sinistres.

3.4 La surprime est-elle indispensable pour les jeunes conducteurs ?

Nous sommes posées plusieurs interrogations suite à ces résultats : nous avons montrer qu'il existe une proportion non négligeable de jeunes conducteurs, qui peut être considérée des bons risques. Nous avons confirmé que la sinisralité est liée au risque intrinsèque de l'assurées, alors les bons risques resteront toujours des bons risques. Il vient ainsi de savoir s'il est intéressant pour l'assureur de ne pas appliquer une surprime aux jeunes conducteurs afin de fidéliser les bons risques parmi eux et si ceci mettrait-il en cause la pérennité de la compagnie d'assurance.

Afin de répondre à ces questions, nous nous sommes appuyés sur la démarche décrite par Grun-Rehomme et al. (2009) qui permet de comparer les enjeux financiers pour l'assureur dans le cas où il applique la surprime aux jeunes conducteurs et dans l'autre cas où il vise la fidélisation de sa nouvelle clientèle en leur proposant le même prix d'assurance que les conducteurs expérimentés. Le détail de leur méthodologie se présente comme suit.

Un jeune conducteur devient un conducteur expérimenté, au bout de deux années de contrat d'assurance, après lesquelles il peut résilier son contrat et choisir une autre compagnie d'assurance.

Notons [2] :

- n_1 : le nombre de jeunes conducteurs souscrivant une assurance, indépendamment de la suppression de la surprime. Ce nombre peut diminuer au cours du temps, précisément après deux années d'assurance.
- a : le taux d'entrée supplémentaire de jeunes conducteurs attirés par la suppression de la surprime.
- b : le taux de départ la $3^{ème}$ et $4^{ème}$ année parmi les a assurés attirés par l'absence de surprime.
- p_1 et p_2 : les primes pures pour les jeunes conducteurs et les conducteurs expérimentés, respectivement.

[2]Nous utilisons les mêmes notations que Grun-Rehomme et al. (2009)

– C_1 et C_2 : les primes payées par les jeunes conducteurs et les conducteurs expérimentés, respectivement. Nous supposons que le calcul de ces primes est basé sur la prime pure en tenant compte d'un taux de chargement β identique pour tous les assurés. Plus précisément, $C_i = p_i + \beta C_i$.

La prime pure est calculée après avoir effectué une classification des risques basée sur les caractéristiques observables et disponibles à l'assureur. Une définition plus détaillée de la prime pure est donnée dans la sous-section [4.4.7.1]. Tout au long de la description de la démarche suivante, p_i $(i = 1, 2)$ est en réalité p_{ik}, où k correspond à la classe de risque qui y est associée (k est supprimé pour simplifier la lecture).

Afin de répondre aux questions que nous nous sommes posées plus haut, Il faut savoir au bout de combien d'années les comptes de l'assureur reviennent à l'équilibre, si la compagnie d'assurance n'impose pas de surprime aux jeunes conducteurs et que ces derniers payent ainsi la même prime que les conducteurs expérimentés pendant les deux premières années d'assurance.

a/ Le gain de l'assureur pendant la première et la deuxième année après la souscription de nouveaux contrats d'assurance

Nous supposons que pendant les deux premières années C_1 ne change pas. Alors, la différence entre les primes payées sur les deux premières années ($t = 1,\ 2$), en rajoutant l'entrée supplémentaire des jeunes conducteurs attirés par la suppression de la surprime, est :

$$\Delta p_t = n_1(1 + a)C_2 - n_1 C_1$$

La différence de coût de sinistralité pour l'assureur sur les deux premières années est donc égale à :

$$\Delta cout_t = n_1(1+a)p_1 - n_1p_1 = n_1ap_1$$

La fonction de gain pour les deux premières années est :

$$\Delta G_t = \Delta p - \Delta cout = n_1(1+a)C_2 - n_1C_1 - n_1ap_1$$

b/ Le gain de l'assureur pendant la troisième et la quatrième année après la souscription de nouveaux contrats d'assurance

Nous supposons que, pendant la $3^{ème}$ et la $4^{ème}$ année, un taux b d'assurés, parmi ceux qui ont été attiré par la suppression de la surprime pendant la première et la deuxième année, résilient leur contrat d'assurance et quittent la compagnie d'assurance. Par la suite, le nombre de départ est $b * a * n_1$.

La différence de prime payée la $3^{ème}$ année est :

$$\Delta p_3 = n_1a(1-b)C_2$$

Et la $4^{ème}$ année est :

$$\Delta p_4 = n_1a(1-b)^2C_2$$

La différence de coût de sinistralité, respectivement pour la $3^{ème}$ et la $4^{ème}$ années est égale donc à :

$$\Delta cout_3 = n_1a(1-b)p_2$$

et

$$\Delta cout_4 = n_1 a(1-b)^2 p_2$$

Les fonctions de gain pour la $3^{ème}$ et la $4^{ème}$ années sont :

$$\Delta G_3 = \Delta p_3 - \Delta cout_3 = n_1 \beta' a(1-b) p_2$$

et

$$\Delta G_4 = \Delta p_4 - \Delta cout_4 = n_1 \beta' a(1-b)^2 p_2$$

Nous supposons que pour les années qui suivent il n'y a plus de départ supplémentaire et que la fonction de gain de l'assureur ne change pas.

Notons ρ le taux d'actualisation. L'équation d'équilibre à l'horizon h ($h > 3$), pour laquelle la somme des gains de l'assureur est nul, est définie comme suit :

$$\Delta G_1(1+\rho) + \Delta G_3 \rho^2 + \Delta G_4 \sum_{k=3}^{h-1} \rho^k = 0$$

Cette équation est indépendante du nombre n_1 des jeunes conducteurs assurés indépendamment de la suppression de la surprime.

L'objectif est donc de chercher la plus petite valeur de h pour laquelle les gains de l'assureur doivent être positifs et donc l'équation d'équilibre est supérieure à zéro.

$\Delta G_1(1+\rho) + \Delta G_3 \rho^2 + \Delta G_4 \sum_{k=3}^{h-1} \rho^k > 0$

$\Rightarrow \Delta G_1(1+\rho) + \Delta G_3 \rho^2 + \Delta G_4 \rho^3 \left(\frac{1-\rho^{h-3}}{1-\rho}\right) > 0$

$\Rightarrow h = 3 + \frac{1}{ln\rho} ln\left[(1-\rho)\frac{\Delta G_1(1+\rho)+\Delta G_3\rho^2}{\Delta G_4 \rho^3} + 1\right]$

Ce qui équivaut, en fonction de a, b, $\beta^{'}$, ρ et des primes pures, à :

$$h = 3 + \frac{1}{ln\rho} ln \left[1 + (1-\rho)^{\frac{(1+\rho)\left[(1+a)(1+\beta^{'})p_2 - (1+a+\beta^{'})p_1\right] + \left[a\beta^{'}(1-b)p_2\right]}{\rho^3 \beta^{'} a(1-b)^2 p_2}} \right]$$

La formule de l'horizon h peut être écrit en fonction du rapport entre la prime pure des jeunes conducteurs et celle des conducteurs expérimentés. Nous notons $\lambda = \frac{p_1}{p_2}$. L'expression de h devient comme suit :

$$h = 3 + \frac{1}{ln\rho} ln \left[1 + (1-\rho)^{\frac{(1+\rho)\left[(1+a)(1+\beta^{'}) - (1+a+\beta^{'})\lambda\right] + \left[a\beta^{'}(1-b)\right]}{\rho^3 \beta^{'} a(1-b)^2}} \right]$$

Nous fixons les paramètres $\beta^{'}$ et ρ conformément aux règles et pratiques du marché d el'assurance en France. Le taux de chargement peut, en partie, varier d'une compagnie à une autre selon la stratégie commerciale envisagée. Grun-Rehomme et al. (2009) utilisent un taux de chargement égal à 0.3, soit $\beta^{'} = 0.42857$. Nous fixons le taux d'intêtrét réel à 2%, soit $\rho = 98\%$.

Pour un taux de départ et un taux d'entrée fixés, l'horizon h est une fonction croissante et convexe de λ. L'horizon d'équilibre augmente donc avec le différentiel de la prime de référence entre les jeunes conducteurs et les conducteurs expérimentés. Il est croissant en fonction de niveau de la surprime appliquée au jeune conducteurs. Nous analysons l'effet du taux d'entrée a et du taux de départ b après le calcul du différentiel de la prime entre les jeunes et les expérimentés. C'est ce que nous effectuons à la sous-section suivante.

3.4.1 Calcul de la prime pure

La prime pure correspond à l'espérance mathématiques des coûts des sinistres déclarés tout au long de la période de couverture (bien évidemment, ceci ne tient compte que sous l'hypothèse de *la loi des grands nombres*, c'est-à-dire avec un large portefeuille contenant des risques indépendant et identiquement distribués).

La tarification actuarielle est basée sur la fréquence des sinistres ainsi que leurs coûts. Plus précisément, soit :

- S_i une variable aléatoire de comptage correspondant au nombre des sinistres déclarés par l'assuré i durant la période de couverture [3],
- et c_{ji} une variable aléatoire continue représentant le montant du $j^{\text{ème}}$ sinistre déclaré par l'assuré i, avec $j = 1, 2, 3...S_i$.

La charge totale des sinistres C_i associée à l'assuré i suit une loi composée. Elle est de la forme :

$$C_i = \sum_{j=1}^{S_i} c_{ij}$$

La prime pure est alors égale à[4]

$$\begin{aligned} E(C_i) &= E\left(\sum_{j=1}^{S_i} c_{ij}\right) \\ &= \sum_{k=1}^{+\infty} Pr(S_i = k) E\left(\sum_{j=1}^{k} c_{ij}\right) \\ &= \left(\sum_{k=1}^{+\infty} Pr(S_i = k) k\right) E(c_{i1}) \\ &= E(S_i) E(c_{i1}) \end{aligned}$$

Ceci n'est valable que sous les hypothèses d'indépendance entre les coûts et le nombre des sinistres, et l'indépendance et l'identique distribution des coûts.

[3]Dans notre échantillon, la période de couverture de chaque assuré est d'un an. Il n'y a eu ni résiliation de contrat, ni souscription de nouveaux contrats pendant l'année.

[4]Pour plus de détail sur le calcul de la prime pure, voir Charpentier, A. et Denuit, M. (2004) "*Mathématiques de l'assurance non-vie, Tome 1 : principes fondamentaux de théorie du risque*", pages 105-180.

L'assureur fait face à des risques hétérogènes se côtoyant au sein du portefeuille de l'assurance. Il soulève en partie ce problème d'hétérogénéité en se basant sur les différentes caractéristiques observables et disponibles des assurés, afin de constituer plusieurs classes de risques homogènes.

La compagnie d'assurance fixe donc le montant de la prime pure dans chaque classe de risque en se basant sur les caractéristiques observables connues *a priori*, c'est la tarification *a priori*. Ce montant est ensuite personnalisé pour chaque assuré en prenant en considération l'historique individuel des sinistres. Des corrections individuelles *a posteriori* sont ainsi appliquées. Ceci permet de tenir compte de l'écart entre le risque individuel et le risque collectif de la classe. La composante individuelle constituant l'élément de correction *a posteriori* est le coefficient de réduction majoration.

Ce que nous venons d'évoquer nous mène à donner l'expression exacte de la prime pure. Elle n'est plus $E(C)$, mais $E(C/X)$, c'est-à-dire l'espérance mathématique des coûts conditionnellement aux différentes informations disponibles à l'assureur pour chaque classe de risque constituée. Une remarque importante doit être soulevée à ce niveau.

Remarque 5 *Comme nous l'avons évoqué plus haut, le calcul de la prime pure se base sur l'espérance mathématique sous l'hypothèse de la loi des grands nombres, c'est-à-dire un large effectif avec des risques indépendants et identiquement distribués. L'assureur, en regroupant les assurés présentant les mêmes caractéristiques en différentes classes de risque, peut faire face à des classes avec de très faibles effectifs. Ainsi, l'une des hypothèses de la loi des grands nombres n'est plus satisfaite. L'assureur, afin de déterminer les tarifs, a donc recours aux modèles de régressions fonctions des différentes caractéristiques.*

Cette remarque nous conduit à définir les lois conditionnelles que nous appliquons par la suite aux nombres de sinistres et à leurs coûts. Comme il a été appliqué dans les sections précédentes, nous supposons que les coûts suivent une loi Gamma, et nous choisissons la distribution binomiale négative pour le nombre des sinistres.

Ce qui est différent par rapport aux régressions précédentes est que, dans les estimations suivantes, les sinistres ainsi que leurs coûts ne dépendent que des variables de tarification *a priori* de l'assureur. En d'autres termes, afin de déterminer la prime pure de l'assuré dans chaque classe de risque, les variables indépendantes expliquant la fréquence des sinistres ainsi que les coûts des sinistres ne contiennent pas la variable du bonus malus. Les modélisations des coûts et du nombre des sinistres sont appliquées sur la totalité de l'échantillon. Nous comparons par la suite les tarifs calculés dans les différentes classes de risques entre les jeunes conducteurs et les conducteurs expérimentés.

Afin de faire ressortir les différentes classes de risque, nous catégorisons les variables continues avant d'appliquer les régressions. Nous avons appliqué la même catégorisation des variables continues que la compagnie qui nous a fourni les données.

- Groupe d'Ancienneté du véhicule en années : $[0, 2]$, $[3, 5]$, $[6, 10]$ et > 10.
- Groupe de Puissance du véhicule en chevaux DIN : ≤ 65, $[66, 85]$, $[86, 110]$ et > 110.
- Groupe d'Age du conducteur en années : $[18, 25]$, $[26, 35]$, $[36, 55]$ et > 55.

Pour faire ressortir les deux catégories des jeunes conducteurs et des conducteurs expérimentés, nous transformons l'ancienneté du permis de conduire en une variable catégorielle égale à : *novice* si l'ancienneté du permis < 3, *expérimenté* si l'ancienneté du permis ≥ 3.

3.4.1.1 Les coûts des sinistres

Nous utilisons la même méthode de détection des valeurs extrêmes que dans la section précédente [4.4.6] sur tout l'échantillon, à savoir la méthode par approximation de la loi Pareto Généralisée. Le tableau 3.8 ainsi que la figure 3.4 illustrent le quantile extrême estimé à 99.9%.

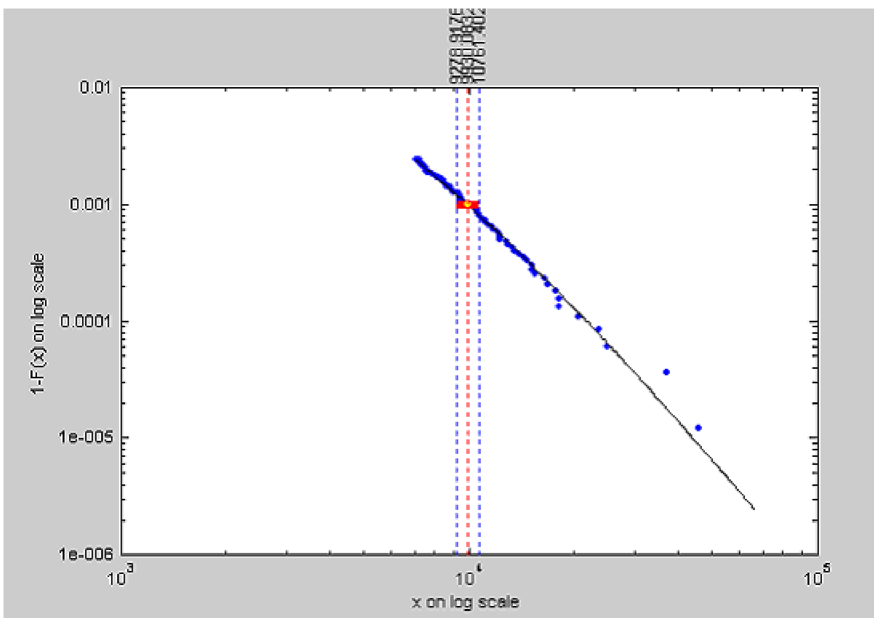


FIG. 3.4 – **Estimation du quantile extrême à 99,9% par l'approximation GPD, l'échantillon total**

TAB. 3.8: **Le choix du seuil par approximation GPD**

	en euros
Limite inférieure	9279
Quantile extrême estimé	9930
Limite supérieure	10761

Ainsi, nous écartons les polices ayant des coûts supérieurs à 9930 *euros*, qui sont au nombre de 40 polices. Nous ajustons ensuite la loi Gamma sur les coûts standards des sinistres en introduisant un poids s_i correspondant au nombre de sinistres.

La moyenne est donc égale à :

$$\mu_i = E(c_{ij}/x_i) = exp(\beta_0^{coût} + \sum_{k=1}^{p} \beta_k^{coût} x_{ik})$$

avec p le nombre de variables explicatives.

La variance est égale à :

$$V(c_{ij}/x_i) = \frac{\mu_i^2}{\upsilon s_i}$$

L'analyse du Type 3 fait ressortir les variables contribuant significativement dans le modèle, au seuil de 5%. Ces variables sont présentées dans le tableau 3.9. Les résultats de la régression Gamma sont illustrés dans le tableau 3.10.

TAB. 3.9: **Statistiques du rapport de vraisemblance pour l'analyse de Type 3**

Variables	DF	Khi2	Pr>Khi2
Type du conducteur	1	6.91	0.0086
Profession	1	12.51	0.0004
Age du conducteur	3	41.17	<.0001
Ancienneté du véhicule	3	49.03	<.0001
Type du contrat	1	27.75	<.0001

Nous commençons par exclure les variables les moins significatives et seules les variables significativement différentes de zéro au seuil de 5% sont présentées, à savoir le type du conducteur, la profession, l'âge du conducteur, l'ancienneté du véhicule et le type du contrat.

TAB. 3.10: **Résultats de la regression Gamma sur les coûts des sinistres, le modèle final**

Paramètre		Estimation	Std Error	Khi2	Pr>Khi2
Constante		7.3528[a]	0.0373	38910.4	<.0001
Type du conducteur	Principal	0.0542	0.0205	6.96	0.0083
Type du conducteur	Secondaire	0.0000	0.0000	.	.
Profession	Enseignant	-0.0678	0.0192	12.47	0.0004
Profession	Autres	0.0000	0.0000	.	.
Age du conducteur	18 - 25	0.2149	0.0395	29.65	<.0001
Age du conducteur	26 - 35	0.0366	0.0303	1.45	0.2281
Age du conducteur	36 - 55	-0.0170	0.0224	0.57	0.4486
Age du conducteur	> 55	0.0000	0.0000	.	.
Ancienneté du véhicule	0 - 2	0.1960	0.0340	33.31	<.0001
Ancienneté du véhicule	3 - 5	0.1582	0.0323	23.95	<.0001
Ancienneté du véhicule	6 - 10	0.0644	0.0317	4.12	0.0423
Ancienneté du véhicule	>10	0.0000	0.0000	.	.
Type du contrat	RC	-0.1053	0.0200	27.73	<.0001
Type du contrat	Tous risques	0.0000	0.0000	.	.
Scale		0.0195	1.0007		

a : $\beta_0^{\text{coût}} = 7.3528$

3.4.1.2 La fréquence des accidents

Nous avons montré dans la section [4.4.3] que l'ajustement du modèle binomial négatif sur le nombre des accidents est plus adéquat que le modèle de Poisson. Nous choisissons donc d'effectuer la regression négative binomiale sur la fréquence des sinistres en fonction de toutes les variables qui nous sont disponibles. Le coefficient du bonus malus en est exclu. Il constitue une variable de tarification *a posteriori*.

Pour tout assuré i, la moyenne du nombres de sinistres standards suivant la loi binomiale négative est égale à :

$$\mu_i = E(S_i/x_i) = exp(\beta_0^{freq} + \sum_{k=1}^{p} \beta_k^{freq} x_{ik})$$

De la même manière que les coûts des sinistres, nous présentons les résultats de l'analyse de Type 3, dans le tableau 3.11 faisant ressortir les variables à contribution significative au seuil de 5%. Les résultats de la régression sont exposés dans le tableau 3.12.

TAB. 3.11: **Statistiques du rapport de vraisemblance pour l'analyse de Type 3**

Variables	DF	Khi2	Pr>Khi2
Sexe	1	7.02	0.0081
Age du conducteur	3	13.89	0.0031
Ancienneté du permis	1	6.11	0.0134
Ancienneté du véhicule	3	247.37	<.0001
Puissance du véhicule	3	33.70	<.0001
Type du contrat	1	119.27	<.0001

Nous remarquons que les variables qui influencent significativement, au seuil de 5%, la fréquence de sinistres sont le sexe et l'age du conducteur, l'ancienneté et la puissance du véhicule, le type du contrat ainsi que l'ancienneté du permis de conduire.

TAB. 3.12: **Résultats de la regression négative binomiale sur la fréquence des sinistres, le modèle final**

Paramètre		**Estimation**	**Std Error**	**Khi2**	**Pr>Khi2**
Constante		-2.1184^a	0.0782	734.66	<.0001
Sexe	homme	-0.1016	0.0383	7.02	0.0080
Sexe	femme	0.0000	0.0000	.	.
Age du conducteur	18 - 25	0.2964	0.0949	9.75	0.0018
Age du conducteur	26 - 35	0.0581	0.0585	0.99	0.3202
Age du conducteur	35 - 55	0.1199	0.0444	7.31	0.0069
Age du conducteur	> 55	0.0000	0.0000	.	.
Ancienneté du permis	<3	0.2900	0.1178	6.06	0.0138
Ancienneté du permis	≥ 3	0.0000	0.0000	.	.
Ancienneté du véhicule	0 - 2	0.8294	0.0625	176.32	<.0001
Ancienneté du véhicule	3 - 5	0.8161	0.0569	205.44	<.0001
Ancienneté du véhicule	6 - 10	0.6962	0.0529	173.11	<.0001
Ancienneté du véhicule	>10	0.0000	0.0000	.	.
Puissance du véhicule	≤ 65	-0.2367	0.0620	14.57	0.0001
Puissance du véhicule	66 - 85	-0.1538	0.0632	5.93	0.0149
Puissance du véhicule	86 - 110	0.0265	0.0580	0.21	0.6478
Puissance du véhicule	>110	0.0000	0.0000	.	.
Type du contrat	RC	0.4335	0.0396	119.55	<.0001
Type du contrat	Tous risques	0.0000	0.0000	.	.
Scale		0.2200	9.2186		

$a : \beta_0^{freq} = -2.1184$

3.4.1.3 Les tarifs de la prime pure

En se basant seulement sur les coûts standards, la formule de la prime pure est :

$$E(C_i) = E(S_i)E(c_{i1}) = exp\left(\beta_0^{freq} + \beta_0^{\text{coût}} + \sum_{k=1}^{p}\left(\beta_k^{freq} + \beta_k^{\text{coût}}\right)x_{ik}\right)$$

A partir de cette formule, le tarif de la prime pure de la classe de référence est égal à :

$$exp\left(\beta_0^{freq} + \beta_0^{\text{coût}}\right) = exp(-2.1184 + 7.3528) = 187.62 euros$$

Ce montant correspond à la prime pure d'une femme conductrice expérimentée, âgée de plus de 55 ans, ne travaillant pas dans le secteur de l'enseignement, assurée en tant que conductrice secondaire pour un véhicule ayant plus de 10 ans d'ancienneté avec une puissance supérieure à 110 chevaux DIN et choisissant un contrat tous risques plutôt qu'un contrat RC.

Les estimations des coefficients des différentes variables nous permettent de corriger ce montant selon l'appartenance de l'assuré aux différentes classes de risques. Les pourcentages de corrections se trouvent dans le tableau 3.13

TAB. 3.13: **Influence, en pourcentage, des différentes variables sur la prime pure**

Variable		Influence du nombre des sinistres	Influence du coût	Influence totale
Sexe	homme	89,84	100	89,84
Sexe	femme	100	100	100,00
Type du conducteur	Principal	100	105,42	105,42
Type du conducteur	Secondaire	100	100	100,00
Profession	Enseignant	100	93,22	93,22
Profession	Autres	100	100	100,00
Age du conducteur	18 - 25	129,64	121,49	157,50
Age du conducteur	26 - 35	105,81	103,66	109,68
Age du conducteur	36 - 55	111,99	98,3	110,09
Age du conducteur	> 55	100	100	100,00
Ancienneté du véhicule	0 - 2	182,94	119,6	218,80
Ancienneté du véhicule	3 - 5	181,61	115,82	210,34
Ancienneté du véhicule	6 - 10	169,62	106,44	180,54
Ancienneté du véhicule	>10	100	100	100,00
Puissance du véhicule	<65	76,33	100	76,33
Puissance du véhicule	66 - 85	84,62	100	84,62
Puissance du véhicule	86 - 110	102,65	100	102,65
Puissance du véhicule	>110	100	100	100,00
Ancienneté du permis	<3	129	100	129,00
Ancienneté du permis	≥ 3	100	100	100,00
Type du contrat	RC	143,35	89,47	128,26
Type du contrat	Tous risques	100	100	100,00

Ce tableau nous permet, par exemple, de calculer la prime pure d'un enseignant, ayant entre 18 et 25 ans, déclaré en tant que conducteur principal expérimenté, assuré pour un véhicule très neuf et très puissant et optant pour un contrat RC. Cette prime est égale à :

187.62 *euros* $\times$ 89.84% (la correction pour le fait d'être un homme)
$\times$ 105% (la correction pour pour le fait d'être conducteur principal)
$\times$ 93.22% (la correction pour pour le fait d'être enseignant)
$\times$ 157.5% (la correction pour avoir entre 18 et 25 ans)
$\times$ $218,8\%$ (la correction pour avoir un véhicul d'une ancienneté $\leq$ 2 ans)
$\times$ 100% (la correction pour avoir un véhicul d'une puissance $>$ 110 $CV.DIN.$)
$\times$ 128.26% (la correction pour avoir opté pour un contrat RC)
$\times$ 100% (la correction pour être un conducteur principal)
$=$ 726.56 *euros*

De la même manière, nous pouvons calculer les différents tarifs pour les différentes classes de risques définies par toutes les variables significatives disponibles à l'assureur.

En prenant en considération les coûts extrêmes, la prime pure s'élève alors à :

$$exp\left((\beta^{freq} + \beta^{\text{coût}})' x_i\right) + p_i E(L)$$

avec p_i est la probabilité que l'assuré i déclare au moins un sinistre grave et $E(L)$ est l'espérance mathématique des coûts extrêmes.

Nous ne procédons pas au calcul de la prime pure finale. Ce qui nous intéresse le plus est l'influence de l'ancienneté du permis de conduire que nous pouvons déterminer directement du tableau 3.13. En effet l'objectif de cette section : est-il indispensable de faire payer aux jeunes conducteurs une surprime par rapport aux conducteurs expérimentés ? Dans le cas où l'assureur préfère la fidélisation de sa nouvelle jeune clientèle, c'est-à-dire dans le cas où l'assureur n'exige pas de surprime aux conducteurs novices et propose le même tarif que les conducteurs expérimentés, au bout de combien d'années cet assureur équilibrera ses comptes ?

Nous rappelons l'expression de l'horizon h :

$$h = 3 + \frac{1}{ln\rho} ln \left[1 + (1-\rho) \frac{(1+\rho)\left[(1+a)(1+\beta')-(1+a+\beta')\lambda\right] + \left[a\beta'(1-b)\right]}{\rho^3 \beta' a(1-b)^2} \right]$$

Avec $\lambda = \frac{p_1}{p_2}$, le rapport entre la prime pure des jeunes conducteurs et celle des conducteurs expérimentés.

La valeur de lambda est donnée par le tableau 3.13. $\lambda = 1.29$. En général, « le différentiel du montant des sinistres est compris entre 10% et 30%, voire jusqu'à 40% »[5]

En fonction des paramètres d'actualisation ρ et du taux de chargement β que nous avons fixés conformément aux pratiques du marché de l'assurance automobile en France, nous calculons l'horizon h selon les variations du taux d'entrée et du taux de départ à la concurrence. Le taux d'entrée est celui des jeunes conducteurs attirés par la suppression de la surprime. Le tableau 3.14 présente les différentes valeurs de l'horizon h en fonction des variations des taux d'entrée a et des taux de départ b (nous avons supposé les mêmes valeurs de a et de b que Grun-Rehomme et *al.*).

[5] Grun-Rehomme, M. et *al.*(2009), *Les jeunes conducteurs : surprimes ou fidélisation ?*, page 119.

TAB. 3.14: **Valeur de h**

(a,b)	$(0.2, 0.2)$	$(0.2, 0.5)$	$(0.2, 0.8)$
h	11.5	31	infini *
(a,b)	$(0.3, 0.2)$	$(0.3, 0.5)$	$(0.3, 0.8)$
h	7.3	16.6	infini
(a,b)	$(0.4, 0.2)$	$(0.4, 0.5)$	$(0.4, 0.8)$
h	5.3	10.7	infini
(a,b)	$(0.5, 0.2)$	$(0.5, 0.5)$	$(0.5, 0.8)$
h	4.2	7.5	63.4
(a,b)	$(0.6, 0.2)$	$(0.6, 0.5)$	$(0.6, 0.8)$
h	3.5	5.5	34
(a,b)	$(0.8, 0.2)$	$(0.8, 0.5)$	$(0.8, 0.8)$
h	2.5	3	12
(a,b)	$(1, 0.2)$	$(1, 0.5)$	$(1, 0.8)$
h	1.5	1.6	2.3

* L'infini indique que le nombre d'année dépasse les 150 ans

La figure 3.5 montre l'évolution de l'horizon h en fonction des taux d'entrée et de départ des jeunes conducteurs.

Ces résultats appellent aux commentaires suivants :

- Pour un taux de départ fixé, la durée de retour à l'équilibre diminue avec le taux d'entrée. Par exemple, pour un taux de départ fixé à 20%, l'horizon passe de 11.5 années pour un taux d'entrée égal à 20%, à deux ans et demi pour un taux d'entrée supplémentaire de 80% de jeunes conducteurs attirés par la suppression de la surprime.

- Pour un taux d'entrée fixé, l'horizon augmente avec le taux de départ. L'augmentation est très rapide dès que le taux d'entrée est faible. Nous remarquons par exemple, que pour un taux d'entrée égale à 40%, l'horizon passe de 5.3 années à l'infini si le taux de départ varie entre 20% et 80%. À 30% de taux d'entrée et 80% de taux départ, l'assureur met plus que 150 ans pour que ses comptes retournent à l'équilibre !

- Nous pouvons remarquer que la durée du retour à l'équilibre des comptes de l'assureur semble supportable dès que le taux d'entrées supplémentaires des jeunes assurés attirés par la suppression de la surprime dépasse les 50%. Pour des taux de départ supérieurs à 60%, le retour à l'équilibre est très difficile, voire même impossible. Ainsi, pour répondre à nos intérrogations de départ, la pérennité de la compagnie d'assurance peut ne pas être touchée à condition que le taux d'entrée des jeunes conducteurs attirés par l'offre de la suppréssion de la surprime dépasse un certain seuil. Il faut bien évidemment avoir connaissance de seuil en dessous duquel il ne serait pas conseillé de supprimer la surprime, et surtout si cela est accompagné par une forte vague de départ. Il faudrait alors faire une étude de marché pertinente avant d'appliquer une telle démarche.

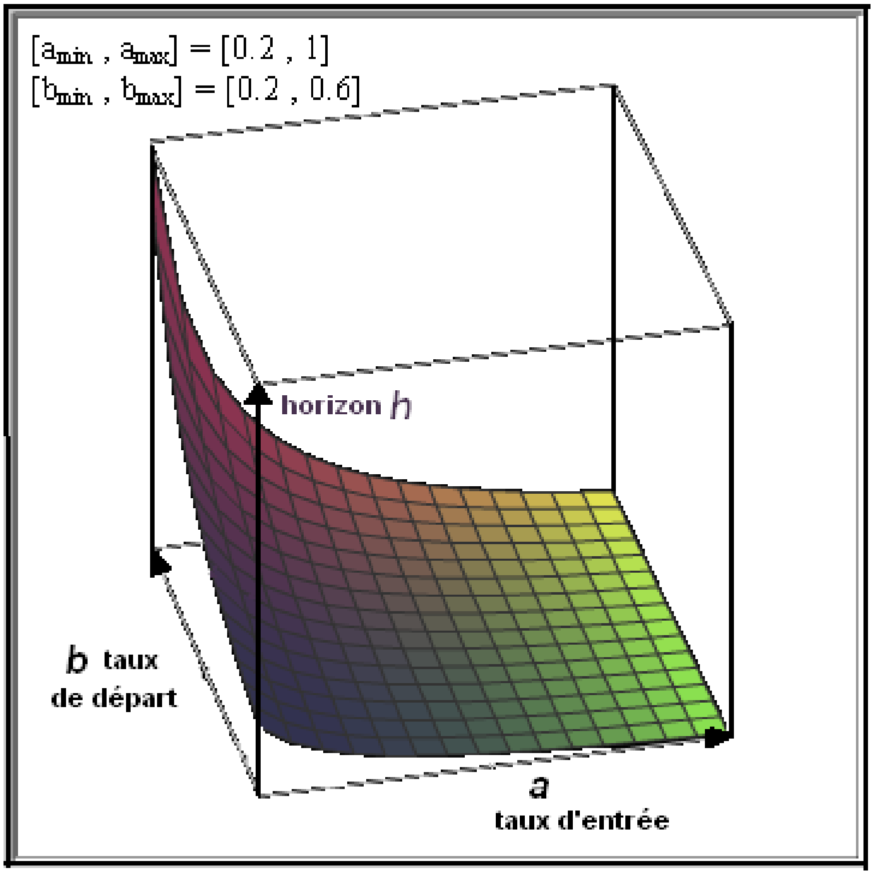


FIG. 3.5 – **Durée h du retour à l'équilibre en fonction de a et b**

Ce graphique est conditionnée aux valeurs suivantes : $[a_{min}, a_{max}] = [20\%, 100\%]$ et de $[b_{min}, b_{max}] = [20\%, 60\%]$ (Pour des taux de départs $b > 60\%$, l'horizon h est égal à l'infini). Ce qui donne des durées de départs h entre : $[h_{min}, h_{max}] = [1.522, 60.72]$.

3.5 Conclusion

Dans ce chapitre, nous avons utilisé plusieurs modèles économétriques afin d'estimer la sinistralité en termes de fréquence et de coûts. Ces modélisations ont révélé d'importantes informations sur le lien existant entre les variables disponibles à l'assureur et le risque de l'assuré exprimé en termes de sinistralité (fréquence et coûts). Nous avons pu détecté les différents pouvoirs prédictifs que ces variables présentent sur la sinistralité. Ceci nous a permis de calculer la prime pure pour les différentes classes de risques, la valeur de la prime étant une variable confidentielle que la compagnie d'assurance ne nous a pas fournie.

Grâce au calcul de la prime pure que nous avons pu répondre à la question de la surprime chez les jeunes conducteurs : Est-il indispensable d'appliquer une surprime aux conducteurs novices contrairement au conducteurs expérimentés ? Notre réponse est non. Nous avons montré qu'il est intéressant pour l'assureur d'opter pour la fidélisation d'une clientèle potentielle en éliminant la surprime, à condition que la part des jeunes conducteurs attirés au début de période dépasse un certain seuil. Estimer le taux d'entrée des jeunes conducteurs sur le marché d'assurance automobile est donc indispensable avant d'appliquer la suppression de la surprime.

L'éventualité d'avoir un taux d'entrée important n'est, à notre avis, pas improbable. Une surprime constitue pour les jeunes conducteurs un obstacle financier, les décourageant à contracter une assurance et par la suite se priver d'utiliser leur véhicule ("La conduite et la possession d'un véhicule permettent aux jeunes d'être mobiles, de participer aux activités économiques et éducatives..." (Grun-Rehomme et *al.* (2009)). Bien évidemment si la suppression de la surprime est une idée intéressante à prendre en considération, elle doit néanmoins être accompagnée d'une compagne de sécurité routière destinée aux jeunes conducteurs (ce qui est le cas dans plusieurs compagnies d'assurance), ou bien d'une formation.

Conclusion générale

Ce travail s'inscrit dans un cadre théorique et empirique de l'étude des problèmes informationnels entre assureur et assuré, l'analyse de la sinistralité et son effet sur la tarification.

Dans le premier chapitre, nous avons présenté un modèle de principal-agent appliqué sur le marché de l'assurance automobile. Ce modèle a couplé à la fois les deux problèmes d'asymétrie d'information : l'aléa moral et la sélection adverse. L'aversion à l'effort a constitué la principale caractéristique propre à l'agent, inobservable par l'assureur. L'effort de prévention de l'agent a été pris en considération comme étant influençable par la nature du contrat d'assurance souscrit. Nous avons montré qu'un équilibre mélangeant peut exister contrairement à Rothschild et Stiglitz (1976).

En effet, à partir d'un certain seuil de remboursement d'assurance, un même contrat d'assurance souscrit par une proportion d'agents incités à fournir de l'effort et une autre proportion d'agents ayant un niveau d'aversion à l'effort plus élevé et ne fournissant aucun effort de prévention, peut exister. Par conséquent, d'après notre modèle, il peut exister des contrats d'équilibre séparateurs mais aussi des contrats d'équilibre mélangeants. Un deuxième contrat d'équilibre peut aussi exister. C'est le contrat d'équilibre à la Rothscild et Stiglitz. En effet, nous avons montré que si l'aversion à l'effort est suffisamment faible, alors tous les agents choisissent un contrat avec la plus faible couverture qui les incitent à fournir l'effort de prévention nécessaire pour se prémunir contre le risque. Nous pouvons penser notamment au contrat obligatoire de responsabilité civile qui offre le remboursement minimum aux assurés.

Ce résultat théorique a été testé empiriquement, en se fondant sur des données en coupe instantanée provenant d'une compagnie d'assurance française. Afin de contrôler la variable de l'effort des individus, nous avons créé une nouvelle variable décrivant la sinistralité passée. Cette variable est construite en fonction du coefficient du bonus malus de l'assuré. La construction d'un indicateur de sinistralité passée, proche du CRM, mais qui fait aussi intervenir l'ancienneté de permis, permet d'élargir la classe des bons conducteurs aux jeunes conducteurs qui sont souvent considérés comme des mauvais risques. Du fait que cet indicateur a un bon pouvoir prédictif de la sinistralité, l'assureur pourrait l'utiliser pour accorder une réduction de prime à ces assurés afin de les fidéliser. Ils constitueront son portefeuille dans le futur. Les mauvais conducteurs restent de mauvais conducteurs.

L'utilisation des différents modèles bivariés et trivariés a permis de confirmer que l'estimation séparée des équations de choix de contrat et de sinistralité comporte des biais d'endogénéité. En effet, il existe plusieurs facteurs inobservés ou inobservables par l'assureur, aussi bien pour les jeunes conducteurs que pour les conducteurs expérimentés. L'hypothèse d'avantage informationnel de l'assuré par rapport à l'assureur est ainsi validée. Cette asymétrie d'information est interprétée principalement par la sélection adverse.

L'efficacité des variables détenues par l'assureur dans l'estimation de la sinistralité ont été aussi évaluée par des modélisations économétriques sur les coûts et les nombres d'accidents. Même si ces variables ne révèlent pas en totalité le type de risque des assurés, elles représentent néanmoins d'importantes informations qui nous ont permis de traiter la question de la surprime et son application pour les jeunes conducteurs : Est-il indispensable d'appliquer une surprime aux conducteurs novices contrairement au conducteurs expérimentés ? Notre réponse est non. Nous avons montré qu'il est intéressant pour l'assureur d'opter pour la fidélisation d'une clientèle potentielle en éliminant la surprime, à condition que la part des jeunes conducteurs attirés au début de période dépasse un certain seuil. Estimer le taux d'entrée des jeunes conducteurs sur le marché d'assurance automobile est donc indispensable avant d'appliquer la suppression de la surprime. Bien évidemment si la suppression de la surprime est une idée intéressante à prendre en considération, elle doit néanmoins être accompagnée d'une compagne de sécurité routière destinée aux jeunes conducteurs.

Notre recherche a permis d'appréhender les problèmes d'asymétrie d'information entre assureur et assuré ainsi que les questions relatives à la sinistralité sur le marché d'assurance automobile. Il est néanmoins important de souligner quelques limites de nos travaux. La principale limite concerne la base de données dont nous disposons. Dans le chapitre 2, nous avons créé une *proxy* de l'effort fondée sur la sinistralité passée (croisement entre le bonus malus, l'ancienneté du permis) et la sinistralité actuelle. Il en est ressorti qu'une partie des assurés a été désignée comme des *conducteurs moyens* ou ayant une aversion moyenne à l'effort. Nous soulignons que ces assurés peuvent être de mauvais conducteurs comme ils peuvent être de bons conducteurs. Mais le fait de ne pas observer les fréquences exactes des sinistres survenus avant l'année 2004 nous laisse incapable de faire cette distinction et d'être par conséquent imprécis sur le niveau d'effort fournis pas les assurés. La deuxième limite se rattache à l'année pendant laquelle les données ont été collectées. Ces données concernent l'année 2004, année après l'instauration de nouvelles mesures de sécurité routière en France, telles que la limitation de vitesse, du taux d'alcoolémie au volant, la mise en application du permis probatoire, les implantations de radars supplémentaires, etc. L'effet de ces mesures n'a pas été pris en considération dans l'analyse empirique de l'asymétrie d'information.

Ces derniers points ouvrent la voie à de nouvelles pistes de recherche, notamment en matière de comparaison entre le comportement des assurés avant et après l'instauration de nouvelles mesures sécuritaires afin d'étudier leurs effets sur le choix de la couverture d'assurance ainsi que sur la sinistralité. Il serait ainsi intéressant de travailler sur des données dynamiques françaises s'étalant avant et après l'année 2004. Ceci nous permet de réévaluer la sinistralité passée, ainsi que la variable décrivant l'aversion à l'effort dans une étude empirique sur l'hypothèse de l'existence de sélection adverse et d'aléa moral sur le marché d'assurance automobile français.

Tout au long du second chapitre, les modèles bivariés et trivariés appliqués sur la sinistralité et la couverture d'assurance supposent que la corrélation entre les termes résiduels des équations suit la loi normal. Le choix de ces modèles paramétriques a été bien évidemment volontaire pour que nous puissions situer nos résultats par rapport aux travaux de Chiappori et Salanié (2000) et les différents auteurs s'appuyant sur ce type de modèles paramétriques. Une deuxième perspective est alors envisageable. Elle concerne l'application de modèles non paramétriques analysant la relation risque-couverture, où l'hypothèse de normalité de la corrélation entre les résidus sera relâchée.

Une troisième perspective concerne l'extension envisageable de notre modèle théorique. Il serait intéressant de prendre en considération l'effet de la variation du coefficient du bonus malus sur l'espérance d'utilité des agents et par la suite sur la détermination du contrat optimal.

Annexes

Annexe 1

Sexe du conducteur	1	Masculin
	2	Féminin
Type de conducteur	1	L'assuré est un conducteur principal
	2	L'assuré est un conducteur secondaire
Usage de véhicule	1	Usage promenade-trajet
	2	Usage professionnel
Numéro de département	01 à 95	Département de métropole
Contrat d'assurance	1	RC
	2	Formule 1
	3	Formule 2
	4	Formule 3

	Catégorie socio-professionnelle
0	Absence d'information
1	Enseignant du primaire
2	Enseignant du secondaire ;
3	Enseignant du supérieur et chercheur
4	Enseignant d'établissement de formation professionnelle
5	Educateur technique ou spécialisé, animateur, moniteur
6	Profession artistique, culturelle et sportive
7	Profession libérale
8	Artisan
9	Commerçant
10	Exploitant agricole
11	Cadre supérieur
12	Cadre moyen
13	Technicien, agent de maîtrise, contremaître
14	Agent ou employé administratif
15	Personnel de service ou d'entretien
16	Ouvrier
17	Étudiant
18	Chômeur ou demandeur d'emploi
19	Homme ou femme au foyer
50	Au service national
51	En conge (formation, convenance personnelle)
52	A la retraite
90	Autres

Annexe 2

Analyse des données

Les principaux résultats de l'analyse bivariée entre les différentes variables sont décrits ci dessous.

TAB. 3.17: **Les sinistres et les coûts selon le sexe, le type du conducteur et la profession**

	sinistres>0	**coûts>0**		
	%	**Moyenne**	**Médiane**	**Ecart-type**
Sexe				
Homme	11,3	1764	1143.7	2158
Femme	12,34	1720.8	1152	1706.8
test du χ^2 = 10.65,		test de de comparaison de WMW : $Z_{WMW} = -1.1407$		
$Pr = 0.0011$		$Two - SidedPr(> \|Z\|) = 0.2540$, $One - SidedPr(< Z) = 0.1270$		
Type du conducteur				
Principal	12,25	1753	1143.7	1934.8
Secondaire	11,08	1717.5	1148	1927.3
test du χ^2 = 12.34,		test de de comparaison de WMW : $Z_{WMW} = -0.7943$		
$Pr = 0.0004$		$Two - SidedPr(> \|Z\|) = 0.4270$, $One - SidedPr(< Z) = 0.2135$		
Profession				
Enseignant	12	1736.9	1143.7	2078.3
Non enseignant	11,58	1747.5	1164	1686.4
test du χ^2 = 1.76,		test de de comparaison de WMW : $Z_{WMW} = 1.4438$		
$Pr = 0.1840$		$Two - SidedPr(> \|Z\|) = 0.1488$, $One - SidedPr(< Z) = 0.0744$		
Zone de résidence				
Zone urbaine	11.14	1805.1	1212.8	1849.1
Zone rurale	12.41	1835.7	1227.4	2128.2
test du χ^2 = 16.01,		test de de comparaison de WMW : $Z_{WMW} = -0.6156$		
$Pr < .0001$		$Two - SidedPr(> \|Z\|) = 0.5382$, $One - SidedPr(< Z) = 0.2691$		

Le tableau 3.17 présente plus de femmes sinistrées que d'hommes ayant au moins un accident responsable dans l'année : 12.34% des conductrices assurées ont déclaré au moins un sinistre responsable à leur assureur contre 11.3% pour les sociétaires masculins. D'après le test du χ^2, nous confirmons qu'il existe une relation statistiquement significative entre l'occurrence d'accidents responsables et le sexe de l'assuré. Ceci n'est pas le cas de ce qui est généralement constaté. Selon la Fédréation Française des Sociétés d'Assurance (FFSA), les femmes sont moins dangereuses que les hommes en matière de sécurité routière[6].

En termes de coûts, la moyenne des dépenses positives des sinistres causés par les femmes est supérieure à celle des hommes. Mais la différence des coûts entre les hommes et les femmes n'est pas significative d'après le test de comparaison de Wilcoxon-Mann-Whitney (WMW) [7] ($Two - SidedPr(> |Z|) = 0.1488$). Nous pouvons observer comme cas particulier les quelques sociétaires masculins qui ont causé des accidents graves ayant des valeurs élevées. La valeur maximum des coûts des sinistres est de 45972 euros pour les hommes, tandis que les femmes ont engendré au maximum des coûts de l'ordre de 23792 euros.

Le type du conducteur ainsi que la zone de résidence ont aussi un effet statistiquement significatif sur l'occurrence des sinistres responsables. Ces variables, par contre, n'ont pas d'impact significatif sur le coût des sinistres. En ce qui concerne la zone de résidence des assurés, nous remarquons que son effet sur la sinistralité

[6] C'est pour cette raison qu'il existe des avantages tarifaires pour les femmes par rapport aux hommes dans des prestations offertes par certaines compagnies d'assurance automobile.

[7] Le test de Wilcoxon-Mann-Whitney est une alternative non paramétrique au test t de student. Il est appliqué quelque soit la distribution de probabilité caractérisant les données (i.e. normale ou non). De plus, il est moins sensible aux valeurs extrêmes que le test classique de Student.

est différent de ce qui est généralement observé en matière de sécurité routière. Dans notre échantillon, il y a plus de sinistrés résidants dans les zones ruraux que ceux qui habitent en ville.

La catégorie socioprofessionnelle, quant à elle ne fait ressortir aucun lien significatif avec la sinistralité ni en termes de fréquence, ni en termes de coût.

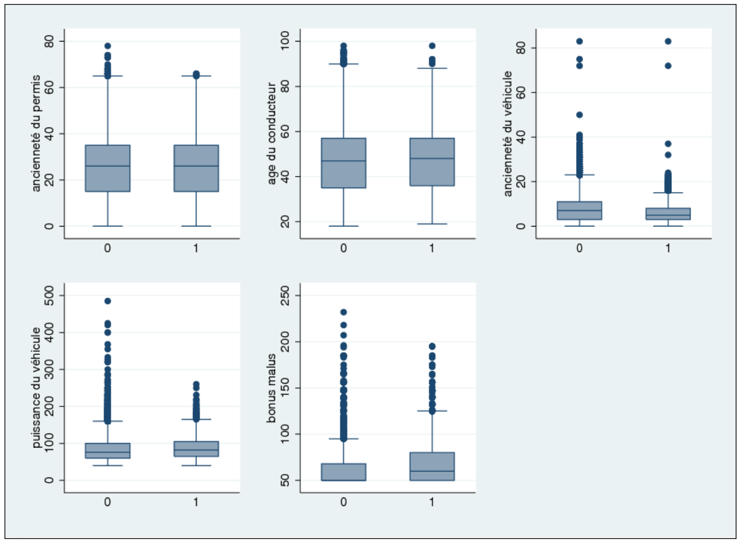
Le tableau 3.18 ainsi que les box plots de la figure 3.6 décrivent les caractéristiques observables de l'assureur et de son véhicule en fonction de la déclaration d'un accident responsable ou plus.

TAB. 3.18: **La déclaration d'un accident responsable ou plus et les caractéristiques observables**

Variables	**Moyenne**	**Ecart-type**	Z_{WMW}	
Ancienneté du permis	26	13	3.7850	***
Age du conducteur	47.42	14.2	4.0975	***
Ancienneté du véhicule	5.78	4.1	-21.5668	***
Puissance du véhicule	87.5	26.19	11.7603	***
Bonus Malus	66.45	22.78	16.709	***

$***$: significativité au seuil de 1%.
Z_{WMW} est la statistique du test de comparaison des distributions de chaque caractéristique du conducteur selon l'occurrence ou non des sinistres responsables.

Nous constatons, d'après le test de comparaison de Wilcoxon-Mann-Whitney, qu'il existe des différences statistiquement significatives des distributions de toutes les variables continues en fonction de la déclaration des sinistres. En ce qui concerne la variable du bonus malus, environ 6% des sociétaires qui ont le coefficient maximal de réduction-majoration ont déclaré au cours de l'année au moins un sinistre



$y = 1$: au moins un sinistre responsable sinon $y = 0$

FIG. 3.6 – **Les caractéristiques observables en fonction de l'occurrence des sinistres responsables**

responsable. Ils verront leur bonus diminuer et leur prime augmenter l'année suivante. Un assuré ayant atteint le meilleur coefficient réduction majoration aurait cependant intérêt à préserver son bonus et serait donc incité à plus de prudence au volant et ne déclarer aucun sinistre responsable. Une explication peut être donnée à ce niveau : en France, l'assureur n'applique aucune majoration, c'est-à-dire malus, pour le premier sinistre survenu après une période d'au moins de trois ans au cours de la quelle le coefficient de réduction majoration est égal à 50. Ainsi, parmi les sociétaires sinistrés pendant l'année ayant le bonus maximal, il existe ceux qui ne verront ni leur malus ni leur prime augmenter. Nous ne pouvons pas donner leur pourcentage exact car l'information qui concerne le nombre enregistré des accidents passés n'est pas disponible.

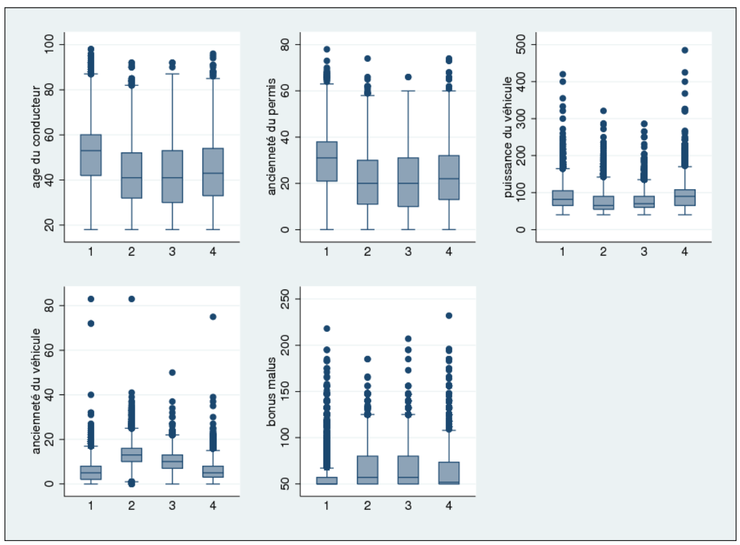
Les résultats croisés entre les contrats souscrits et les caractéristiques du conducteur et du véhicule sont décrits ci-dessous.

TAB. 3.19: **La répartition du portefeuille selon les sinistres responsables et leurs coûts par sexe, type du conducteur et par profession**

%	RC	Formule 1	Formule 2	Formule 3
Sexe				
Homme	38.99	19.31	9.62	32.08
Femme	45.42	14.79	8.38	31.41
$\chi^2 = 244.23,\ Pr < .0001$				
Type du conducteur				
Principal	43.28	16.19	8.67	31.85
Secondaire	40.40	18.50	9.56	31.54
$\chi^2 = 57.11,\ Pr < .0001$				
Profession				
Enseignant	45.49	14.77	8.40	31.34
Non enseignant	37.52	20.30	9.85	32.33
$\chi^2 = 355.56,\ Pr < .0001$				
Zone de résidence				
Zone urbaine	43.53	15.88	8.83	31.76
Zone rurale	41.18	17.96	9.13	31.72
$\chi^2 = 40.49,\ Pr < .0001$				
Total	42.25	17.02	8.99	31.74

Les sociétaires masculins ont plus de préférence pour les contrats tous risques que les femmes dont 45% ont souscrit la garantie obligatoire. En ce qui concerne le type du conducteur, la profession du sociétaire et la zone de résidence, c'est chez les conducteurs principaux, les enseignants et les urbains où il y a le plus de souscription de contrat de responsabilité civile par rapport aux autres formes de contrat d'assurance *tous risques*.

Les box plots de la figure 3.7 ainsi que le tableau 3.20 décrivent la distribution de l'échantillon par contrat d'assurance souscrit selon l'âge du conducteur, l'ancienneté du permis de conduire, l'ancienneté et la puissance du véhicule et le coefficient du bonus malus.



1, 2, 3 et 4 représentent respectivement les contrats RC, Formule 1, Formule 2 et Formule 3

FIG. 3.7 – **L'âge du conducteur, l'ancienneté du permis de conduire, l'ancienneté et la puissance du véhicule et le coefficient du bonus malus par contrat d'assurance**

Le test de comparaison des distribution de Kruskal-Wallis[8] indique qu'il existe une relation significative entre le type du contrat d'assurance choisi et les différentes caractéristiques observables prises une par une (l'ancienneté du permis, l'age du conducteur, l'ancienneté et la puissance du véhicule et le bonus malus).

TAB. 3.20: **Les type du contrat et les caractéristiques observables : le test de comparaison de Kruskal-Wallis**

Variables	Test de Kruskal-Wallis : χ^2=	
Ancienneté du permis	3659.6	***
Age du conducteur	3582.3	***
Ancienneté du véhicule	13762.6	***
Puissance du véhicule	2357.4	***
Bonus Malus	2863.1	***

$***$: significativité au seuil de 1%

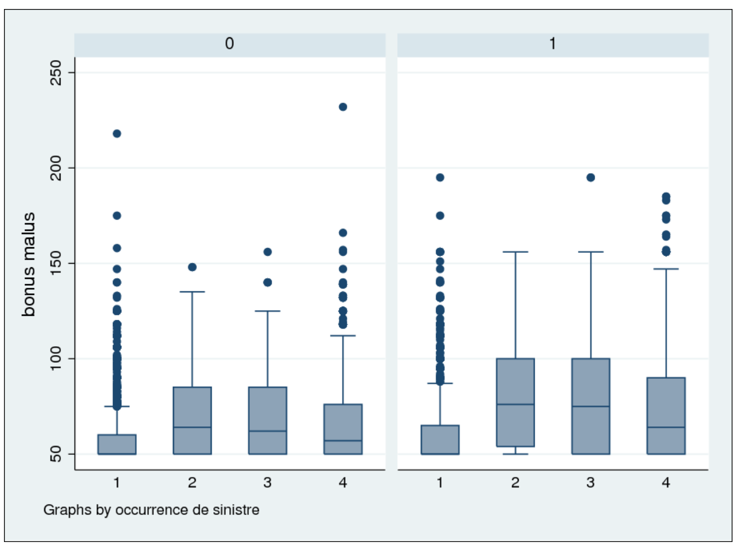
Le test de Kruskal-Wallis compare les distributions de chaque caractéristique du conducteur en fonction du type du contrat d'assurance

La principale remarque que nous énonçons à ce niveau est que trois quarts des sociétaires ayant souscrit le contrat obligatoire (77%) ont des coefficients de réduction majoration ne dépassant pas 57 : les bas risques optent plus pour les contrats RC. Ceci n'est pas le cas pour le reste des assurés ayant choisi des cotisations dommage au véhicule, pour les quels le malus est plus important. Ceci semble justifier l'hypothèse de la sélection adverse qui stipule que les hauts risques optent pour les contrats à forte couverture. Bien évidemment, ce n'est qu'une simple observation

[8]Le test de Kruskal-Wallis est la généralisation à K populations du test de la somme des rangs de Wilcoxon-Mann-Whitney bilatéral. On le considère comme l'alternative non paramétrique de l'ANOVA dès que la distribution sous-jacente des données n'est plus gaussienne.

qui doit être vérifiée et éventuellement appuyée par une étude économétrique en prenant en compte toutes les informations décrites par les variables dont nous disposons. En plus, la notion du risque sur laquelle est basée cette intuition ne prend en compte que la variable du bonus malus décrivant la sinistralité passée et non pas les accidents durant l'année.

En croisant l'occurrence des sinistres et les contrats d'assurance avec les autres variables, nous constatons principalement que les plus bas risque - c'est-à-dire les assurés qui ont le bonus maximal et qui n'on enregistré aucun sinistre responsable pendant l'année - ont opté pour la garantie de la responsabilité civile. Une grande partie des plus hauts risques en termes de bonus-malus, c'est à dire ceux qui ont déclaré des sinistres responsables et dont le coefficient du bonus malus est élevé, a souscrit des garanties tous risques (voir figure 3.8).



1, 2, 3 et 4 représentent respectivement les contrats RC, Formule 1, Formule 2 et Formule 3

FIG. 3.8 – **Le coefficient du bonus malus par contrat d'assurance et occurrence des sinistres**

Annexe 3

Jeunes conducteurs *vs* Conducteurs expérimentés : analyse exploratoire des données

La fréquence des sinistres Les distributions des sinistres responsables déclarés par les jeunes conducteurs et les conducteurs expérimentés sont décrits dans les tableaux 3.21.

TAB. 3.21: **Le nombre d'accidents observés chez les jeunes conducteurs et les conducteurs expérimentés**

Nombre de sinistres	**Jeunes conducteurs**		**Conducteurs expérimentés**	
0	1029	(89.56%)	35451	(88.13%)
1	36	(3.13%)	1449	(3.6%)
2	36	(3.13%)	1750	(4.35%)
>3	48	(4.18%)	1576	(3.91%)
Test de comparaison de WMW :				
$Z_{WMW} = -1.3648,\ Two - SidedPr(> \vert Z\vert) = 0.1723,\ One - SidedPr(< Z) = 0.0862$				

Le nombre maximum de sinistres déclarés par un assuré vaut 9 pour les novices et 10 pour les expérimentés. Environ 11% des assurés (les deux catégories confondues de conducteurs) ont enregistré au moins un accident aucours de l'année 2004. La fréquence moyenne des sinistres pour les jeunes conducteurs et les conducteurs expérimentés est respectivement de 27.24% et 27% par an. Les jeunes conducteurs, bien qu'ils représentent 3% de la totalité de l'échantillon, déclarent en moyenne autant d'accidents responsables que les conducteurs expérimentés (mais avec un écat-type plus élévé de 0.965 contre 0.861 pour les conducteurs expérimentés). Ceci est confirmé par le test de Wilcoxon-Mann-Whitney (WMW)

qui compare la distribution de la sinistralité entre les jeunes conducteurs et les conducteurs expérimentés. Ce test révèle que la différence de distribution de la sinistralité entre les jeunes conducteurs et les conducteurs expérimentés n'est pas statistiquement significative au seuil de 5%.

TAB. 3.22: **L'occurrence des sinistres pour les jeunes et les expérimentés**

	Accidents = 0	**Accidents > 0**
Jeunes conducteurs	89.56	10.44
Conducteurs expérimentés	88.13	11.87
Test du $\chi^2 = 2.1794$, $Pr = 0.1399$		

Le tableau 3.22 montre qu'il n'y a pas de relation significative entre l'occurrence d'au moins un sinistre responsable et le fait d'être jeune conducteur ou expérimenté. Cela confirme le résultat précédent.

Nous dressons dans les tableaux 3.23 la fréquence des accidents selon le sexe de l'assuré pour les deux catégories de conducteurs.

TAB. 3.23: **Le nombre des accidents observés selon le sexe chez les jeunes conducteurs et les conducteurs expérimentés**

Nombre de sinistres	**Jeunes conducteurs**				**Conducteurs expérimentés**			
	Homme		**Femme**		**Homme**		**Femme**	
0	551	(87.88%)	478	(91.57%)	17511	(88.72%)	17940	(87.56%)
1	22	(3.51%)	14	(2.68%)	698	(3.54%)	751	(3.67%)
2	23	(3.67%)	13	(2.49%)	810	(4.10%)	940	(4.59%)
>3	31	(4.94%)	17	(3.25%)	718	(3.65%)	858	(4.18%)
Total	627	(54,57%)	522	(45,43%)	19737	49.07%)	20489	(50.93%)

La comparaison entre l'occurrence des sinistres des jeunes hommes et des hommes expérimentés, de la même manière pour les conductrices femmes est présentée dans le tableau 3.25 suivant.

TAB. 3.24: **Le sexe et l'occurrence des sinistres pour les jeunes et les expérimentés**

	Homme		**Femme**	
	Accidents = 0	**Accidents > 0**	**Accidents = 0**	**Accidents > 0**
Jeunes conducteurs	87.88	12.12	91.57	8.43
Conducteurs expérimentés	88.72	11.28	87.56	12.44
	Test du $\chi^2 = 0.4306$		**Test du** $\chi^2 = 7.5727$	
	$Pr = 0.5117$		$Pr = 0.0059$	

Le test du χ^2 a révélé, en ce qui concerne les sociétaires hommes, qu'il n'y a pas de relation significative entre l'occurrence des sinistres et le fait d'être jeune ou expérimenté. Par contre, pour les femmes assurées, la différence est statistiquement significative. Il y a plus de femmes expérimentées qui ont déclaré au moins un sinistre responsable que de femmes qui ont moins de deux ans d'expérience de conduite. Nous avons aussi tester la liaison entre le sexe et l'occurrence des sinistres au sein de chaque groupe de conducteurs et nous avons trouvé des χ^2 significatifs.

Tab. 3.25: **L'occurrence des sinistres en fonction de la profession, le lieu de résidence et le type du conducteur pour les jeunes et les expérimentés**

		Jeunes		Expérimentés	
	%	**Accidents = 0**	**Accidents > 0**	**Accidents = 0**	**Accidents > 0**
Profession	**Enseignant**	89.57	10.43	87.95	12.05
	Autres	89.54	10.46	88.39	11.61
		Test du $\chi^2 = 0.4306$		**Test du** $\chi^2 = 0.4306$	
Lieu de résidence	**Urbain**	90.60	9.40	88.81	11.19
	Rural	88.69	11.31	87.56	12.44
		Test du $\chi^2 = 0.4306$		**Test du** $\chi^2 = 0.4306 * **$	
le type du conducteur	**Principal**	88.31	11.69	87.75	12.25
	Secondaire	89.75	10. 25	88.86	11.14
		Test du $\chi^2 = 0.4306$		**Test du** $\chi^2 = 0.4306 * **$	

$* * *$: significativité au seuil de 1%

Le tableau 3.25 montre qu'il existe, seulement pour les conducteurs expérimentés, des effets significatifs du lieu de résidence et du type du conducteur sur l'occurrence des sinistres. Pour les jeunes conducteurs, aucune relation n'est significative.

TAB. 3.26: **La déclaration d'un accident responsable ou plus et les caractéristiques observables**

	Jeunes		**Expérimentés**			
Variables	**Moyenne**	**Ecart-type**	**Moyenne**	**Ecart-type**	Z_{WMW}	
Ancienneté du permis	1,62	0.55	26.5	12.4	-18.74	***
Age du conducteur	21.9	5	48	13.7	-17.83	***
Ancienneté du véhicule	8.7	5	5.7	4	7.11	***
Puissance du véhicule	67	14.89	88	26	-9.27	***
Bonus Malus	107	19.7	65.4	21.8	16.46	***

$***$: significativité au seuil de 1%.
Z_{WMW} est la statistique du test de comparaison des distributions
de chaque caractéristique du conducteur selon la catégorie des conducteurs

Le tableau 3.26 décrit les distributions des caractéristiques des jeunes conducteurs et des conducteurs expérimentés qui ont déclaré au moins un accident responsable durant l'année 2004. Le test de WMN montre que les différences de distribution entre les jeunes sinistrés et les expérimentés sinistrés sont statistiquement significatives selon l'age, l'ancienneté du permis, les caractéristiques du véhicule et le bonus malus.

Le risque assuré est mesuré non seulement en terme de fréquence de sinistres mais aussi en terme de coût. Il est ainsi indispensable d'analyser les coûts des sinistres qui représente la charge supportée par la compagnie d'assurance.

Les coûts des sinistres Dans cette sous section, nous comparons les coûts des sinistres entre les jeunes conducteurs et les conducteurs expérimentés. D'après l'analyse précédente de la fréquence des sinistres, environ 10% des assurés, qu'ils soient jeunes ou expérimentés, ont déclaré au moins un accident causant des charges à la compagnie d'assurance. Nous nous intéressons à la sinistralité de ces assurés afin de mesurer la gravité des accidents déclarés. Nous analysons donc les coûts strictement positifs. Ces derniers représentent soit les coûts cumulés des sinistres causés par chaque assuré et réellement dépensés par la compagnie d'assurance, soit les coûts évalués *à priori* par l'assureur.

TAB. 3.27: **La distribution des coûts réels et évalués des sinistres chez les jeunes et les expérimentés**

	Moyenne	**Ecart-type**	**Min.**	**Max.**
les coûts réels cumulés des sinistres				
Jeunes conducteurs	2080	2144.5	505	17754
Conducteurs expérimentés	1816	2009	500	45971
Test de comparaison de WMW :				
$Z_{WMW} = 1.4018$, $Two - SidedPr(> \|Z\|) = 0.1610$, $One - SidedPr(< Z) = 0.0805$				
les coûts évalués cumulés des sinistres				
Jeunes conducteurs	709	880	122.5	8914
Conducteurs expérimentés	555	966	100.5	9299
Test de comparaison de WMW :				
$Z_{WMW} = 4.4074$, $Two - SidedPr(> \|Z\|) =< .0001$, $One - SidedPr(< Z) =< .0001$				

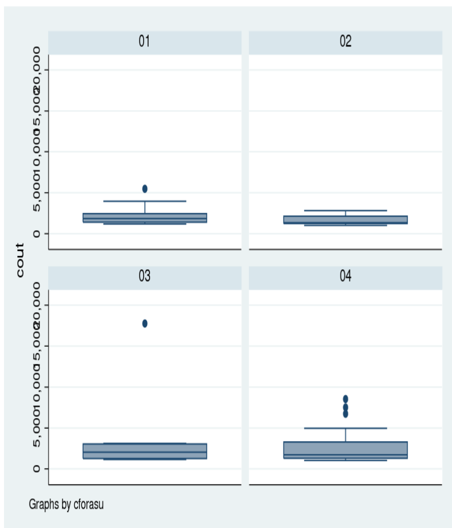
Le tableau 3.27 décrit les distributions des coûts cumulés positifs et des coûts évalués de sinistres. Malgré leur faible effectif dans l'échantillon total des sociétaires, nous remarquons que les jeunes conducteurs déclarent des sinistres aussi coûteux en moyenne que ceux déclarés par les conducteurs expérimentés. Ce point est confirmé principalement par le test de comparaison des distributions des coûts réels entre les jeunes conducteurs et les conducteurs expérimentés, qui montre qu'il n'existe pas de différence statistiquement significative (*Two-Sided Pr (> |Z|)= 0.1610*).

Les coûts des sinistres et les contrats d'assurance Le tableau suivant 3.28 ainsi que la figure 3.9 illustrent la distribution des coûts positifs des sinistres selon les différents contrats choisis par les jeunes conducteurs et les conducteurs expérimentés.

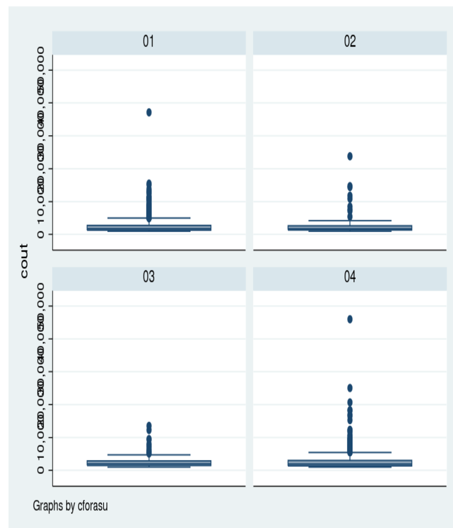
TAB. 3.28: **La distribution des coûts des sinistres pour les jeunes conducteurs et les conducteurs expérimentés selon le choix du contrat d'assurance**

	Contrats	Moyenne	Ecart-type	Min.	Max.
Jeunes conducteurs	**RC**	1919	1229.5	747.5	5486.4
	Formule 1	1292	579.3	512.7	2824.6
	Formule 2	2154.5	3664	545.9	17753.7
	Formule 3	2167.6	1758.6	505.3	8531.2
Test de Kruskal-Wallis : $\chi^2 = 7.3113$, $Pr = 0.0626$					
Conducteurs expérimentés	**RC**	1638.5	1678.7	500.4	37147
	Formule 1	1753.8	2257.2	501.2	23791.9
	Formule 2	1787.5	1761.4	501.2	13529
	Formule 3	1891.7	2264.9	500.1	45971
Test de Kruskal-Wallis : $\chi^2 = 28.9787$, $Pr < .0001$					

D'après le tableau 3.28 et la figure 3.9, les jeunes conducteurs ainsi que les conducteurs expérimentés ayant souscrit les contrats avec les plus fortes couvertures déclarent les accidents les plus coûteux en moyenne, plus spécialement pour les contrats *Formule 2* et *Formule 3*. Ces résultats vont dans le même sens que l'occurrence des sinistres en fonction des contrats d'assurance. La fréquence et le coût des sinistres sont ainsi plus élevés pour les polices ayant choisi des contrats *tous risques*. Cependant, d'après le test de Kruskal-Wallis, il n'existe pas de relation statistiquement significative entre le type de contrat et le coût des accidents responsable déclarés par les jeunes conducteurs. Pour les conducteurs expérimentés, la différence des coûts selon les quatre contrats d'assurance est significative ($\chi^2 = 28.9787$, $Pr < .0001$).



Jeunes conducteurs



Conducteurs expérimentés

FIG. 3.9 – **Les coûts réels des sinistres selon les contrats d'assurance**

Annexe 4

TAB. 3.29: **Les effets marginaux**

Variables	**Jeunes conducteurs**	**Conducteurs expérimentés**
L'equation de la couverture		
Homme	n.s.	0.0558628
Homme	n.s.	0.000723815
		0.0025622
		0.0289192
Age du conducteur	n.s.	-0.000111881
		-0.000396039
		-0.0044701
Enseignant	-0.0245151	-0.000820831
	0.0044723	-0.0029056
	0.0436475	-0.0327954
Zone urbaine	n.s.	n.s.
Ancienneté du permis	n.s.	-0.000050407
		-0.000178433
		-0.0020140
Conducteur principal	n.s.	0.000020388
		0.000072172
		0.000814598
Bonus malus	n.s.	0.000038090
		0.000134832
		0.0015218
Ancienneté du véhicule	0.0100724	0.000164558
	-0.0018375	0.000582511
	-0.0179332	0.0065747
Puissance du véhicule	n.s.	0.000030094
		0.000106526
		0.0012024
L'equation du sinistre		

Homme	n.s.	-0.0012427
Age du conducteur	n.s.	-0.0014657
Enseignant	n.s.	-0.0144990
Zone urbaine	n.s.	n.s.
Ancienneté du permis	0.0623750	0.000748881
Conducteur principal	n.s.	0.0081382
Bonus malus	0.0065101	0.0035279
Puissance du véhicule	n.s.	0.000872303
Formule 1	-0.2006199	-0.2544361
Formule 2	-0.2202370	-0.2575197
Formule 3	-0.2541055	-0.3647596

Chaque valeur représente l'effet marginal de la variable X_k sur la probabilité que $C = 1,\ 2,\ ou\ 3$

ns signifie que la variable est non significative

Bibliographie

[1] J.H. Abbring, P.A. Chiappori, J.J. Heckman et J. Pinquet : Testing for Moral Hazard on Dynamic Insurance Data. *THEMA Working Papers, 2002-24*, 2002.

[2] J.H. Abbring, P.A. Chiappori et J. Pinquet : Moral hazard and dynamic insurance data. *Journal of the European Economic Association*, 1(4):767–820, 2003 (a).

[3] J.H. Abbring, P.A. Chiappori et T. Zavadil : Better safe than sorry ? ex ante and ex post moral hazard in dynamic insurance data. Discussion Paper 2008-77, Tilburg University, Center for Economic Research, 2008.

[4] J.H. Abbring, J.J. Heckman, P.A. Chiappori et J. Pinquet : Adverse selection and moral hazard in insurance : Can dynamic data help to distinguish ? *Journal of the European Economic Association*, 1(2-3):512–521, 2003 (b).

[5] Y. AIarie, G. Dionne et L. Eeckhoudt : Increases in Risk and the Demand for Insurance. *Contributions to insurance economics*, pages 275–294, 1992.

[6] G. Akerlof : The market for lemons : Qualitative uncertainty and the market mechanism. *Quarterly Journal of Economics*, 84(3):488–500, 1970.

[7] J.F. Angers, D. Desjardins, G. Dionne et F. Guertin : Vehicle and fleet random effects in a model of insurance rating for fleets of vehicles. *Astin Bulletin*, 36(1):25–77, 2006.

[8] R. Arnott : Moral hazard and competitive insurance markets. *in Contributions to Insurance Economics, G. Dionne (ed)*, pages 325–358, 1992.

[9] R.J. Arnott et J.E. Stiglitz : The basic analytics of moral hazard. *The Scandinavian Journal of Economics*, 90(3):383–413, 1988.

[10] R.J. Arnott et J.E. Stiglitz : Equilibrium in competitive insurance markets with moral hazard. *National Bureau of Economic Research Working Paper, 3588*, 1991.

[11] K.J. Arrow : *Essays in the theory of risk-bearing.* North-Holland, 1974.

[12] B. Belhadji et G. Dionne : Development of an expert system for automatic detection of automobile insurance fraud. *Cahier de recherche, HEC Montréal-Chaire de gestion des risques.*, (97-06), 1997.

[13] D. Bernoulli : Specimen theoriae novae de mensura sortis. *Commentarii Academiae Scientiarum Imperialis Petropolitanae. 1738 - Traduit en anglais dans Econometrica*, 22:23–36, 1954.

[14] P. Bolton et M. Dewatripont : *Contract theory.* The MIT Press, 2005.

[15] M. Boyer et G. Dionne : Riscophobie et étalement à moyenne constante : analyse et applications. *L'Actualité économique*, 59(2):208–229, 1983.

[16] M. Boyer et G. Dionne : The economics of road safety. Cahiers de recherche 8554, Universite de Montreal, Departement de sciences economiques, 1985.

[17] M. Boyer et G. Dionne : La tarification de l'assurance automobile et les incitations a la securite routiere : une etude empirique. Cahiers de recherche 8553, Universite de Montreal, Departement de sciences economiques, 1985.

[18] M. Boyer et G. Dionne : An empirical analysis of moral hazard and experience rating. *The Review of Economics and Statistics*, 71(1):128–134, 1989.

[19] M. Boyer et G. Dionne : More on insurance, protection, and risk. *Canadian Journal of Economics*, 22(1):202–204, 1989.

[20] M Boyer, G Dionne et R Khilstrom : Insurance and the value of publicly available information. *Studies in the Economics of Uncertainty*, page 137–156, 1989.

[21] M. Boyer, G. Dionne et R. Kihlstrom : Public information and experience rating. Cahiers de recherche 8627, Universite de Montreal, Departement de sciences economiques, 1986.

[22] M. Boyer, G. Dionne et C. Vanasse : Econometric Models of Accident Distribution. *Contributions to Insurance Economics*, pages 169–213, 1992.

[23] A. Chassagnon : Heterogenous risk in a moral hazard economy with adverse selection. *mimeo, Paris School of Economics*, 2005.

[24] A. Chassagnon et P.A. Chiappori : Insurance under moral hazard and adverse selection : The case of pure competition. *Delta working paper, 28*, 1997.

[25] A. Chassagnon et P.A. Chiappori : Insurance under moral hazard and adverse selection : efficiency and competition. *mimeo*, 2005.

[26] P.A. Chiappori : *Risque et assurance*. Flammarion, 1996.

[27] P.A. Chiappori : Asymmetric information in automobile insurance : an overview. *Automobile Insurance : Road Safety, New Drivers, Risks, Insurance Fraud and Regulation*, pages 1–12, 1999.

[28] P.A. Chiappori : Econometric Models of Insurance under Asymmetric Information. *Handbook of Insurance*, pages 365–394, 2001.

[29] P.A. Chiappori et J.J. Heckman : Testing Moral Hazard vs. Adverse Selection in Insurance Markets. *manuscrito, Universidad de Chicago, Chicago*, 2000.

[30] P.A. Chiappori, B. Jullien, B. Salanie et F. Salanie : Asymmetric information in insurance : General testable implications. *Rand Journal of Economics*, 37(4):783–798, 2006.

[31] P.A. Chiappori, I. Macho, P. Rey et B. Salanié : Repeated moral hazard : The role of memory, commitment, and the access to credit markets. *European Economic Review*, 38(8):1527–1553, 1994.

[32] P.A. Chiappori et B. Salanié : Testing Contract Theory : A Survey of Some Recent Work. *Advances in Economics and Econometrics, Cambridge University Press*, pages 115–150, 2003.

[33] P.A. Chiappori et B. Salanie : Testing for asymmetric information in insurance markets. *Journal of Political Economy*, 108(1):56–78, 2000.

[34] A. Cohen : Asymmetric information and learning : Evidence from the automobile insurance market. *Review of Economics and Statistics*, 87(2): 197–207, 2005.

[35] A. Cohen et L. Einav : Estimating risk preferences from deductible choice. *The American economic review*, 97(3):745–788, 2007.

[36] A. Cohen et P. Siegelman : Testing for adverse selection in insurance markets. *Journal of Risk and Insurance*, 77(1):39–84, 2010.

[37] M. Cohen et J.M. Tallon : Décision dans le risque et l'incertain : L'apport des modèles non additifs. *Revue d'économie politique*, 110(5):631–681, 2000.

[38] K. Crocker et A. Snow : The theory of risk classification. *Handbook of insurance*, pages 245–276, 2000.

[39] J.D. Cummins : *Deregulating property-liability insurance : restoring competition and increasing market efficiency.* AEI-Brookings Joint Center for Regulatory Studies, 2002.

[40] D. M. Cutler, A. Finkelstein et K. McGarry : Preference Heterogeneity and Insurance Markets : Explaining a Puzzle of Insurance. *American Economic Review*, 98(2):157–62, May 2008.

[41] K. Dachraoui, G. Dionne, L. Eeckhoudt et P. Godfroid : Comparative mixed risk aversion : definition and application to self-protection and willingness to pay. *Journal of Risk and Uncertainty*, 29(3):261–276, 2004.

[42] M. Dahchour et G. Dionne : Pricing of automobile insurance under asymmetric information : a study on panel data. THEMA Working Papers 2002-12, 2002.

[43] M. Dahchour et G. Dionne : Pricing of automobile insurance under asymmetric information : A study on panel data. *THEMA Working Papers, numéro 2002-12*, 2002.

[44] B. Dahlby : Testing for asymmetric information in Canadian automobile insurance. *Contributions to Insurance Economics*, pages 423–443, 1992.

[45] B.G. Dahlby : Adverse selection and Pareto improvements through compulsory insurance. *Public Choice*, 37(3):547–558, 1981.

[46] B.G. Dahlby : Adverse selection and statistical discrimination : : An analysis of Canadian automobile insurance. *Journal of Public Economics*, 20(1):121–130, 1983.

[47] P.J. Danaher et M.S. Smith : Modeling multivariate distributions using copulas : Applications in Marketing. *Journal of Marketing Science (forthcoming)*, 2009.

[48] D. De Meza et D.C. Webb : Advantageous selection in insurance markets. *RAND Journal of Economics*, 32(2):249–262, 2001.

[49] M. Denuit et A. Charpentier : Mathématiques de l'assurance non-vie Tome 1 : principes fondamentaux de théorie du risque (coll. économie et statistiques avancées). 2004.

[50] M. Denuit et A. Charpentier : Mathématiques de l'assurance non-vie Tome 2 : tarification et provisionnement (Coll. Economie et statistiques avancées). 2005.

[51] M. Denuit, X. Maréchal, S. Pitrebois et J.F. Walhin : *Actuarial modelling of claim counts.* John Wiley and Sons, Ltd, 2007.

[52] G. Dionne : Les mesures empiriques des problèmes d'information. *L'Actualité Economique*, 74(4):585–606, 1998.

[53] G. Dionne : Offre d'assurance non vie : une revue de la littérature récente. *Cahier de recherche, 98-01, HEC Montréal*, 1998.

[54] G. Dionne : *Handbook of insurance*. Kluwer Academic Publishers, 2000.

[55] G. Dionne : Commitment and automobile insurance in france, quebec and japan. THEMA Working Papers 2001-10, 2001.

[56] G. Dionne : Introduction to the Special Issue on Long-Term Care Insurance and Health Insurance. *Journal of Risk and Insurance*, 76(1):1–4, 2009.

[57] G. Dionne, N. Doherty et N. Fombaron : Adverse Selection in Insurance Markets. *Handbook of insurance*, pages 185–244, 2001.

[58] G. Dionne, C. Fluet et D. Desjardins : Predicted risk perception and risk-taking behavior : The case of impaired driving. *Journal of Risk and Uncertainty*, 35(3):237–264, 2007.

[59] G. Dionne, C.D. Fluet, D. Desjardins et S. Messier : La perception des risques d'accident et d'arrestation lors de conduite avec facultés affaiblies. *Cahier de recherche numéro 04-02, THEMA*, 2004.

[60] G. Dionne, C. Gouriéroux et C. Vanasse : Testing for evidence of adverse selection in the automobile insurance market : A comment. *Journal of Political Economy*, 109(2):444–453, 2001.

[61] G. Dionne et P. Lasserre : Adverse Selection, Repeated Insurance Contracts and Annoucement Strategy. *The Review of Economic Studies*, 52(4):719–723, 1985.

[62] G. Dionne et P. Lasserre : Dealing with moral hazard and adverse selection simultaneously. Cahiers de recherche 8559, Universite de Montreal, Departement de sciences economiques, 1985.

[63] G. Dionne, P.C. Michaud et M. Dahchour : Separating Moral Hazard from Adverse Selection and Learning in Automobile Insurance : Longitudinal Evidence from France. *American Risk and Insurance Association Meeting, Washington*, 2006.

[64] G. Dionne et J. Pinquet : Mesure des effets incitatifs à la prudence au volant créés par les sanctions et évaluation du pouvoir prédictif des infractions sur le risque routier. *Cahier de recherche numéro 05-06, HEC Monréal*, 2005.

[65] G. Dionne, J. Pinquet, M. Maurice et C. Vanasse : Incentive Mechanisms for Safe Driving : A Comparative Analysis with Dynamic Data. *The Review of Economics and Statistics*, 92(4), 2010.

[66] G. Dionne et C. Vanasse : A generalization of automobile insurance rating models : the negative binomial distribution with a regression component. *ASTIN*, 19(2):199–213, 1989.

[67] G. Dionne et C. Vanasse : Automobile insurance ratemaking in the presence of asymmetrical information. *Journal of Applied Econometrics*, 7(2):149–165, 1992.

[68] G. Dionne et C. Vanasse : Une evaluation empirique de la nouvelle tariffication de l'assurance automobile (1992) au Quebec. *Actualité Économique*, 73:47–80, 1997.

[69] N.A. Doherty et G. Dionne : Insurance with undiversifiable risk : Contract structure and organizational form of insurance firms. *Journal of Risk and Uncertainty*, 6(2):187–203, 1993.

[70] I. Ehrlich et G.S. Becker : Market insurance, self-insurance, and self-protection. *The Journal of Political Economy*, 80(4):623–648, 1972.

[71] M.-C. Fagart et B. Kambia-Chopin : Aléa moral et sélection adverse sur le marché de l'assurance. *CREST Working Papers, 2003-39*, 2003.

[72] M.C. Fagart : Concurrence en contrats, antisélection et structure d'information. *Annales d'Economie et de Statistique*, 43:1–27, 1996.

[73] M.C. Fagart : Wealth effects, moral hazard and adverse selection in a principal-agent model. *CREST Working Papers, numéro 2002-15*, 2002.

[74] A. Finkelstein et J. Poterba : Adverse selection in insurance markets : Policyholder evidence from the UK annuity market. *Journal of Political Economy*, 112(1):183–208, 2004.

[75] J. Geweke : Efficient simulation from the multivariate normal and student-t distributions subject to linear constraints and the evaluation of constraint probabilities. *Computing science and statistics : proceedings of the 23rd symposium on the interface*, pages 571–578, 1991.

[76] J. Geweke, M. Keane et D. Runkle : Alternative computational approaches to inference in the multinomial probit model. *The Review of Economics and Statistics*, pages 609–632, 1994.

[77] R.D. Gibbons et V. Wilcox-Gok : Health Service Utilization and Insurance Coverage : A Multivariate Probit Analysis. *Journal of the American Statistical Association*, 93(441):63–72, 1998.

[78] C. Gollier : Risque et incertitude. *Encyclopedia Universalis Notions*, pages 904–906, 2004.

[79] C. Gourieroux : *Statistique de l'assurance*. Economica, 1999.

[80] C. Gourieroux, A. Monfort et A. Trognon : Pseudo maximum likelihood methods : Applications to poisson models. *Econometrica*, 52(3):701–20, 1984.

[81] C. Gourieroux, A. Monfort et A. Trognon : Pseudo maximum likelihood methods : Theory. *Econometrica*, 52(3):681–700, 1984.

[82] C.W.J. Granger : Investigating causal relations by econometric models and cross-spectral methods. *Econometrica : Journal of the Econometric Society*, pages 424–438, 1969.

[83] W.H. Greene : *Econometric analysis, fifth edition*. prentice Hall, 2003.

[84] M. Grun-Réhomme : Prévision du risque et tarification : le rôle du bonus-malus français. *Assurance et gestion des risques, Montréal*, 68(1):21–30, 2000.

[85] M. Grun-Réhomme : Sélection adverse, sinistralité et choix de contrat chez les jeunes conducteurs : La pertinence des modèles. *Risques, les cahiers de l'assurance*, 75:133–141, 2008.

[86] M. Grun-Réhomme et N. Benlagha : Choix de contrat et sinistralite chez les jeunes conducteurs. *Assurance et gestion des risques, HEC Montréal*, 74(4):505–532, 2007.

[87] M. Grun-Réhomme et V. Joly : Risque individuel et choix de contrat : Le cas de l'assurance automobile. *Assurance et gestion des risques, Montréal*, 71(1):145–162, 2003.

[88] V. Hajivassiliou, D. McFadden et P. Ruud : Simulation of multivariate normal orthant probabilities : Methods and programs. *Journal of Econometrics*, 72(1-2):85–134, 1995.

[89] V. Hajivassiliou, D. McFadden et P. Ruud : Simulation of multivariate normal rectangle probabilities and their derivatives theoretical and computational results* 1. *Journal of Econometrics*, 72(1-2):85–134, 1996.

[90] V.A. Hajivassiliou : Simulation estimation methods for limited dependent variable models. *Handbook of Statistics, ed. Maddala, Rao and Vinod, New York : Elsevier Science Publishing*, 1993.

[91] J.J. Heckman et G.J. Borjas : Does unemployment cause future unemployment ? Definitions, questions and answers from a continuous time model of heterogeneity and state dependence. *Economica*, 47(187):247–283, 1980.

[92] J.J. Heckman et E.E. Leamer : *Handbook of econometrics*. North-Holland, 2007.

[93] D. Henriet et J.C. Rochet : *Microéconomie de l'assurance*. Economica, Paris, 1991.

[94] B. Holmström : Moral hazard and observability. *The Bell Journal of Economics*, 10(1):74–91, 1979.

[95] B. Jullien, B. Salanie et F. Salanie : Screening risk-averse agents under moral hazard : Single-crossing and the CARA case. *Economic Theory*, 30(1):151–169, 2007.

[96] M.P. Keane : A computationally practical simulation estimator for panel data. *Econometrica : Journal of the Econometric Society*, 62(1):95–116, 1994.

[97] H. Kim, D. Kim, S. Im et J.W. Hardin : Evidence of Asymmetric Information in the Automobile Insurance Market : Dichotomous Versus Multinomial Measurement of Insurance Coverage. *Journal of Risk and Insurance*, 76(2):343–366, 2009.

[98] F.H. Knight : Risk, uncertainty and profit. *Boston and New York*, 1921.

[99] M. Kouki-Zekri et D. Gaumont : Is there any asymmetric information in a cross-sectional insurance data set ? The impact of Aversion To Effort. *ERMES Working Paper*, (10-02), 2010.

[100] J.J. Laffont : *Economie de l'incertain et de l'information, volume 2 du cours de Théorie Microéconomique.* Economica, Paris, 1991.

[101] J.J. Laffont et D. Martimort : *The theory of incentives : the principal-agent model.* Princeton University Press, 2002.

[102] E. Lawrence : Simulated Maximum Likelihood via the GHK Simulator : An Application to the 1988 Democratic Super Tuesday Primary. *Political Science Association, Chicago*, 1997.

[103] S. Lerman et C. Manski : On the use of simulated frequencies to approximate choice probabilities. *Structural analysis of discrete data with econometric applications*, 10:305–319, 1981.

[104] S. Lollivier : Endogénéité d'une variable explicative dichotomique dans le cadre d'un modèle bivarié. *Annales d'économie et de statistique*, 62:251–269, 2001.

[105] M.J. Machina et M. Rothschild : Risk. *The New Palgrave Dictionary of Economics, 2nd edition, ed. by Steven N. Durlauf and Lawrence E. Blume*, pages 190–197, 2008.

[106] G.S. Maddala : *Limited-dependent and qualitative variables in econometrics*. Cambridge University Press, 1986.

[107] C.F. Manski et D. McFadden : *Structural analysis of discrete data with econometric applications*. MIT Press Cambridge, 1981.

[108] D. McFadden : A Comment on Discriminant Analysis 'Versus' Logit Analysis. *Annals of Economic and Social Measurement*, 5(4):511–523, 1976.

[109] O. Morgenstern et J. Von Neumann : *Theory of games and economic behavior*. Princeton University Press Princeton, NJ, 1947.

[110] J. Mossin : Aspects of rational insurance purchasing. *The Journal of Political Economy*, 76(4):553–568, 1968.

[111] J. Pinquet, G. Dionne, C. Vanasse et M. Maurice : Point-record incentives, asymmetric information and dynamic data. Working Papers hal-00243056, HAL, 2007.

[112] J. W. Pratt : Risk Aversion in the Small and in the Large. *Econometrica*, 32(1/2):122–136, 1964.

[113] R. Puelz et A. Snow : Evidence on adverse selection : Equilibrium signaling and cross-subsidization in the insurance market. *Journal of Political Economy*, 102(2):236–257, 1994.

[114] D. Richaudeau : Automobile insurance contracts and risk of accident : An empirical test using French individual data. *The GENEVA Papers on Risk and Insurance-Theory*, 24(1):97–114, 1999.

[115] J.G. Riley : Informational equilibrium. *Econometrica : Journal of the Econometric Society*, 47(2):331–359, 1979.

[116] M. Rothschild et J. Stiglitz : Equilibrium in competitive insurance markets : An essay on the economics of imperfect information. *The Quarterly Journal of Economics*, 90(4):629–649, 1976.

[117] K. Saito : Testing for asymmetric information in the automobile insurance market under rate regulation. *Journal of Risk and Insurance*, 73(2):335–356, 2006.

[118] B. Salanié : *The economics of contracts : a primer*. The MIT Press, 1997.

[119] G. Saporta : *Probabilités, analyse des données et statistique*. Editions Technip, 2006.

[120] L.J. Savage : *The Foundations of Statistics*. New York : John Wiley and Sons, 1954.

[121] S.H. Seog : *The Economics of Risk and Insurance*. Wiley-Blackwell, 2010.

[122] S. Shavell : On moral hazard and insurance. *The Quarterly Journal of Economics*, 93(4):541–562, 1979.

[123] M. Spence : Product differentiation and performance in insurance markets. *Journal of Public Economics*, 10(3):427–447, 1978.

[124] P. Srivastava et X. Zhao : Impact of Private Health Insurance on the Choice of Public versus Private Hospital Services. *Health, Econometrics and Data Group (HEDG) Working Papers*, 2008.

[125] J. Stewart : The welfare implications of moral hazard and adverse selection in competitive insurance markets. *Economic Inquiry*, 32(2):193–208, 1994.

[126] S. Tufféry : *Data mining et statistique décisionnelle : l'intelligence des données*. Editions TECHNIP, 2007.

[127] O.A. Vasechko, M.O. Albizzati et M. Grun-Rehomme : Les jeunes conducteurs : surprimes ou fidélisation ? *Assurance et gestion des risques, Montréal*, 74(4):109–127, 2009.

[128] C. Wilson : A model of insurance markets with incomplete information* 1. *Journal of Economic Theory*, 16(2):167–207, 1977.

[129] J.M. Wooldridge : *Econometric analysis of cross section and panel data*. The MIT press, 2002.

[130] G. Young, E.A. Valdez et R. Kohn : Multivariate probit models for conditional claim-types. *Insurance : Mathematics and Economics*, 44(2): 214–228, 2009.