

## Choix de contrat et sinistralité chez les jeunes conducteurs

Michel Grun-Réhomme and Nouredine Benlagha

Volume 74, Number 4, 2007

URI: <https://id.erudit.org/iderudit/1091737ar>

DOI: <https://doi.org/10.7202/1091737ar>

[See table of contents](#)

Publisher(s)

Faculté des sciences de l'administration, Université Laval

ISSN

1705-7299 (print)

2371-4913 (digital)

[Explore this journal](#)

Cite this article

Grun-Réhomme, M. & Benlagha, N. (2007). Choix de contrat et sinistralité chez les jeunes conducteurs. *Assurances et gestion des risques / Insurance and Risk Management*, 74(4), 505–532. <https://doi.org/10.7202/1091737ar>

Article abstract

In this relationship between the insurer and the policy-holder, between a contract of guarantees and a remuneration (premium or contribution), the insurance company faces up to a risk that is straight connects the information asymmetry which exists between it and the policy-holder.

In a recent paper, A. Cohen (2005) shows that adverse selection is present with young drivers because they have a lack of experience, results which agree with the findings of Chiappori and Salanié (2000).

The study presented in this paper comes within the framework of this last work. There is the equally question on the pertinence of the links existing between the accidents and the choice of contract of the young drivers. Our study is new on two points:

- It concerns 4 types of guarantees (but no two), that is nearer to the reality in comparison to using a bivariate model.

- These data are recent. They concern the year 2004, after the implementation of new security measures at the circulation: automatic radars, more strict control of blood alcohol content and more severe punishment for infractions. In addition, this new environment has brought about a decline in mortality from car accidents.

In this modelling, the characteristics that explain the accidents and the choice of guarantee are correlated. In this case, the autonomous estimation of the equation of accidents can include an endogenous bias. The individual characteristics of the young drivers that explain the guarantee choice, explain also positively the probability of accidents. The hypothesis of adverse selection is verified among the young drivers who choose a contract "all risks" with a variable deductible.

*Assurances et gestion des risques*, vol. 74(4), janvier 2007, 505-532  
*Insurance and Risk Management*, vol. 74(4), January 2007, 505-532

## **Choix de contrat et sinistralité chez les jeunes conducteurs**

**par Michel Grun-Réhomme et Noureddine Benlagha**

### **RESUME**

Dans les rapports entre l'assureur et l'assuré, entre un contrat de garanties et une rémunération (prime ou cotisation), la compagnie d'assurance fait face à un risque qui est directement lié à l'asymétrie d'information qui existe entre elle et l'assuré.

Dans un article récent, A. Cohen (2005) montre, conformément aux résultats de Chiappori et Salanié (2000), que l'hypothèse de sélection adverse est mise en défaut chez les jeunes conducteurs. Les jeunes ont une perception imprécise (absence d'expérience) de leurs risques.

L'étude présentée dans cet article s'inscrit dans la suite de ces travaux. On s'est également interrogé sur la pertinence des liens existants entre la sinistralité et le choix de contrat chez les jeunes conducteurs. Notre étude est nouvelle sur deux points :

- Elle porte sur quatre types de garanties (et non deux), ce qui est plus proche de la réalité, en utilisant un modèle polytomique ordonné bivarié.
- Ces données sont récentes. Elles concernent l'année 2004, après la mise en place de nouvelles mesures sécuritaires au niveau de la circulation (radars automatiques, contrôles plus stricts de l'alcoolémie, alourdissement des peines), et la baisse des accidents mortels.

Dans cette modélisation, les caractéristiques qui expliquent la sinistralité et le choix de garantie sont corrélées. Dans ce cas, l'estimation autonome de l'équation de la sinistralité peut comporter un biais d'endogénéité. Les caractéristiques individuelles des jeunes conducteurs, qui expliquent le choix de garantie, expliquent aussi positivement la probabilité de sinistralité. L'hypothèse de sélection adverse est toutefois vérifiée parmi les jeunes conducteurs qui choisissent un contrat « tous risques » variable selon le montant de la franchise.

**Mots clés** : Assurance automobile, sélection adverse, choix de garantie, sinistralité, modèles bivariés.

---

### **Les auteurs :**

Université Paris 2, ERMES-UMR7017-CNRS, 92 rue d'Assas, 75006 Paris, France. courriel : [grun@u-paris2.fr](mailto:grun@u-paris2.fr) et [blnouri2002@yahoo.fr](mailto:blnouri2002@yahoo.fr)

In this relationship between the insurer and the policy-holder, between a contract of guarantees and a remuneration (premium or contribution), the insurance company faces up to a risk that is straight connects the information asymmetry which exists between it and the policy-holder.

In a recent paper, A. Cohen (2005) shows that adverse selection is present with young drivers because they have a lack of experience, results which agree with the findings of Chiappori and Salanié (2000).

The study presented in this paper comes within the framework of this last work. There is the equally question on the pertinence of the links existing between the accidents and the choice of contract of the young drivers. Our study is new on two points:

- It concerns 4 types of guarantees (but no two), that is nearer to the reality in comparison to using a bivariate model.
- These data are recent. They concern the year 2004, after the implementation of new security measures at the circulation: automatic radars, more strict control of blood alcohol content and more severe punishment for infractions. In addition, this new environment has brought about a decline in mortality from car accidents.

In this modelling, the characteristics that explain the accidents and the choice of guarantee are correlated. In this case, the autonomous estimation of the equation of accidents can include an endogenous bias. The individual characteristics of the young drivers that explain the guarantee choice, explain also positively the probability of accidents. The hypothesis of adverse selection is verified among the young drivers who choose a contract "all risks" with a variable deductible.

**Keywords:** Automobile insurance, adverse selection, choices of guarantee, accident, damage, bivariate models.

## **I. INTRODUCTION**

Si effectivement l'assurance automobile est obligatoire pour couvrir les dégâts causés aux autres (Responsabilité Civile), le conducteur peut choisir de compléter cette garantie de base en s'assurant pour les dégâts causés à son propre véhicule selon la cause (vol, incendie, ...) et l'étendue des dommages (niveau de franchise). L'existence, chez les assurés, de différents degrés d'aversion pour le risque permet aux assureurs de proposer ces divers contrats.

Dans cet échange entre l'assureur et l'assuré, entre un contrat de garanties et une rémunération (prime ou cotisation), la compagnie d'assurance fait face à un risque qui est directement lié à l'asymétrie d'information qui existe entre elle et l'assuré. En effet, l'information sur les risques n'étant pas partagée, ceci compromet l'optimalité de l'échange.

L'assureur dispose pourtant d'informations légales (Coefficient Réduction Majoration, sinistralité sur les trois dernières années) et

demande des informations au conducteur sur lui-même et son véhicule qui lui permettent d'appréhender la réalité statistique du risque. Il a donc des informations mesurables sur le risque, mais il lui manque des informations subjectives ou non mesurables (nombre de kilomètres parcourus, infractions au code de la route, état du véhicule, type de conduite, ...).

Rappelons que l'hypothèse de sélection adverse en économie stipule que les assurés à hauts risques ont tendance à choisir la garantie la plus élevée en terme de couverture (Rothschild-Stiglitz (1976), Wilson (1977), Spence (1978)). Et le risque moral est, par définition, le risque supplémentaire que crée lui-même un assuré en diminuant sa prévention du risque fondamental parce que, justement, il est assuré.

Trois études initiales empiriques ont montré l'existence d'un phénomène de sélection adverse, mais leurs résultats ont été fortement critiqués dans des études postérieures.

Les deux premières études de Dahlby (1983, 1992) présentent une faiblesse au niveau des données utilisées. En effet Dahlby, au cours de ces deux études, ne disposait pas des données sur la couverture des risques.

Puelz et Snow (1994) ont utilisé des données individuelles obtenues des compagnies d'assurance de l'état de Georgie, mais d'autres études ont mis en question leurs résultats. Puelz et Snow n'avaient pas d'informations sur les risques de l'assuré, telle que l'expérience du conducteur ou l'historique des coûts des sinistres.

Dionne, Gouriéroux et Vanasse (2001) ont formulé une deuxième objection aux travaux de Puelz-Snow. Ils montrent que la classification faite par les assureurs est suffisante (dans le sens où il n'existe pas de sélection adverse résiduelle dans chaque classe de risque des portefeuilles des assureurs), mais par contre des effets non linéaires existent qui doivent donc être pris en compte.

D'autres études récentes empiriques (Dionne, Doherty et Fombaron (2001), Dionne, Gouriéroux et Vanasse (2001)) sur le marché des assurances automobiles montrent l'existence de preuves en contradiction avec la théorie de sélection adverse.

On s'était déjà interrogé (Grun-Réhomme et Joly, 2003) sur cette question de savoir si le choix de la formule de garantie par l'assuré révèle une information sur les risques que ce dernier fait courir à l'assureur (et à lui-même) et sur son aversion au risque dans cette situation d'échange en information asymétrique (risque moral et sélection adverse).

À partir des données de l'ensemble du portefeuille d'une mutuelle française (3,3 millions d'observations), on avait montré que l'assuré, à travers le choix de la formule de garanties traduit principalement son aversion à la perte de l'investissement qu'il vient de réaliser. Il y est sensible en valeur absolue et non en termes de perte patrimoniale. L'aversion au risque de l'individu correspond non pas à l'espérance mathématique de la perte, mais au niveau de perte maximale. L'hypothèse faite souvent par la théorie économique, à savoir que le risque de perte du véhicule est relativisé par son poids dans le patrimoine global de l'assuré, est mise en défaut.

Il faut toutefois noter que les assurés ont conscience de leur risque objectif puisque, à valeur de véhicule constante, ils cherchent à se couvrir plus fortement lorsque ce dernier augmente. Par contre, le CRM, censé représenter l'expérience de conduite, joue de façon inverse : les assurés les moins risqués (ou les plus expérimentés), à valeur de véhicule constante, privilégient une forte couverture, à l'inverse des moins expérimentés qui souscrivent des formules moins protectrices. On retrouvait aussi les résultats de Dionne, Doherty et Fombaron (2001), à savoir que les bas risques ont plus d'aversion pour le risque que les hauts risques, ainsi que les résultats de Dionne, Gouriéroux et Vanasse (2001) sur l'absence de sélection adverse résiduelle dans les classes de risque.

Cette hypothèse de sélection adverse est pourtant vérifiée à travers plusieurs études empiriques dans d'autres domaines que l'assurance automobile; en assurance santé, Cutler et Zeckhauser (2000) et en assurance vie, Finkietien et Poterba (2000).

Dans un article récent, A. Cohen (2005) s'est intéressé à la relation qui pourrait exister entre le choix du contrat d'assurance et la sinistralité chez les jeunes conducteurs ayant une ancienneté de permis inférieure à trois ans. Les données utilisées sont celles d'une compagnie d'assurance privée Israélienne; il s'agit d'informations sur 104639 conducteurs ayant souscrit une police d'assurance dans la période qui s'étale entre les années 1995 et 1999. A. Cohen a testé le problème des effets non linéaires, pour appréhender au mieux la question de sélection adverse et il montre que l'hypothèse de sélection adverse est mise en défaut chez les jeunes conducteurs. Cette conclusion est conforme aux résultats de Chiappori et Salanié (2000) qui modélisent la relation entre le choix de la franchise et le type de risque par un modèle Probit bivarié. Les jeunes ont une perception imprécise (absence d'expérience) de leurs risques.

Notre étude présentée dans cet article s'inscrit dans la suite des travaux de Chiappori, Salanié et Cohen. On s'est également interrogé

sur la pertinence du lien existant entre la sinistralité et le choix de contrat chez les jeunes conducteurs. On a aussi considéré comme jeune conducteur tout assuré dont l'ancienneté de permis était inférieure ou égale à trois ans.

Par rapport à l'article de Chiappori et Salanié, notre étude est nouvelle sur deux points :

- Elle porte sur quatre types de garanties (et non deux),
- Ces données récentes concernent l'année 2004, après la mise en place de nouvelles mesures sécuritaires au niveau de la circulation (radars automatiques, contrôles plus stricts de l'alcoolémie, alourdissement des peines).

Et contrairement à l'étude de Cohen, l'assureur français dispose d'informations sur le passé de l'assuré à travers le bonus-malus ou Coefficient Réduction Majoration (CRM).

La base de données est présentée dans la section 2. Les résultats sont élaborés en trois étapes : la première d'analyse exploratoire repose sur des analyses univariées et bivariées afin de dégager quelques tendances (section 3). La deuxième partie économétrique est classique et repose sur des modèles explicatifs univariés (section 4). La troisième partie recherche des effets non linéaires dans les modèles explicatifs de la sinistralité (section 5). Enfin une conclusion termine cette présentation. Les analyses sont réalisées avec le logiciel SAS.

Nous remercions la mutuelle qui nous a fourni les données et qui a accepté la publication de ces résultats. Toutefois les imprécisions numériques qui demeurent sont nécessaires pour garder une certaine confidentialité des données. En particulier, nous sommes restés discrets sur le montant des sinistres et le zonier (zone de garage habituel).

## **2. LES DONNÉES**

Rappelons qu'en France, on distingue deux catégories d'assurances automobiles, les sociétés d'assurances regroupées au sein de la FFSA (Fédération Française des Sociétés d'Assurance) et les mutuelles regroupées dans le GEMA (Groupement des Entreprises Mutuelles d'Assurance).

## 2.1 Les contrats

Notre mutuelle propose 4 formules pour l'assurance d'un véhicule 4 roues de tourisme :

- 1- Responsabilité Civile (RC, assurance minimale obligatoire); sont inclus dans cette formule des garanties Défense-recours – Attentats – Catastrophes Naturelles – Corporel du conducteur – Assistance.
- 2- Dommages au véhicule (DV1) : RC + Garantie Dommage au véhicule toutes causes avec une franchise importante (par exemple, 600 Euros pour une Renault Clio et 830 pour une Renault Laguna).
- 3- DV2 : RC + Garantie Dommage au véhicule toutes causes avec une franchise moyenne (200 Euros pour une Renault Clio et 250 pour la Laguna).
- 4- DV3 : RC + Garantie Dommage au véhicule toutes causes avec une franchise faible (70 Euros pour une Renault Clio et 83 pour une Laguna).

Le caractère obligatoire de l'assurance, lié au risque de responsabilité civile de l'assuré, a pour avantage de limiter les effets de sélection adverse susceptibles de conduire à l'inassurabilité du risque, au cas où seuls les conducteurs certains d'être sinistrés s'assureraient.

Avec un contrat d'assurance contenant une franchise (DV1-DV3), aucune indemnité n'est versée lorsque la perte ne dépasse pas le montant de la franchise. Dans le cas contraire, l'indemnité est égale à la perte dont on déduit la franchise. Ce contrat, simple à mettre en application, est un exemple d'assurance partielle.

## 2.2 Le fichier

La base de données contient des jeunes conducteurs (moins de trois ans de permis) assurés pour des véhicules 4 roues de tourisme durant l'année 2004. Dans sa classification, la FFSA appelle « novices », les titulaires d'un permis de moins de trois ans. Le fichier contient 28 148 observations dont 15 200 hommes et 12 948 femmes. Un enregistrement correspond à un risque / sinistre garanti (si un assuré a deux véhicules, on aura deux enregistrements).

Le descriptif complet du fichier et la codification des variables se trouvent dans l'annexe 1.

Les données concernant les sociétaires sont réparties en trois groupes de variables :

– **Les caractéristiques du conducteur :**

*Sexe* : Il s'agit du sexe du conducteur principal déclaré.

*Type du conducteur*, il exprime la qualification du conducteur principal déclaré au regard du véhicule (le conducteur principal déclaré est ou non le sociétaire).

*Age du conducteur*, exprimé en années.

*Profession* : c'est la profession du conducteur, codée en 17 modalités (cf. annexe 1).

*Ancienneté de permis* : de 0 à 3 ans.

*Formule de garantie* : 4 modalités; RC, DV1-DV3.

*Coefficient Bonus-malus* : il est compris entre 0,50 et 3,50 inclus (exprimé en %), conformément à la législation française en vigueur.

*Période de couverture* : période, en mois, au cours de laquelle l'assuré est couvert par la police qu'il a souscrit, le plus souvent cette période est d'une année.

Le coefficient réduction majoration (CRM) ou coefficient bonus-malus est un mécanisme qui permet d'ajuster la prime d'assurance payée par l'assuré, en fonction de sa conduite observée au cours des périodes antérieures. Ce coefficient augmente avec les sinistres responsables et diminue avec l'absence de sinistres responsables, mais cette variable ne contient pas d'information pour les nouveaux conducteurs qui viennent juste d'avoir leur permis (et leur assurance) et seulement une information récente pour les autres nouveaux conducteurs. Pour plus de détails sur le CRM, on peut se référer à l'article de M. Grun-Réhomme (2000). Cette variable est appelée KCLAANN dans le fichier.

– **Les caractéristiques du véhicule :**

*Ancienneté de véhicule* : elle exprime le millésime de l'année du modèle du véhicule. *Puissance réelle du véhicule* : elle exprime 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). Par exemple, la Fiat Uno, la Renault Clio et la Peugeot 205 ont 60 ch. Din, et les Peugeot 206, Renault Laguna et Citroen Picasso, 110 ch. Din.



La variable sur l'usage du véhicule n'a pas été retenue puisque la quasi-totalité des sociétaires en avait un usage promenade-trajet et non un usage professionnel.

– **Les caractéristiques des sinistres :**

*Nombre de sinistre* : pour l'année de référence, ici 2004.

*Responsabilité du conducteur* : variable binaire qui indique la responsabilité du conducteur en cas de sinistre.

*Montant de dépenses cumulées* : lorsque ce montant figure pour un sinistre, il représente les dépenses cumulées depuis l'enregistrement de l'événement ou depuis sa dernière remise en cours. Il est réinitialisé lors d'une remise en cours.

*Montant de l'évaluation du sinistre* : ce montant correspond à une évaluation a priori du montant du sinistre.

Pour les accidents corporels, l'assureur peut connaître le montant total de ses dépenses seulement après plusieurs mois, voire plusieurs années, d'où cette nécessité de faire une prévision de ses dépenses. Le coût des sinistres corporels comprend plusieurs composantes : indemnités aux personnes physiques, soins, tierce personne, préjudices personnels et économiques. Les sinistres corporels représentent environ 3 % du nombre total de sinistres indemnisés, mais plus du quart du coût total des sinistres.

### **3. ANALYSE EXPLORATOIRE DES DONNÉES**

Nous présentons dans cette section, les premiers résultats significatifs sur une variable et le croisement de deux variables.

Le fichier contient 54 % d'hommes et 81 % des jeunes conducteurs ne sont pas sociétaires (en général le sociétaire est l'un des parents du jeune). La moyenne d'âge est de 22 ans et la médiane de 21 ans. 46 % des jeunes conducteurs ont entre 20 et 21 ans et 90 % moins de 30 ans. 25 % des conducteurs ont un CRM supérieur à 100 et un quart inférieur à 90.

L'ancienneté des véhicules est en moyenne de 9 ans, dont 25 % moins de 6 ans et 25 % plus de 13 ans. La médiane de la puissance des véhicules est de 60 ch. Din (moyenne de 66), 25 % des véhicules ont une puissance inférieure à 58 et 25 % une puissance supérieure à 75.

L'annexe 2 présente un tableau des statistiques descriptives des différentes variables qui seront par la suite retenues dans les modèles.

8,5 % des jeunes conducteurs ont eu au moins un sinistre responsable (partiel ou total) dans l'année (les autres n'ont pas eu de sinistre, 80 %, ou au moins un sinistre non responsable sans sinistre responsable, 11,5 %). Il s'agit bien sûr des sinistres déclarés à l'assureur. On sait que souvent des automobilistes ne déclarent pas un petit accrochage pour éviter d'être pénalisés au niveau du coefficient bonus-malus ou lorsque le montant du sinistre est inférieur à la franchise.

### Le choix de contrat

Le tableau suivant présente la répartition de la population des jeunes conducteurs selon la formule de garantie choisie et selon l'ancienneté de permis :

Garantie	%	Ancienneté de permis	%
RC	13	0	12
DV1	36	1	25
DV2	18	2	30
DV3	33	3	33

Soit  $Y$  la variable binaire qui mesure la sinistralité observée en 2004 et qui est définie par :

$$Y_i = \begin{cases} 0 & \text{si l'assuré } i \text{ n'a pas de sinistre responsable} \\ 1 & \text{si l'assuré } i \text{ a au moins un sinistre responsable} \end{cases}$$

On observe une indépendance de cette variable  $Y$  avec l'âge du conducteur, le sexe et la profession.

À 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. Cette variable est donc à manier avec précaution.

Pour le croisement de  $Y$  avec le contrat et l'ancienneté de permis, on obtient les résultats suivants, exprimés en pourcentage :

Contrat	$Y$		Anc_ permis	$Y$	
	$Y = 0$	$Y = 1$		$Y = 0$	$Y = 1$
RC	68	32	0	97	3
DV1	92	8	1	86	14
DV2	79	21	2	74	26
DV3	74	26	3	77	23

On peut constater dans cette phase exploratoire que la sinistralité est plus importante pour les sociétaires qui choisissent un contrat RC (garantie minimale), mais parmi ceux qui choisissent un contrat DV, les hauts risques prennent une garantie plus grande. Mais ces premiers résultats exploratoires doivent être confirmés ou infirmés à travers des approches multidimensionnelles.

#### 4. MODÈLES ÉCONOMÉTRIQUES UNIVARIÉS

Une régression logistique a été utilisée pour expliquer la variable binaire de sinistralité (pas de sinistre responsable, modalité de référence, ou au moins un sinistre responsable), une loi de Poisson pour la variable de comptage du nombre de sinistres et un modèle log-linéaire pour le montant cumulé des coûts des sinistres. Dans un article récent, Melgar M.C., Ordaz Sanz J.A., Guerrero F.M. (2005) montrent que différents types de modèles de comptage donnent des résultats qualitativement semblables.

Toutes ces modélisations vont dans le même sens. Les coûts des sinistres sont en moyenne plus élevés chez les hommes que les femmes, mais la différence n'est pas très importante; elle est de l'ordre de 10 %. La variabilité des sinistres est aussi plus importante chez les hommes que les femmes, cela tient essentiellement à la présence de quelques sinistres graves (valeurs extrêmes) chez les hommes. Le coût moyen des sinistres responsables est sensiblement le même selon la formule de garantie choisie. Le modèle log linéaire sur les coûts n'est pas de bonne qualité.

Le modèle Logit, par rapport au modèle Probit, facilite l'interprétation des paramètres associés aux variables explicatives, même s'il tend à attribuer aux événements « rares » une probabilité plus forte que la loi normale.

Pour la régression logistique, la sélection des variables explicatives et l'ajustement du modèle aux données sont effectués sur 75 % de la population des jeunes et la validation sur un échantillon distinct. Les coefficients du modèle, obtenus sur le premier échantillon, sont utilisés dans le deuxième de manière à calculer des probabilités a posteriori.

Seuls les résultats nécessaires ont été gardés. Par exemple, seule la dernière étape de la méthode stepwise a été conservée.

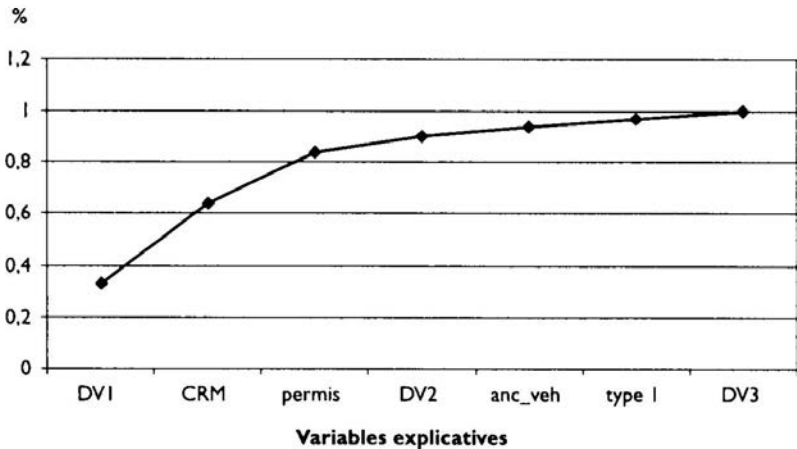
Résultats de la modélisation avec quatre types de contrats :

Paramètres	DF	Esti- mation	Ecart type	Chi- Square	Pr > ChiSq
Constante	1	5,6952	0,5243	117,9968	< ,0001
Type1 (*)	1	0,4769	0,1728	7,6139	0,0058
DV1	1	2,1692	0,2553	72,1983	< ,0001
DV2	1	0,7975	0,2161	13,6154	0,0002
DV3	1	0,4049	0,1771	5,2264	0,0222
Ancienneté du véhicule	1	-0,0485	0,0169	8,2779	0,0040
Ancienneté de permis	1	-0,4651	0,0700	44,1895	< ,0001
Bonus-Malus	1	-0,0400	0,00484	68,3770	< ,0001

(\*) Le conducteur est le sociétaire

Au niveau de la qualité du modèle, on trouve 75 % de paires concordantes. Au seuil de 91,5 % (qui correspond à la probabilité de  $Y = 0$ ), la sensibilité est égale à 44 % (pourcentage de vrais valeurs  $Y = 0$  estimées par le modèle) et 88 % pour la spécificité (pourcentage de vrais valeurs  $Y = 1$  estimées par le modèle). L'annexe 3 présente la courbe ROC du modèle et le calcul des effets marginaux des variables explicatives.

Toutes choses égales par ailleurs, la sinistralité augmente avec l'ancienneté de permis, l'ancienneté du véhicule et le CRM (Bonus-Malus), et elle diminue lorsque le conducteur est le sociétaire. On peut hiérarchiser les variables exogènes selon leur pouvoir explicatif dans la modélisation à l'aide du rapport de vraisemblance, comme le montre le graphique suivant.



Si on fait une analyse séparée pour les deux types de conducteur (le sociétaire est ou n'est pas le conducteur), on trouve des résultats semblables, avec trois variables significatives qui sont dans l'ordre, le contrat DV1, l'ancienneté de permis et le CRM.

L'analyse de Chiappori et Salanié (2000) faite sur des données françaises, ne prend en compte que deux types de garantie. Dans notre étude, le modèle est meilleur en prenant en considération les 4 types de garanties, aussi bien au niveau des critères d'information que des paires concordantes (75 % vs 70 %) (cf annexe 4).

La modélisation confirme l'analyse exploratoire.

Il est intéressant de noter que si l'on considère seulement deux garanties RC et DV (en regroupant en une seule modalité les contrats « tous risques » quel que soit le niveau de franchise), on vérifie, comme dans l'article de Chiappori et Salanié (2000), l'absence de sélection adverse. Mais parmi ceux qui ont choisit un contrat DV, la sinistralité augmente avec la couverture (avec la diminution de la franchise).

Le tableau suivant présente les résultats de l'estimation des paramètres du modèle Logit avec seulement deux types de contrats.

Le modèle présente une moins bonne qualité quand on ne considère que deux types de contrat (cf. annexe 4) et la hiérarchisation des variables dans le modèle change.

Paramètres	DF	Esti- mation	Ecart type	Chi- Square	Pr > ChiSq
Constante	1	5,8690	0,5134	130,7060	< ,0001
Type I (*)	1	0,5384	0,1695	10,0873	0,0015
DV	1	0,7185	0,1706	17,7353	<,0001
Ancienneté du véhicule	1	0,0335	0,0136	6,1136	0,0134
Ancienneté de permis	1	-0,4767	0,0683	48,6553	< ,0001
Bonus-Malus	1	-0,0391	0,00472	68,4433	< ,0001

(\*) Le conducteur est le sociétaire

## 5. MODÈLES BIVARIÉS

Deux études comparatives sont menées, l'une avec un modèle probit bivarié (et en ne considérant que deux types de garantie) et l'autre avec un modèle polytomique ordonné bivarié (avec quatre types de garantie).

Des modèles bivariés ou multivariés ont été utilisés récemment dans de nombreux domaines. On peut citer, par exemple, pour le modèle probit bivarié, des applications dans le domaine de l'éducation, Greene (1998) ou dans le champ social, Jaenicke (2003).

Pour le modèle ordinal bivarié, on peut consulter les papiers de Aitchison et Sisley (1957), Ashford et Sowden (1970), McKelvey et Zavoina (1975), McCullagh (1980), les ouvrages de Collett (1991), Agresti (2002) et pour une approche bayésienne, Albert (1993) ou O'Brien et Dunson (2004). On trouve les premières applications en biologie, Gurland, Lee et Dahm (1960). On trouve également des articles en médecine, Kim (1995), Todem, Kim et Lesaffre (2002) ou en épidémiologie, Biswas et Das (2002). Curieusement ces modèles d'équations simultanées sont encore peu utilisés dans le domaine économique.

### 5.1 Modèle probit bivarié

D'abord, comme dans les études empiriques de Chiappori et Salanié, on étudie un modèle bivarié où  $Y$  est la variable endogène de sinistralité et l'autre variable endogène  $Z$  correspond au choix de contrat avec dans le cadre d'un modèle probit bivarié :

$$Z_i = \begin{cases} 0 & \text{si l'assuré } i \text{ a un contrat } RC \\ 1 & \text{si l'assuré } i \text{ a un contrat } DV \end{cases}$$

Le modèle s'écrit :

$$Z^* = \alpha_1 X_1 + \varepsilon$$

$$Y^* = \alpha_2 X_2 + \beta Z + \varepsilon$$

où  $Y^*$  et  $Z^*$  sont deux variables latentes continues, pour lesquelles on observe  $Y = I(Y^* > 0)$  et  $Z = I(Z^* > 0)$ ,  $X_1$  et  $X_2$  correspondent aux caractéristiques de l'assuré (les variables ou modalités significatives ne sont pas nécessairement les mêmes dans les deux équations).

Dans le modèle, les erreurs sont supposées corrélées :

$$\begin{pmatrix} \varepsilon \\ e \end{pmatrix} \rightarrow N \left[ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \rho \sigma_1 \sigma_2 \\ \rho \sigma_1 \sigma_2 & \sigma_2^2 \end{pmatrix} \right]$$

Dans ce problème, les paramètres peuvent être déterminés à une constante multiplicative près. On impose alors aux termes d'erreur d'avoir une variance unitaire ( $\sigma_1 = \sigma_2 = 1$ ). Les paramètres du modèle sont estimés à l'aide d'un algorithme de Newton-Raphson.

Cette modélisation est utilisée pour déterminer ce qui des caractéristiques de l'assuré ou du véhicule et du choix de garantie est le plus déterminant dans la probabilité de sinistre ou de non sinistre. Un modèle bivarié qui estime simultanément la sinistralité et le choix de garantie permet de tester l'endogénéité du choix de garantie sur la sinistralité et donne des résultats robustes quant à l'estimation des probabilités de sinistre. Le choix de garantie pourrait aussi refléter l'influence de caractéristiques individuelles inobservées.

Le tableau suivant présente les estimations des paramètres du modèle probit bivarié :

Paramètres	Estimation	Ecart type	t value	Pr >  t
z.constante	-1,093115	0,393875	-2,78	0,0055
z.Puissance du véhicule	0,009206	0,002757	3,34	0,0008
z.Ancienneté du véhicule	-0,131453	0,010773	-12,20	< ,0001
z.Ancienneté de permis	-0,111785	0,040809	-2,74	0,0062
z.Bonus-malus	-0,008132	0,003270	-2,49	0,0129
y.constante	-3,137948	0,335480	-9,35	< ,0001
y.z	-0,495393	0,253975	-1,97	0,0491
y.Type I (*)	-0,304652	0,094734	-3,22	0,0013
y.Ancienneté du permis	0,281061	0,038689	7,26	< ,0001
y.Bonus-malus	0,022719	0,002685	8,46	< ,0001
Rho	-0,002307	0,145850	-0,02	0,9874
(*) le conducteur est le sociétaire				

On constate que l'assuré augmente sa couverture avec la puissance du véhicule et la diminue avec l'ancienneté du véhicule, l'ancienneté de permis et le CRM. Le souscripteur diminue sa couverture avec l'ancienneté de son permis; on peut penser que le jeune conducteur prend de l'assurance après ses premières expériences de conduite. Dans ce cas, son degré d'aversion au risque diminue, le conducteur sous estime ses risques et il choisit donc des garanties plus faibles.

Pour une gamme donnée de véhicules, le prix augmente en général avec la puissance et donc le propriétaire d'un véhicule puissant est plus prévoyant pour protéger la valeur de son véhicule, d'autant plus qu'un jeune a un patrimoine relativement faible. Et à l'inverse, il lui paraît inutile de prendre une forte couverture pour un véhicule ancien qui a une faible valeur.

Le coefficient  $\rho$  de corrélation entre les erreurs des deux équations, c'est-à-dire la corrélation entre les variables observées et inobservées des deux situations est ici non significative. Les deux équations pouvaient être estimées séparément.



## 5.2 Modèle ordinal bivarié

Dans le cadre du modèle polytomique ordonné, la variable endogène  $Z$  est définie par :

$$Z_i = \begin{cases} 0 & \text{si l'assuré } i \text{ a un contrat } RC \\ 1 & \text{si l'assuré } i \text{ a un contrat } DV1 \\ 2 & \text{si l'assuré } i \text{ a un contrat } DV2 \\ 3 & \text{si l'assuré } i \text{ a un contrat } DV3 \end{cases}$$

Le modèle se présente sous la même forme que le modèle probit bivarié, mais la variable latente continue  $Z^*$  associée à la variable ordinaire  $Z$  est alors définie de la façon suivante :

$$Z = \begin{cases} 0 & \text{si } Z^* \leq l_1 \\ 1 & \text{si } l_1 < Z^* \leq l_2 \\ 2 & \text{si } l_2 < Z^* \leq l_3 \\ 3 & \text{si } Z^* < l_3 \end{cases}$$

Dans ce modèle ordinal bivarié, les paramètres et les limites sont inconnus. Pour s'assurer que les paramètres sont identifiables, il est nécessaire d'imposer, sans perte de généralité, que la première limite  $l_1$  est nulle, et dans ce cas la constante est conservée dans le modèle. C'est l'option utilisée par défaut dans la procédure Qlim de SAS. On peut aussi estimer cette limite  $l_1$  et enlever la constante (avec l'option de Model, limit1 = varying).

Les paramètres et les limites peuvent être estimés par une méthode de Newton-Raphson ou par des méthodes bayésiennes, Albert et Chib (1993), Zhang, Strawderman, Cowen et Wells (2006).

La probabilité que  $z_i^*$  appartienne à la  $j^{\text{ième}}$  modalité de  $Z$  est définie par :

$$P[l_{j-1} < z_i^* < l_j] = F(l_j - x_i' \beta) - F(l_{j-1} - x_i' \beta)$$

où  $F(\cdot)$  est la fonction de répartition de la loi logit ou de la loi normale standard. Dans cette présentation,  $l_0 = -\infty$  et  $l_4 = +\infty$ .

Dans le cas polytomique ordonné, le modèle est basé sur les probabilités cumulées des modalités de  $Z$  et non sur les probabilités de réalisation d'une modalité de  $Z$ .

À l'exception de la puissance du véhicule et de l'ancienneté de permis pour  $Z$ , les variables explicatives des deux équations restent

les mêmes. Le CRM n'est pas significatif au seuil de 5 % dans la première équation où la formule de garantie est endogène. À ce propos il faut rappeler que les 2/3 des jeunes conducteurs ont un CRM compris entre 90 et 100. Les jeunes conducteurs choisissent une couverture plus faible quand le véhicule est plus ancien, traduisant ainsi une sensibilité à la perte de l'investissement réalisé, ce qui est conforme aux résultats trouvés antérieurement sur l'ensemble du portefeuille (Grun-Rehomme, Joly, 2003).

Le tableau suivant présente les estimations des paramètres du modèle ordinal bivarié :

Paramètres	Esti- mation	Ecart type	t value	Pr >  t
cforasu <sup>(*)</sup> .constante	1,197113	0,213101	5,62	< ,0001
cforasu.Ancienneté du véhicule	-0,047540	0,005081	-9,36	< ,0001
cforasu.Bonus-malus	0,004054	0,002126	1,91	0,0566
Limit2.cforasu	1,089948	0,036973	29,48	< ,0001
Limit3.cforasu	1,593508	0,041834	38,09	< ,0001
y.constante	-2,082383	0,509587	-4,09	< ,0001
y.DV1	-1,576471	0,171888	-9,17	< ,0001
y.DV2	-1,286077	0,291437	-4,41	< ,0001
y.DV3	-1,589008	0,431940	-3,68	0,0002
y.Type1	-0,255349	0,090541	-2,82	0,0048
y.Ancienneté de permis	0,241538	0,039778	6,07	< ,0001
y.Bonus-malus	0,022188	0,002767	8,02	< ,0001
Rho	0,494946	0,155233	3,19	0,0014

(\*) Formule d'assurance à quatre modalités

Mais dans cette modélisation, le coefficient  $\rho$  de corrélation est significatif, on peut donc en conclure que les caractéristiques qui expliquent la sinistralité et le choix de garantie sont corrélées. Dans ce cas, l'estimation autonome de la seconde équation peut comporter un biais d'endogénéité (cf. Lollivier, 2001). Et comme  $\rho$  est positif, les caractéristiques individuelles des jeunes conducteurs, qui expliquent le choix de garantie, expliquent aussi positivement la probabilité de

sinistralité. On met donc bien en évidence une corrélation entre le choix de garantie et la sinistralité, à condition de considérer les quatre formules de garantie. Le modèle bivarié permet de mesurer l'effet réel (et non biaisé) des caractéristiques des assurés sur la sinistralité. L'effet d'endogénéité de  $Z$  se traduit également par des estimations différentes des coefficients de cette variable dans les modèles univariés et bivariés.

La sinistralité augmente avec l'ancienneté de permis et le CRM et elle diminue lorsque le conducteur est le sociétaire et avec les différentes formules de garantie « tous risques »

L'hypothèse de sélection adverse est vérifiée parmi la sous population des jeunes conducteurs qui choisissent un contrat DV. À ce propos, il faut noter que la garantie DV1 (tous risques avec la franchise la plus élevée) est atypique dans l'ensemble du portefeuille. Elle est souscrite essentiellement par deux catégories d'assurés, ceux qui possèdent une voiture récente et puissante (on en trouve peu chez les jeunes conducteurs) et d'autre part ceux qui ont une voiture ancienne. 74 % des jeunes conducteurs qui ont un véhicule ancien (plus de 12 ans) choisissent cette garantie DV1. Ils représentent 23 % de la population de ces jeunes.

Un phénomène d'aléa moral peut également expliquer en partie ce résultat. Chiappori et Salanié (2000) avaient également laissé de côté cette question de la distinction entre sélection adverse et aléa moral. Il est effectivement difficile de faire la différence entre ces deux aspects avec des données statiques. Des données de panel sont nécessaires.

Les limites de seuil estimées par ce modèle structurel pour les garanties s'interprètent ainsi :

$$Z = \begin{cases} 0 & \text{si } Z^* \leq 0 \\ 1 & \text{si } 0 < Z^* \leq 1,09 \\ 2 & \text{si } 1,09 < Z^* \leq 1,59 \\ 3 & \text{si } Z^* > 1,59 \end{cases}$$

L'étendue d'un intervalle obtenu à partir des limites estimées dans la modélisation reflète la part de la population qui a la modalité correspondante à cet intervalle. Par exemple, l'étendue de l'intervalle pour  $Z = 1$  est de 1,09 et correspond à 36 % de la population, alors que pour la modalité 2 de  $Z$ , l'étendue est de 0,5 et correspond à 18 % des jeunes conducteurs.

Les modèles ont été estimés à l'aide de la procédure Qlim de SAS (version 9).

## 6. CONCLUSIONS ET PERSPECTIVES

L'étude statistique de la relation entre la sinistralité et le choix de garantie ou implicitement l'étude du phénomène de sélection adverse sur un marché d'assurance doit figurer parmi les priorités des services statistiques dans les compagnies et les mutuelles d'assurance.

On sait qu'en général la corrélation positive entre le risque et la couverture est d'autant plus grande que le marché est concurrentiel. Dans notre cas, cette corrélation positive entre le risque et le contrat n'est pas si élevée, ce qui s'explique en grande partie par le fait que cette mutuelle n'est que peu concurrencée dans sa niche. Conformément aux résultats que nous avons obtenus sur l'ensemble du portefeuille de cette assurance (Grun-Réhomme et Joly, 2003), cette corrélation montre que les variables dont dispose l'assureur traduisent, en partie, l'aversion pour le risque des assurés.

À travers cette étude, l'assureur peut mesurer la sinistralité des jeunes conducteurs en termes de fréquence et de coût selon la garantie choisie. Cela lui permet d'optimiser son offre de contrat et éventuellement d'inciter les nouveaux assurés à choisir un contrat plutôt qu'un autre par l'intermédiaire des gestionnaires de son réseau. Au moment de la souscription, le choix de l'assuré peut être influencé par le gestionnaire du contrat; les assurés sont d'ailleurs très demandeurs de conseils à ce moment là et dans ce cas, l'employé de l'assurance préfère souvent proposer un contrat « tous risques » afin d'éviter, en cas de sinistre, le mécontentement de l'assuré.

Cette étude met en évidence un certain nombre de variables explicatives de la sinistralité et du choix de contrat, mais on peut s'interroger sur l'existence de variables inobservées, explicatives du choix de contrat. En particulier, il peut y avoir un lien entre le niveau de cotisation et le niveau de revenu. On peut également penser à faire intervenir le patrimoine du sociétaire, à la place de son revenu, mais cette information est plus difficile à obtenir et elle doit être corrélée avec le revenu.

La variable sexe n'est pas significative pour expliquer la sinistralité chez les conducteurs novices; on peut simplement noter que le coût moyen des sinistres chez les hommes est environ 10 % plus important que pour les femmes. Ceci est une surprise, car quand on avait examiné cette question en 1998 (Grun-Réhomme) sur le portefeuille de la même mutuelle, on avait trouvé que la sinistralité des jeunes conducteurs hommes était bien plus importante que celle des jeunes femmes (en termes de fréquence et de coût). Par ailleurs dans

le cadre d'une convention passée entre l'État et les assureurs, les mutuelles consacrent 0,5 % du total de leurs primes RC annuelles à des facteurs de sécurité routière, dont la moitié à des stages de perfectionnement à la conduite pour des conducteurs novices, seniors ou accidentogènes. On peut avancer que ses différentes mesures préventives, ajoutées aux mesures répressives ont eu un effet positif sur la sinistralité des jeunes conducteurs hommes. Cette hypothèse est toutefois à confirmer.

L'autre perspective d'étude concerne le niveau supportable de la cotisation qui repose sur l'idée qu'il existe un seuil à partir duquel on ne peut plus souscrire un contrat « tous risques ». En effet, on peut souscrire un niveau de couverture élevé si on est assez riche pour payer la prime correspondante et en même temps craindre la perte financière liée à l'absence de ce niveau de couverture. Comme le montrent Dionne, Doherty et Fombaron (2001), avec des modèles dynamiques, le choix de la garantie peut dépendre du coût de ce contrat. Une enquête devrait être réalisée en ce sens.

Une macro SAS (McMillan, Hanson, 2005) qui complète la Proc Nlmixed, étant maintenant disponible, on pourrait étudier simultanément la sinistralité en fréquence et en coût et estimer les facteurs de risque à l'aide d'un modèle de régression de Dale bivarié (Dale, 1986). La dépendance entre la fréquence et le coût est alors modélisée par un modèle global de « cross-ratio » (Molenberghs, Lesaffre, 1994).

## ANNEXE I DESCRIPTION DU FICHIER

### Liste des variables disponibles dans le fichier

Variables	Nom des variables
CSEX	Code sexe
CFORASU(*)	Formule d'Assurance
QDI N	Puissance réelle du véhicule
CPROSOC	La profession du sociétaire
AGE CONDU	Âge du conducteur
AGEP	Ancienneté de permis
ANCIEN_VEH	Ancienneté du véhicule
TCDD2(**)	Type du conducteur
KCLAANN	Coefficient bonus-malus
NBSIN	Nombre de sinistres dans l'année
MDEPCUM	Montant des dépenses cumulées
MEVA2	Montant des évaluations des sinistres
RESP	Responsabilité du conducteur
(*) Contrat RC ou DV (pour deux modalités), RC et DVI-3 (pour quatre modalités)	
(**) Variable binaire codée 1 si le sociétaire est le conducteur et 0 dans le cas contraire	

**ANNEXE I (suite)**  
**DESCRIPTION DU FICHER**

**Codification des variables**

**Responsabilité du conducteur**

Code	Responsabilité de conducteur
0	Pas de sinistre
1	Responsabilité du conducteur non engagée dans le sinistre
2	Responsabilité du conducteur engagée dans le sinistre

**Profession**

Code	Profession
0	Non renseigné
1	Agriculteur, exploitant
2	Artisan (*)
3	Commerçant
4	Chef d'entreprise (*)
5	Profession libérale
6	Profession de l'information des arts et des spectacles
7	Cadre, ingénieur
8	Enseignant, formateur, chercheur
9	Educateur, animateur, moniteur
10	Profession intermédiaire de la santé et du travail social
11	Technicien
12	Contremaître, agent de maîtrise (*)
13	Agent, employé
14	Ouvrier
15	Etudiant
16	Demandeur d'emploi
17	Homme ou femme au foyer

(\*) : Certaines de ces professions ne se trouvent pas parmi les jeunes (artisan, chef d'entreprise, agent de maîtrise). Cette variable n'est pas explicative dans les modèles.

**ANNEXE 2**  
**TABLEAU DE STATISTIQUES DESCRIPTIVES**

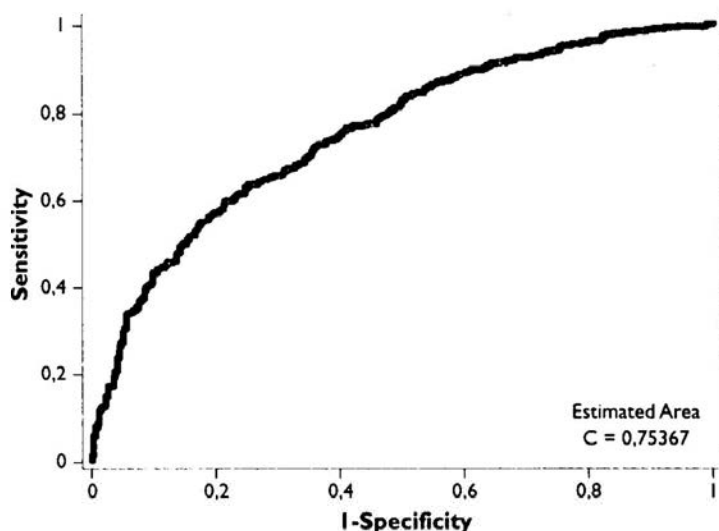
	<b>Min</b>	<b>Q<sub>1</sub></b> <b>(25 %)</b>	<b>Médiane</b>	<b>Moyenne</b>	<b>Q<sub>3</sub></b> <b>(75 %)</b>	<b>Max</b>
<b>Ancienneté du véhicule</b>	0	6	9	9,41	13	39
<b>Âge du conducteur</b>	18	20	21	22	22	63
<b>Puissance du véhicule</b>	34	57	60	66	75	142
<b>Bonus-Malus</b>	50	90	90	95,3	100	185



### ANNEXE 3 COURBE ROC ET CALCUL DES EFFETS MARGINAUX DES VARIABLES EXOGÈNES

La courbe ROC (Receiver Operating Characteristic) permet de visualiser les variations de la spécificité et de la sensibilité du modèle pour différentes valeurs du seuil de discrimination. L'aire sous la courbe est un estimateur de l'efficacité globale du modèle.

ROC Curve for the Selected Model



La courbe ROC correspond au critère  $c$  :

$$c = \frac{n_c + 0,5(t - n_c - n_d)}{t}$$

où  $n_c$  est le nombre de paires concordantes (entre la réalité et le modèle),  $n_d$  le nombre de paires discordantes,  $t$  le nombre total de paires et  $n$  l'effectif total. Si le nombre de paires ex æquo (c'est-à-dire  $t - n_c - n_d$ ) est négligeable, alors  $c$  correspond à la proportion de paires concordantes.

Les valeurs des critères d'information figurent dans l'annexe 2, ainsi que les estimations des paramètres du modèle.

Le signe des coefficients du modèle indique si la variable associée influe à la hausse ou à la baisse sur la probabilité de l'événement considéré. Rappelons que la valeur numérique d'un coefficient n'a pas d'intérêt. Toutefois il est possible de calculer l'effet marginal associé à chaque variable explicative. Si on désigne par  $X$  le vecteur des variables explicatives  $x^j$  et par  $\beta$  le vecteur des paramètres associés, l'effet marginal associé à la  $j^{\text{ème}}$  variable explicative  $x^j$  est défini par :

### ANNEXE 3 (suite)

$$\frac{\partial p_i}{\partial x_i'} = \frac{e^{-x_i \beta}}{(1 + e^{-x_i \beta})^2} \beta_i \text{ où } p_i = \text{Prob}(Y_i = 1).$$

Calcul des effets marginaux des variables explicatives de la sinistralité dans le modèle Logit

Variables	Type I	DV1	DV2	DV3
Effets marginaux	0,06374	0,28994	0,10659	0,05413
	Ancien_veh	agep	kclaann	
Effets marginaux	-,006482	-,06216	-,00534	

### ANNEXE 4 COMPARAISON DE LA QUALITÉ DES MODÈLES AVEC 2 CONTRATS ET 4 CONTRATS

Qualité du modèle selon les critères d'information

*Avec deux contrats (RC-DV)*

Critères	Constante seule	Constante et variables
AIC	1758,939	1607,783
SC	1764,432	1640,740
-2 Log L	1756,939	1595,783

*Avec quatre contrats (RC-DV1, 2, 3)*

Critères	Constante seule	Constante et variables
AIC	1758,939	1538,055
SC	1764,432	1576,997
-2 Log L	1756,939	1522,055

## Références

- Agresti (2002), *Categorical Data Analysis*, Wiley and Sons Inc., Hoboken, N.J., second edition.
- Aitchison, J. et Silvey, S.D. (1957), "Polychotomous Quantal Response by Maximum Indicant", *Biometrika*, 44, 131-140.
- Albert, J.H. et Chib, S. (1993), "Bayesian Analysis of Binary and Polychotomous Response Data", *JASA*, 88, 422, 669-679.
- Ashford, J. et Sowden, R.R. (1970), "Multivariate Probit Analysis", *Biometrics*, 26, 235-246.
- Biswas, A. et Das, K. (2002), "A bayesian analysis of bivariate ordinal data: Wisconsin epidemiologic study of diabetic retinopathy revisited", *Statistics in Medicine*, 21, 549-559.
- Chiappori, P.A. et Salanié, B. (2000), "Testing for Asymmetric Information in Insurance Markets", *Journal of Political Economy*, 108, 1, 56-78.
- Cohen, A. (2005), "Asymmetric information and learning: evidence from the automobile insurance market", *The Review of Economics and Statistics*, 87, 197-207.
- Collett (1991), *Modelling Binary Data*, Chapman and hall, London.
- Crocker, K.J. et Snow, A. (2000), *The Theory of Risk Classification, Handbook of Insurance*, Kluwer Academic Publishers, Boston, Chapter 8.
- Cutler, D.M. et Zeckhauser, R.J. (2000), *The Anatomy of Health Insurance*, In handbook of Health Economics, Newhouse, Eds.
- Dahlby, B. (1983), "Adverse Selection and Statistical Discrimination: an Analysis of Canadian Automobile Insurance", *J.P.E.*, 20, 121-130.
- Dahlby, B. (1992), *Adverse Selection and Pareto Improvements through Compulsory Insurance*, Cambridge U. Press, chap. 12.
- Dale, J. (1986), "Global cross-ratio Models for Bivariate, Discrete, Ordered Response", *Biometrics*, 42, 909-917.
- Dionne, G. et Vanasse, C. (1992), "Automobile Insurance Ratemaking in the Presence of Asymmetrical Information", *Journal of Applied Econometrics*, 7, 149-165.
- Dionne, G., Doherty, N. et Fombaron, N. (2001), Adverse Selection in Insurance Markets, Handbook of Insurance, Kluwer Academic Publishers, Boston, 185-243.
- Dionne, G., Gouriéroux, C. et Vanasse, C. (2001), "Testing for Evidence of Adverse Selection in the Automobile Insurance Market", *Journal of Political Economy*, 109, 2, 444-453.
- Finkietien, A. et Poterba, J. (2000), "Adverse Selection in Insurance Markets: Policyholder Evidence from the UK Annuity Market", *NBER Working paper*, n. 8045.
- Greene, W.H. (1998), "Gender economics courses in liberal arts colleges: further results", *J. of Economic Education*, 29, 291-300.
- Grun-Réhomme, M. (1998), « Etude de la stabilité des indicateurs de risque en assurance », *Risques, les cahiers de l'assurance*, 35, 111-119.

- Grun-Réhomme, M. (2000), « Prévion du risque et tarification : le rôle du bonus-malus français », *Revue Assurances et gestion des risques*, 1, 21-30.
- Grun-Réhomme, M. et Joly, V. (2003), « Risque individuel et choix de contrat : Le cas de l'assurance automobile », *Revue Assurances et gestion des risques*, 71-1, 145-162.
- Gurland, J., Lee, I. et Dahm, P.A. (1960), "Polychotomous Quantal Response in Biological Assay", *Biometrics*, 16, 382-398.
- Henriet, D. et Rocher, J.C. (1991), « Microéconomie de l'assurance », *Economica*, Paris, 119-140, 165-187.
- Jaenicke, J. (2003), "Bivariate probit models of social interactions : Some Monte Carlo results", *Bulletin of International Statistical Institute*, 60, 1, 543-544.
- Kim, K. (1995), "A cumulative probit regression model for ordered categorical data", *Statistics in Medicine*, 14, 1341-1352.
- Laffont, J.J. (1985), « Economie de l'incertain et de l'information », *Economica*, Paris, 127-141.
- Lollivier, S. (2001), « Endogénéité d'une variable explicative dichotomique dans le cadre d'un modèle bivarié », *Annales d'économie et de statistique*, 62, 251-269.
- MacCullagh, P. (1980), "Regression Models for ordinal Data", *JRSS, Ser. B*, 109-142.
- McKelvey, R. et Zavoina, W. (1975), "A statistical Model for the Analysis of Ordinal Level Dependent Variables", *J. of Mathematical Sociology*, 4, 103-120.
- McMillan, G. et Hanson, T. (2005), "SAS macro BDM for Fitting the Data Regression Model to Bivariate Ordinal Response Data", *J. of Statistical Software*, 14, 2.
- Melgar, M.C., Ordaz, Sanz J.A. et Guerrero, F.M. (2005), « Diverses alternatives pour déterminer les facteurs significatifs de la fréquence d'accidents dans l'assurance automobile », *Revue Assurances et gestion des risques*, 73-1, 31-54.
- Mollenberghs, G. et Lesaffre, E. (1994), "Marginal Modeling of Correlated Ordinal Data using a Multivariate Plackett Distribution", *JASA*, 89, 633-644.
- O'Brien, S.M. et Dunson, D.B. (2004), "Bayesian Multivariate Logistic Regression", *Biometrics*, 60, 739-746.
- Puelz, R. et Snow, A. (1994), "Evidence on Adverse Selection Equilibrium Signalling and Cross-subsidization in the Insurance Market", *J.P.E.*, 192, 236-257.
- Rothschild, M. et Stiglitz, J. (1976), "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information", *Q.J.E.*, 90, 629-649.
- Salanié, B. (1997). «Théorie des contrats», *Economica*. Paris, 65-86.
- Spence, M.A. (1978), "Product Differentiation and Performance in Insurance Markets", *J.P.E.*, 10, 427-447.

- Todem, D., Kim, K. et Lesaffre, E. (2002), "Latent-Variable Models for Longitudinal Data with Bivariate Ordinal Outcomes", *Technical report*, Department of Biostatistics, Univ. of Wisconsin-Madison.
- Wilson, C.A. (1977), "A Model of Insurance Markets with Incomplete Information", *J. Econ. Theory*, 16, 167-207.
- Zhang, M. et Strawderman, R.L., Cowen M.E., Wells M.T. (2006), "Bayesian Inference for a Two-Part hierarchical Model: An application to Profiling Providers in Managed Health Care", *JASA*, 101, 475, 934-945.