

# LA TARIFICATION PAR GENRE EN ASSURANCE CORRÉLATION OU CAUSALITÉ ?

*Katrien Antonio*

*Professeur, Université catholique de Louvain*

*Arthur Charpentier*

*Maître de conférences, Université de Rennes 1*

*La segmentation en matière d'assurance évoque la classification qu'un assureur opère, selon différents critères, en vue de fixer la cotisation, de telle sorte qu'elle reflète, du mieux possible, le risque représenté par chaque assuré. On parlera ainsi de « segmentation tarifaire ». Et classiquement, les modèles économétriques de régression permettent de capturer les variables les plus corrélées avec la fréquence des sinistres, ou leur coût. Mais, comme le notait Davet [2005], « si la corrélation globale entre âge et coût du risque santé est indéniable, les relations de causalité sont moins simples qu'il n'y paraît ». La corrélation, pourtant importante, entre sinistralité et genre en assurance automobile ne peut plus être évoquée pour créer une discrimination tarifaire depuis décembre 2012 (1). Mais comme nous allons le voir, les objets connectés permettent de récupérer les vraies variables tarifaires (causales) dont le genre n'était alors qu'un proxy.*

---

## Les fondements de la segmentation tarifaire

---

**L**es techniques de l'assurance reposent sur la loi des grands nombres, où des risques homogènes de nature et homogènes de valeur sont agrégés. Il est alors mathématiquement essentiel que les risques soient classés par sous-catégories. Les assureurs évoqueront un principe de justice actuarielle pour justifier la segmentation tarifaire : l'équité entre assurés ne peut être garantie si

la prime n'est pas proportionnelle au risque encouru. L'argument économique associé est lié au fait que l'activité d'assurance est concurrentielle. C'est une question de survie qui pousse les assureurs à segmenter. Si un assureur demande à un assuré une prime supérieure au « risque réel » (au sens estimé par des techniques actuarielles, et perçu par l'assuré), il court le risque de le perdre car il pourrait être attiré par un concurrent qui lui proposerait une prime plus juste, actuariellement parlant. Et demander une prime trop faible signifie perdre de l'argent sur un contrat, sans forcément avoir d'autres assurés pouvant compenser la perte.

## ■ Solidarité et mutualité

La solidarité est une relation entre personnes ayant conscience d'une communauté d'intérêts, qui entraîne pour certaines d'entre elles l'obligation morale de ne pas desservir les autres et de leur porter assistance. La solidarité est un concept moral. Il s'agit d'un engagement de venir en aide aux personnes touchées par une infortune (*"the contribution of the many to the misfortune of the few"*, devise des Lloyd's). Elles se sentent alors solidaires, avec une responsabilité et des intérêts communs. Cette solidarité, qui existait historiquement au sein de la famille, a été transférée au village, puis aux corporations socio-professionnelles. On peut retrouver une expression de cette solidarité dans les régimes de sécurité sociale.

La mutualité est une forme de prévoyance volontaire, par laquelle les membres d'un groupe, moyennant le versement d'une cotisation, s'assurent réciproquement contre certains risques (maladies, blessures, infirmités, etc.) ou se promettent certaines prestations. Le sentiment d'appartenance à une communauté induit une forme de solidarité entre les membres. Mais la mutualité va plus loin, affirmant que les dispositions sont prises à l'avance. De plus, la solidarité qui existe *de facto* n'est pas nécessairement désirée lorsqu'un agent décide de se couvrir contre un risque. L'assurance sera une solution (parmi d'autres), basée sur la solidarité entre ceux qui ont souhaité couvrir un risque similaire, mais d'autres méthodes sont parfois envisageables (on pourra penser aux techniques englobées dans la notion de transfert alternatif des risques). L'assurance privée n'a pas pour fondement la solidarité mais juste le souhait d'agents de transférer leurs risques, et la répartition de la charge globale de ces sinistres entre tous les membres de la mutualité est une forme de solidarité.

Au sein d'une mutualité, un assureur peut décider de ne pas segmenter. Il va alors obliger les risques les plus faibles à subventionner les plus risqués. Cette solidarité subventionnelle est organisée par l'assureur, lorsqu'il constitue des classes tarifaires. On peut noter que la méconnaissance des risques par l'assureur, qui ne dispose que de quelques variables tarifaires (âge du

véhicule, ancienneté du permis de conduire, etc.), lesquelles vont engendrer une incertitude supplémentaire, va induire une solidarité subventionnelle. Formellement, c'est ce que l'on retrouve en utilisant la formule de décomposition de la variance, comme le rappelle Denuit et Charpentier [2005]. Dans un contexte où le risque serait parfaitement connu (noté ici  $\Omega$ ), la variabilité totale se décompose en deux termes

$$\text{Var}[E(S|\Omega)] + E[\text{Var}(S|\Omega)] = \text{Var}[S]$$

Le premier terme (communément appelé *variance inter*) est interprété comme une grandeur mesurant l'inégalité face au risque, alors que le second (appelé *variance intra*) mesure le risque au sein d'une classe tarifaire. Notons que cette formule montre qu'une connaissance parfaite des risques ne tue pas l'aléa : au lieu de dire qu'un conducteur a une probabilité de 8 % d'avoir un accident (comme dans l'ensemble de la population), la segmentation permettra peut-être de dire que sa probabilité d'avoir un accident sur l'année est de 5,63 %. Mais si l'assureur n'a pas accès à  $\Omega$ , mais uniquement à quelques variables dites tarifaires  $X = X_1, \dots, X_k$ , la variance se décompose en

$$\text{Var}[E(S|X)] + E[\text{Var}(S|\Omega)] + E[\text{Var}(E[S|\Omega]|X)]$$

Le terme de gauche désigne toujours la variabilité entre les classes tarifaires. Mais la variance intraclass se décompose ici en deux : un terme incompressible, désigne la variabilité fondamentale (si l'information était parfaite), alors que le second mesure la perte d'information en se contentant de *proxy* du risque sous-jacent. Cette composante, que l'on peut relier à une notion d'information non observée, impose une forme de solidarité entre conducteurs.

## ■ Justice actuarielle et équité

Afin de trouver un équilibre entre solidarité et individualisation du risque, il semble juste de ne retenir que les critères de segmentation qui ont une réelle incidence sur le risque, et que l'assuré peut maîtriser. La puissance du véhicule serait acceptable, mais l'âge ou le genre de l'assuré ne le sont plus. C'est en ce sens

que va la *gender directive*. L'assureur ne peut moduler son tarif qu'en fonction de critères objectivement liés à l'importance du risque, pour autant que l'assuré en ait la maîtrise, en rejetant les critères sur lesquels l'assuré n'a aucune maîtrise. En poursuivant la logique, on considère juste ce qui advient à chacun par l'effet du mérite et de l'effort personnel. On finit alors par remplacer certaines inégalités par d'autres, et le principe de sélection des risques se heurtera toujours à des considérations sociales ou morales.

Cette équité peut aussi trouver du sens au regard des théories économiques. Comme le rappelle Chiappori [1997], un des avantages de la segmentation est d'avoir tendance à réduire l'aléa moral. En effet, l'achat d'assurance incite classiquement à réduire la prudence et la prévention, et ce d'autant plus que la majorité des contrats sont standards (pour ne pas dire identiques). On notera toutefois que l'aléa moral diminuera si la segmentation repose sur des variables que l'assuré maîtrise : il n'y a d'effet incitatif que si la segmentation est fondée sur un critère sur lequel l'assuré a un pouvoir de contrôle (souvent considérées comme des mesures de prévention). Mais il convient de noter qu'il s'agit également d'une demande des consommateurs, qui cherchent souvent le meilleur prix sans payer plus qu'ils ne doivent.

---

## Utilisation de données télématiques en assurance automobile

---

**H**istoriquement, et de manière très pragmatique, les actuaires ont utilisé les variables à leur disposition pour constituer des classes tarifaires (comme l'âge du conducteur principal ou l'ancienneté de son permis de conduire, son genre, son lieu d'habitation ou des informations sur le véhicule, comme la puissance, le type de carburant, etc.). Le type de carburant était souvent utilisé car il est un bon *proxy* de la distance parcourue annuellement. En effet, les conducteurs de véhicules Diesel avaient souvent

tendance à conduire davantage, ce qui augmente la probabilité d'avoir un accident, sur une période d'un an. Le genre du conducteur était utilisé car il était un bon *proxy* du style de conduite ou du type d'usage du véhicule. Ainsi, dans un modèle de type régression poissonnienne, pour modéliser la fréquence annuelle de sinistre le genre est une variable (largement) significative (tout comme le lieu d'habitation, l'ancienneté du permis et la puissance du véhicule).

Mais comme le montre Verbelen *et al.* [2016], en utilisant des données télématiques – comme la distance parcourue par an, l'heure de conduite ou le type de route utilisée –, le genre du conducteur principal n'est alors plus significatif pour expliquer la fréquence annuelle des sinistres (en assurance responsabilité civile automobile). C'est bien la preuve que le genre n'intervenait pas en tant que variable causale dans la survenance des sinistres, mais qu'il s'agissait d'un effet de *confounding* (tel que décrit dans Barry et Charpentier [2015]), le genre étant corrélé avec le style et les habitudes de conduite, variables jusqu'alors non observables.

Les objets connectés permettent aujourd'hui de capturer des effets directement liés (de manière causale cette fois) avec la sinistralité, et permettent d'éviter d'utiliser des variables simplement corrélées, et sur lesquelles l'assuré n'a pas de contrôle. On observe le même effet sur l'âge du conducteur. Cette variable était souvent utilisée par les assureurs pour capturer l'effet de l'expérience des conducteurs, mais dès lors que l'ancienneté du permis est disponible, c'est généralement cette variable qui est la plus significative (montrant que c'est l'inexpérience, plus que l'âge, qui est *pénalisante*) (2).

---

## Vers des primes ajustées

---

**B**ien que les objets connectés permettent d'avoir accès aux vraies variables de risque, il est a priori délicat de les utiliser en tarification. C'est en effet une fois que le contrat d'assurance a été souscrit que le comportement peut

être observé. La solution adoptée par les assureurs est de proposer un contrat d'assurance ajusté, qui tiendra compte du nombre de kilomètres parcourus, des horaires des trajets, du respect des limitations de vitesse, etc. Ces contrats, basés sur une forme de tarification « en temps réel » (ou disons plus dynamique que la révision annuelle classique), proposent de s'adapter à la réalité des habitudes de conduite du conducteur, offrant souvent un remboursement partiel pour encourager les comportements vertueux. Mais si ces nouvelles techniques de tarification n'incitent pas à plus de prudence, le jeu est ici à somme nulle, et les bonus offerts pour des conduites vertueuses devront être compensés par des malus pour les assurés qui conduisent davantage la nuit ou sur des routes jugées dangereuses. À moins que les conducteurs qui ne souhaitent pas installer un boîtier GPS dans leur véhicule n'en fassent les frais.

Les assureurs ont bien compris que les objets connectés vont révolutionner les techniques de tarification. Et la transition ne sera pas sans risque, même si elle promet, à long terme, une différenciation des primes réellement basée sur des facteurs de risque, et non plus sur des variables qui pourraient être jugées comme discriminantes.

## Notes

1. La gender directive correspond ici au mémo IP/11/1581 publié par la Commission européenne le 21 décembre 2012. [http://europa.eu/rapid/press-release\\_IP-12-1430\\_fr.htm](http://europa.eu/rapid/press-release_IP-12-1430_fr.htm)

2. On pourrait toutefois atténuer cette affirmation pour les conducteurs âgés. En effet, avec l'âge, les capacités physiques et mentales diminuent : l'acuité visuelle et auditive se dégrade, la mobilité est réduite, et les réflexes diminuent. Dans une circulation dense et rapide, les réflexes peuvent être inexistantes et les réactions dangereuses. Si un test de santé, ou de réflexe, était disponible, nul doute que cette variable serait plus significative que l'âge en tant que tel.

## Bibliographie

CHIAPPORI P.-A., *Risque et assurance*, Flammarion, 1997.

DAVET J.-L., « L'assurance santé individuelle : des solidarités à l'épreuve de la segmentation », *Risques*, n° 87, 2005.

DENUIT M. ; CHARPENTIER A., *Mathématiques de l'assurance non-vie*, tome 1, Economica, 2005.

VERBELEN R. ; ANTONIO K ; CLAESKENS G., "Unraveling the Predictive Power of Telematics Data in Car Insurance Pricing", KU Leuven Working Paper Series, 2016.

WILS W.P.J., "Insurance Risk Classification in the EC: Regulatory Outlook", *Oxford Journal of Legal Studies*, 1994.