

Mémoire présenté devant l'Institut du Risk Management pour la validation du cursus à la Formation d'Actuaire de l'Institut du Risk Management et l'admission à l'Institut des Actuares le 06/04/2021

Par : **Stéphanie SABLON et François-Xavier DUB**

Titre : **Transférabilité en assurance emprunteur : Impact sur la rentabilité et sur la stratégie de développement d'un assureur**

Confidentialité : NON OUI (Durée : 1 an 2 ans)

Les signataires s'engagent à respecter la confidentialité indiquée ci-dessus

Membres présents du jury de l'Institut des Actuares :

Entreprise :

Nom : **Generali France**

Signature et Cachet :

Membres présents du jury de l'Institut du Risk Management :

Directeur de mémoire en entreprise :

Nom : **Aïcha SOUKI**

Signature :

Autorisation de publication et de mise en ligne sur un site de diffusion de documents actuariels

(après expiration de l'éventuel délai de confidentialité)

Signature du responsable entreprise

Signatures des candidat

Secrétariat :
Bibliothèque :

RÉSUMÉ

Le contrat d'assurance Emprunteur est un contrat d'assurance souscrit au moment de l'octroi d'un prêt, en général dans un contexte d'achat immobilier, dont l'objet est de garantir le paiement des sommes dues à l'organisme prêteur en cas de survenance d'un aléa, en particulier le décès. Historiquement, la faible élasticité des clients au prix de l'assurance emprunteur a conduit les banques à focaliser leurs efforts sur la compétitivité de leurs offres de prêt, en s'assurant par ailleurs des marges élevées sur l'assurance Emprunteur.

Une évolution souhaitée par les acteurs en assurance du marché, qui va encore plus loin dans la simplification des processus est maintenant la suppression des formalités médicales et la garantie du tarif dans le cadre d'un changement de banque prêteuse. Opérationnellement, cela permettrait donc la transférabilité du contrat d'assurance emprunteur. Une possibilité particulièrement utile en cas de renégociation de prêt ou changement d'habitation entraînant le remboursement anticipé et la souscription d'un nouveau prêt. Le bénéfice pour l'assureur d'une telle faculté offerte au client est ambigu : certes, cela lui permettrait de conserver des clients plus longtemps en portefeuille, mais elle a aussi un impact sur sa rentabilité via deux effets : d'une part, un effet sur le volume d'affaires nouvelles souscrites, d'autre part un effet sur la rentabilité individuelle de chaque affaire. L'objectif du mémoire est de fournir à un assureur les analyses et outils lui permettant de prendre la décision d'offrir ou non cette faculté.

L'étude d'un portefeuille d'assurance emprunteur de Generali permet de modéliser, à partir de données récentes, la loi actuelle de décès des assurés à l'aide d'un modèle de Cox et la loi de durée des contrats via une méthode de *machine learning* par arbre de survie. Ces éléments sont exploités pour modéliser et optimiser la rentabilité à long-terme de l'activité d'assurance Emprunteur sous différentes contraintes, d'une part dans le cadre actuel (pas de transférabilité des contrats) et d'autre part dans le cadre proposé (transférabilité des contrats permise).

Notre analyse montre

- que la décision d'offrir la faculté de transférabilité aux clients conduirait à une baisse de la richesse générée à long-terme par l'activité d'assurance emprunteur,
- mais qu'il existe à court-terme une possibilité opportuniste de captation supplémentaire de richesse pour un assureur qui arriverait à aligner certains paramètres pour exploiter l'avantage concurrentiel temporaire que lui donnerait cette décision (dans une phase transitoire, avant que la concurrence ne s'aligne),
- qu'en conséquence l'impact de cette décision sur la richesse de l'entreprise s'apprécie comme la différence entre la valeur durablement détruite et la richesse opportunément captée de manière transitoire. L'observation de cet équilibre fournit ainsi des outils de prise de décision à tout professionnel du secteur qui souhaiterait intégrer à ses offres d'assurance emprunteur l'option de transférabilité.

Notre mémoire présente par ailleurs les limites de notre analyse et ouvre la possibilité d'autres approches de la question de la transférabilité.

Mots clés : Assurance Emprunteur, Tarification, Rentabilité, Kaplan-Meier, Modèle de Cox, Arbre de Survie, Optimisation.

ABSTRACT

Loan insurance is underwritten along a loan, generally for the purchase of a home. Loan insurance covers the lender in case of specific events, most notably the death of the borrower. Low price elasticity on loan insurance has allowed lenders to improve the competitiveness of their loan offering and leverage on higher margins on the insurance.

Clients are now calling in general for more simplification in their customer journey and are more specifically asking for lighter medical procedures and guaranteed prices upon application for the insurance. This would particularly apply to medical procedures upon the application for a new insurance contract in the event of a lender change for an ongoing loan. The suppression of medical tests in this event would allow clients to transfer their existing insurance on any loan with any lender and would offer more room for clients to renegotiate with their lenders without losing the benefits of their insurance. The consequences of such evolution for insurers' profitability is unclear: clients would be retained for a longer time and this would have an impact on profitability through two effects : firstly, the new business volume on the market would decrease, secondly, the profitability of each contract would be impacted. The aim of this thesis is to provide insurers with a toolbox to make a decision regarding this client demand.

The study of a loan insurance portfolio underwritten by Generali allows to model mortality with the help of a Cox model and client retention with the help of survival trees. The results are used to optimize client premiums and forecast profitability in the current paradigm (no transfer of loan insurance with new lenders) and in the new paradigm (loan insurance transfer allowed in case of a new lender).

Our analysis proves :

- The decision to allow clients to transfer their loan insurance with new lenders destructs value in the loan insurance market,
- However there is an opportunity for some insurers in the market to temporarily capture more profits by exploiting the transitory competitive edge they would have (before the whole market follows),
- As a consequence, the impact on value creation for insurers is the sum of short-term, transitory, higher profits, and, later, permanent lower profits. The balance between these two items is sensitive to some parameters, which are the basis of a toolbox for objectifying the decision to allow clients to transfer their loan insurance

Additionally, the thesis outlines the limits of this analysis and opens the way for other approaches to this problem.

Keywords : Loan Insurance, Pricing, Profitability, Kaplan-Meier, Cox Model, Survival Tree, Optimization.

REMERCIEMENTS

Nous tenons à remercier chaleureusement Aïcha SOUKI, notre tutrice en entreprise, Manager de l'équipe Solutions d'Assurance de Personnes, pour son aide, ses précieux conseils et son soutien tout au long de ce mémoire

Nous remercions également Laëtitia GHAMMARTE, Chloé PERUTI et François HATAB, de l'équipe Solutions d'Assurance de Personnes ainsi que Pierre CORREGE, Directeur Pilotage et Performance pour l'attention qu'ils ont portée à ce mémoire et leurs recommandations avisées.

Enfin nous remercions particulièrement nos familles et amis pour nous avoir accompagnés et soutenus durant toute la période d'élaboration de ce mémoire.

RÉSUMÉ	2
ABSTRACT	3
REMERCIEMENTS.....	4
INTRODUCTION.....	7
1 ÉLÉMENTS GÉNÉRAUX SUR L'ASSURANCE EMPRUNTEUR : UN MARCHÉ PORTEUR ET DISPUTÉ.....	8
1.1 CONSIDERATIONS GENERALES SUR L'ASSURANCE EMPRUNTEUR.....	8
1.1.1 Le produit d'assurance.....	8
1.1.2 Souscription	10
1.1.3 Le contexte économique et commercial.....	11
1.1.4 Le contexte législatif	12
1.2 TARIFICATION ET MESURE DE LA RENTABILITE EN ASSURANCE EMPRUNTEUR.....	16
1.2.1 Principes de tarification	16
1.2.2 Principes de mesure de la rentabilité	21
2 MODÉLISATION D'UN PORTEFEUILLE D'ASSURANCE EMPRUNTEUR DANS LE PARADIGME ACTUEL	23
2.1 PRESENTATION DE LA BASE DE DONNEES UTILISEE.....	23
2.1.1 Périmètre	23
2.1.2 Préparation des données.....	24
2.1.3 Définition des variables.....	25
2.1.4 Statistiques descriptives	25
2.2 MODELES DE DUREE : FONDEMENTS THEORIQUES.....	32
2.2.1 Censures et troncatures.....	32
2.2.2 Un modèle non paramétrique : l'estimateur de Kaplan-Meier de la fonction de survie.....	32
2.2.3 Un modèle paramétrique : le modèle de Cox.....	32
2.2.4 Méthodologie de positionnement et critères de validation	33
2.2.5 Théorie des arbres de survie à temps discret	33
2.2.6 Lissage des taux bruts par Splines.....	37
2.2.7 Test d'adéquation du Khi-deux	39
2.2.8 Test des changements de signe	39
2.3 MODELISATION DE LA LOI DE DECES ACTUELLE.....	40
2.3.1 Utilisation d'un modèle à risques proportionnels de Cox.....	41
2.3.2 Positionnement par rapport à la table TH	51
2.3.3 Présentation de la loi de décès simulée.....	61
2.4 MODELISATION DE LA LOI DE RESILIATION ACTUELLE.....	64
2.4.1 Analyse des motifs de résiliation	64
2.4.2 Estimation des taux bruts	65
2.4.3 Recherche graphique des paramètres d'hétérogénéité	66
2.4.4 Approche par un modèle de Cox	73
2.4.5 Approche par arbres de survie.....	76
3 MODÉLISATION DE L'ACTIVITÉ D'ASSURANCE EMPRUNTEUR DANS LE NOUVEAU PARADIGME	87

3.1	RESOLUTION DU PROBLEME D’OPTIMISATION : MAXIMISATION DE LA RENTABILITE GENeree PAR UNE COHORTE DANS LE NOUVEAU PARADIGME	87
3.1.1	Définition du problème d’optimisation	88
3.1.2	Présentation des résultats des modélisations	99
3.2	MODELISATION DE LA RENTABILITE DE L’ACTIVITE D’ASSURANCE EMPRUNTEUR DANS LE NOUVEAU PARADIGME ET PRISE DE DECISION	104
3.2.1	Etude de la phase stable : comparaison des comptes de résultat des cohortes dans le paradigme actuel et dans le nouveau paradigme	105
3.2.2	Projection du compte de résultat de l’activité Emprunteur	106
3.2.3	Limites de l’analyse du nouveau paradigme	109
3.3	DECISION.....	110
	CONCLUSION	111
	BIBLIOGRAPHIE	112
	ANNEXES.....	113

INTRODUCTION

Le contrat d'assurance Emprunteur est un contrat d'assurance souscrit au moment de l'octroi d'un prêt, en général dans un contexte d'achat immobilier, dont l'objet est de garantir le paiement des sommes dues à l'organisme prêteur en cas de survenance d'un aléa, en particulier le décès. Historiquement, la faible élasticité des clients au prix de l'assurance emprunteur a conduit les banques à focaliser leurs efforts sur la compétitivité de leurs offres de prêt, en s'assurant par ailleurs des marges élevées sur l'assurance emprunteur. Les évolutions législatives ont exacerbé la concurrence entre les tenants du marché et leurs challengers : les premiers ont adopté une position défensive afin de contenir la progression de la délégation d'assurance tandis que les seconds ont redoublé d'efforts sur leurs modes de distribution, leurs processus ou leurs exigences de rentabilité. Cet affrontement fait mécaniquement tendre les prix à la baisse, entraîne des évolutions dans les modes de distribution et stimule la créativité des assureurs dans l'approche du risque médical. Après avoir entraîné le marché dans son ensemble vers la sélection médicale, tous les assureurs s'orientent désormais vers une simplification des processus de sélection médicale via une formulation plus simple des questionnaires soumis aux clients et l'allègement des formalités médicales (examens et analyses) exigées pour la décision.

Une évolution souhaitée par les acteurs en assurance du marché, qui va encore plus loin dans la simplification des processus, est maintenant la suppression des formalités médicales et la garantie du tarif dans le cadre d'un changement de banque prêteuse. Opérationnellement, cela permettrait donc la transférabilité du contrat d'assurance emprunteur. Aujourd'hui, la durée moyenne des contrats d'assurance emprunteur est de 8 ans, alignée avec la durée moyenne des prêts immobiliers ; la faculté de transfert de l'assurance économiserait ainsi au client de nouvelles formalités administratives dans le cas où, renégociant son prêt ou changeant d'habitation, il effectue un remboursement anticipé avant de souscrire un nouveau prêt. Le bénéfice pour l'assureur est plus ambigu : certes, cette faculté lui permettrait de conserver des clients plus longtemps en portefeuille et de réduire ainsi son coût moyen d'acquisition, mais elle a aussi un impact sur sa rentabilité via deux effets : d'une part, un effet sur le volume d'affaires nouvelles souscrites, d'autre part un effet sur la rentabilité individuelle de chaque affaire. Aussi, l'assureur qui souhaite maintenir sa rentabilité et sa compétitivité tarifaire doit-il étudier l'opportunité d'offrir cette faculté à la lumière d'une analyse de ces deux effets.

Dans ce mémoire, nous utiliserons le terme *paradigme actuel* pour désigner la situation dans laquelle la possibilité de transférer son contrat d'assurance sans nouvelles formalités médicales n'est pas offerte aux clients. Nous utiliserons le terme *nouveau paradigme* ou *paradigme futur* pour désigner la situation dans laquelle cette possibilité est offerte aux clients.

L'objectif recherché dans ce mémoire est ainsi de déterminer des critères objectifs permettant de juger l'opportunité d'entrer dans le *nouveau paradigme*.

Pour aborder cette problématique, nous proposons :

- Dans un premier temps : d'exposer des éléments généraux sur l'assurance emprunteur et la mesure de la rentabilité de cette activité ;
- Dans un second temps : d'étudier un portefeuille d'assurance emprunteur dans le paradigme actuel, en particulier en modélisant, à partir de données récentes constatées sur un portefeuille Generali, la loi actuelle de durée des contrats et la loi de décès des assurés ;
- Dans un troisième temps : d'étudier une activité d'assurance emprunteur dans le nouveau paradigme selon deux axes :
 - o Un premier axe consistant à identifier, en s'appuyant, sur la modélisation précédente, une optimisation de notre approche tarifaire pour maximiser la rentabilité d'une affaire nouvelle, dans le paradigme actuel ou dans le nouveau paradigme ;
 - o Un second axe consistant à objectiver le processus de décision pour le dirigeant d'une activité d'assurance emprunteur d'entrée ou non dans le nouveau paradigme, à partir d'une analyse chiffrée de la richesse générée ou détruite par cette décision.

1 ÉLÉMENTS GÉNÉRAUX SUR L'ASSURANCE EMPRUNTEUR : UN MARCHÉ PORTEUR ET DISPUTÉ

Dans cette première partie, nous souhaitons exposer quelques éléments généraux sur l'assurance Emprunteur. Nous commencerons par décrire les produits d'assurance emprunteur et les contraintes de marché auxquelles ils sont soumis. Nous verrons ensuite les principes de tarification de ces produits et les éléments de mesure de leur rentabilité.

1.1 Considérations générales sur l'assurance Emprunteur

Dans ce chapitre, nous décrivons les caractéristiques des produits d'assurance emprunteur et le contexte économique, commercial et législatif : après avoir rappelé le contenu des produits d'assurance emprunteur et l'expérience du client souscripteur, nous retraçons l'histoire récente de ce marché, marqué par une opposition entre bancassureurs et assureurs alternatifs, dans laquelle le législateur a cherché à intervenir à de multiples reprises pour imposer une plus grande fluidité.

1.1.1 Le produit d'assurance

1.1.1.1 Objectif de l'assurance emprunteur

Le contrat d'assurance Emprunteur est un contrat d'assurance sur la vie souscrit par une personne physique au moment de l'octroi d'un prêt par sa banque, en général dans un contexte d'achat immobilier à titre de résidence principale. Son objet est de garantir le paiement de tout ou partie des sommes dues à l'organisme prêteur, au titre de l'opération de prêt couverte, en cas de survenance d'un ou plusieurs des aléas suivants au cours de la période de remboursement :

- **Le Décès** de la personne assurée ou sa Perte Totale et Irréversible d'Autonomie (PTIA),
- **L'Incapacité** de la personne assurée, qui peut être Temporaire Totale (ITT) ou Temporaire Partielle (ITP),
- **L'Invalidité** de la personne assurée, qui peut être Permanente Totale (IPT) ou Permanente Partielle (IPP),
- **La Perte d'Emploi** de la personne assurée (PE).

Toutes les personnes physiques sont assurables sur le marché, avec en général une quotité retenue de manière différenciée entre les deux co-emprunteurs.

L'assurance Emprunteur est un produit spécifique aux marchés européens et particulièrement développé en France, dont l'essor sert les objectifs de nombreux acteurs du marché de l'immobilier :

- Du point de vue de l'Etat et de sa politique sociale : cette assurance permet un accès sécurisé à la propriété immobilière pour des personnes dont les héritiers ne pourraient pas prendre le relais du remboursement du prêt en cas de décès ou de problèmes financiers consécutifs à un des événements couverts.
- Du point de vue des acteurs du marché de l'immobilier : cette assurance solvabilise les emprunteurs et désensibilise les vendeurs, promoteurs ou banquiers des problèmes financiers individuellement rencontrés par les assurés.
- Du point de vue des banques : cette assurance sert une stratégie de multi-équipement de leur clientèle en produits de prévoyance, en exploitant un moment clef de la vie via une conception du produit dédié à la couverture du prêt uniquement et un ciblage marketing adéquat.

1.1.1.2 Garanties

Il n'existe pas d'encadrement juridique des risques couverts par le contrat d'assurance Emprunteur, que chaque assureur est libre de bâtir en fonction de sa cible, en y incluant habituellement les garanties suivantes :

- **Le Décès.**
- **La Perte Totale et Irréversible d'Autonomie (P.T.I.A.) :** invalidité physique ou intellectuelle rendant l'Assuré définitivement incapable d'exercer une activité quelconque procurant gain ou profit et nécessitant l'assistance permanente d'une tierce personne pour accomplir tous les actes ordinaires de la vie.
- **L'Incapacité Temporaire Totale (I.T.T.) :** état médicalement constaté d'incapacité totale et temporaire à l'exercice, par l'Assuré, de son activité professionnelle lui procurant gain ou profit (ou de toutes ses occupations habituelles, si l'Assuré est sans activité professionnelle au moment du sinistre).
- **L'Invalidité Permanente Totale (I.P.T.) :** persistance d'une incapacité au travail (ou d'une incapacité à effectuer toutes ses occupations habituelles, si l'Assuré est sans activité professionnelle au moment du sinistre) ayant donné lieu à la constatation médicale de l'impossibilité d'améliorer l'état de santé par un traitement approprié, d'après les connaissances médicales et scientifiques actuelles. L'I.P.T. correspond à un taux d'incapacité supérieur ou égal à 66%, ce taux étant déterminé par voie d'expertise médicale avec les taux d'incapacité permanente fonctionnelle et professionnelle.
- **L'Incapacité Temporaire Partielle (I.T.P.) :** état médicalement constaté d'incapacité partielle et temporaire à l'exercice de l'activité procurant gain ou profit à l'Assuré, lui permettant une reprise dans le cadre d'un mi-temps thérapeutique. Cet état doit intervenir dans la continuité d'une I.T.T. indemnisée par l'Assureur.
- **L'Invalidité Permanente Partielle (I.P.P.) :** persistance d'une incapacité au travail (ou d'une incapacité à effectuer toutes ses occupations habituelles, si l'Assuré est sans activité professionnelle au moment du sinistre) ayant donné lieu à la constatation médicale de l'impossibilité d'améliorer l'état de santé par un traitement approprié, d'après les connaissances médicales et scientifiques actuelles. L'I.P.P. correspond à un taux d'incapacité supérieur ou égal à 33 %, ce taux étant déterminé par voie d'expertise médicale avec les taux d'incapacité permanente fonctionnelle et professionnelle.
- **La Perte d'emploi (P.E.) :** licenciement de l'Assuré, salarié dans le cadre d'un contrat de travail à durée indéterminée (CDI).

1.1.1.3 Produits

La demande croissante par les clients de transparence et de comparabilité des offres a conduit à une standardisation des produits, encouragée ou actée par les nombreuses évolutions législatives et les consensus de place intégrés dans les avis émis par le *Comité consultatif du secteur financier* (CCSF).

Tous les produits d'assurance emprunteur proposés sur le marché sont aujourd'hui des assemblages des garanties recensées dans une liste issue d'un consensus de place, également dite *liste des critères*, présentée dans l'Annexe 1. L'origine de cette liste et son utilisation sont détaillées dans le paragraphe **Equivalence des garanties** ci-dessous.

1.1.2 Souscription

La souscription du contrat d'assurance emprunteur est historiquement très simple dans la mesure où elle était généralement intégrée à la proposition de prêt de la banque et souscrite conjointement sans d'autres formes de formalités. Sous l'effet du risque de remise en cause de la vente, la souscription a peu à peu été formellement séparée du processus de prêt (bien que concomitante), puis la loi ayant imposé une possible substitution par la concurrence, les assureurs ont sophistiqué leur approche de souscription, en introduisant une sélection du risque rigoureuse. Le processus de souscription actuel est donc à la fois plus respectueux du droit du client à choisir librement son assurance, mais également plus lourd et moins bien interfacé avec le prêt. Deux caractéristiques de la souscription sont présentées ci-dessous :

- La sélection médicale, qui a été progressivement introduite par tous les acteurs du marché pour éviter l'anti-sélection dans un contexte concurrentiel,
- Le processus administratif vu du client.

Les acteurs du marché travaillent à une optimisation sur les deux plans :

- La simplification de la sélection médicale, en réduisant les formalités au minimum nécessaire pour sélectionner convenablement le risque,
- L'automatisation des étapes administratives.

1.1.2.1 Processus de sélection médicale

La sélection médicale est assurée par des formalités à réaliser par le client, qui reposent sur trois volets activés selon le niveau de risque couvert :

- **La Déclaration d'Etat de Santé** : simple déclaration générale pré-remplie, signée par le client. Elle est réputée non opposable juridiquement mais constitue encore le socle de sélection de nombreux assureurs, qui vantent sa simplicité et son effectivité pour des risques réduits (faibles montants, emprunteur jeune...). Un exemple de déclaration d'état de santé est présenté en Annexe 2.
- **Le Questionnaire de Santé** : suite de questions sur les antécédents médicaux, en général une quinzaine de questions portant sur les habitudes de vie, les traitements médicaux, les maladies et les arrêts de travail consécutifs à celles-ci. Plus long, plus complexe à compléter pour le client, le questionnaire est considéré par les assureurs comme l'outil de sélection ayant un bon rapport coût/bénéfice pour des risques moyens. Un exemple de questionnaire de santé est présenté en Annexe 3.
- **Les formalités médicales** : une série d'exams de santé et d'entretiens avec des médecins, demandés par les assureurs pour les risques les plus importants (montants particulièrement élevés, âges plus avancés, antécédent médical signalé dans le Questionnaire de Santé...). Un comparatif des formalités médicales exigées par différents assureurs pour un profil d'emprunteur est présenté dans l'Annexe 4. Un exemple de grille de formalités médicales imposées par un assureur est présenté en Annexe 5.

Les seuils auxquels les assureurs exigent une déclaration d'état de santé, un questionnaire de santé ou des formalités médicales ainsi que le contenu de ces documents et formalités sont variables d'un assureur à un autre.

1.1.2.2 Processus administratif

Vu du client, le processus de souscription du contrat d'assurance est secondaire au processus bancaire dans lequel il est impliqué, lui-même accessoire à un moment clef de vie correspondant à l'achat d'un bien immobilier. Le processus d'assurance est donc un processus administratif complexe au sein d'un processus lui-même complexe et stressant.

Les points clés sont les suivants :

- La compréhension par le client des garanties proposées par l'établissement prêteur dans le cadre de son assurance groupe,
- La compréhension par le client des garanties minimales exigées par l'établissement prêteur en cas de délégation d'assurance,
- La comparaison avec les garanties ou conditions proposées par des assureurs alternatifs,
- L'obtention par le client des documents d'assurance nécessaires à l'octroi de son prêt.

Le parcours bancaire du client est présenté en Annexe 6 et le parcours assurantiel est présenté en Annexe 7.

1.1.3 Le contexte économique et commercial

L'assurance emprunteur représente un marché de près de 9 milliards de primes annuellement collectées, l'essentiel de ce marché, qui croît chaque année entre 2 et 4 %, étant constitué par l'assurance emprunteur adossée à un crédit immobilier. Ce marché fut longtemps captif puisque les emprunteurs n'avaient d'autre choix que d'adhérer à l'assurance proposée par leur prêteur, c'est-à-dire très souvent une de ses filiales (ce phénomène est habituellement désigné sous le terme de « bancassurance »).

Les dernières statistiques publiées par la FFA datent de 2017, elles présentent la répartition des cotisations d'assurance emprunteur selon le type de contrats :

- 88 % au titre des contrats souscrits par un établissement de crédits pour leurs clients (soit 8 014 millions d'euros)
- 12 % au titre des contrats souscrits en délégation d'assurance (soit 1 069 millions d'euros).

Les observateurs du marché de l'assurance emprunteur ont abondamment mis en exergue les dysfonctionnements qui l'affectaient ainsi que la rente de situation que les groupes bancaires s'y étaient créés :

- *l'assurance proposée par une entreprise partenaire – éventuellement du même groupe bancaire – était présentée comme fiable ; elle était surtout profitable pour la banque et symétriquement coûteuse pour le client¹ ;*
- *le véritable enjeu de l'assurance emprunteur, en particulier dans le contexte de baisse historique des taux d'intérêt, est la rémunération qu'elle autorise. Les établissements de crédit proposent aux emprunteurs d'adhérer à une assurance de groupe emprunteur qu'ils ont souscrite auprès d'un assureur et qui leur permet de recueillir un niveau moyen de commission de l'ordre de 45 % (tandis que la rémunération de l'assureur se situe autour de 15 %). Ces assurances de groupe dominent très largement le marché de l'assurance emprunteur puisqu'elles représentent presque 85 % des cotisations, laissant peu d'opportunités aux assureurs dit « alternatifs » [c'est-à-dire non adossés à des banques]² .*

Ainsi, dans des débats récents, le député Pierre-Alain Muet (alors vice-président de la commission des finances à l'Assemblée nationale), indiquait devant la commission mixte paritaire ayant conduit à l'adoption de l'amendement Bourquin : *les banques sont en situation de quasi-monopole sur les assurances emprunteur, avec des taux de marge exorbitants, de 50 % en moyenne – contre 10 % en moyenne sur les contrats d'assurance dans leur ensemble – et jusqu'à 70 % sur les contrats des jeunes ménages .*

Le législateur a pris la mesure de cette situation et est intervenu à plusieurs reprises pour encourager la « déliaison » entre l'offreur de crédit immobilier et l'offreur de l'assurance emprunteur. En d'autres termes, le législateur a continuellement cherché à émanciper les emprunteurs des prêteurs de crédit

¹ (Roussille, Mai 2017)

² (Pelissier, 2016)

immobilier en leur redonnant la liberté nécessaire afin de faire jouer au maximum la concurrence entre les différents offreurs d'assurance.

L'entrée sur le marché de nouveaux assureurs concurrents a conduit à la modification progressive de ce marché qui s'est engagé dans un processus de démutualisation. Trois catégories d'acteurs interviennent aujourd'hui sur le marché de l'assurance emprunteur :

- les assureurs, qui produisent les assurances-emprunteurs et assurent le portage du risque et sa couverture en cas de sinistre. Ces assureurs sont à la fois des sociétés d'assurance et des mutuelles d'assurance ;
- les établissements bancaires, qui distribuent des contrats d'assurance-emprunteur ;
- les intermédiaires, notamment :
 - o les courtiers en assurances ou courtiers en crédit ;
 - o les comparateurs de tarifs (dont certains sont aussi des courtiers).

1.1.4 Le contexte législatif

Nous proposons dans cette partie de retracer le contexte législatif relatif à l'assurance emprunteur et son évolution dans les 20 dernières années. Dès les années 2000, l'attention du législateur a été attiré par les lobbyistes et associations de consommateurs sur ce marché peu connu du grand public et générateur de marges confortables pour ses acteurs. Les premiers pointaient du doigt l'absence d'information claire du consommateur, la liaison imposée entre deux produits (le prêt et l'assurance) de nature différente, un partage de la valeur défavorable au client ; les derniers mettaient en avant l'équilibre économique entre un produit d'appel déficitaire (le prêt) et un produit complémentaire (l'assurance) par ailleurs utile socialement. La confrontation entre ces différents points de vue a conduit le législateur à agir.

Le rapport de l'inspection générale des finances de novembre 2013 dédié à l'assurance emprunteur décrit la genèse de la volonté publique d'encadrement de l'assurance emprunteur. Nous en reproduisons la synthèse effectuée en début de rapport qui éclaire particulièrement bien ce contexte :

L'attention portée à l'assurance-emprunteur n'a cessé de croître au cours des années 2000. Plusieurs événements, en France comme à l'étranger, ont incité les pouvoirs publics à étudier et encadrer ce marché présentant alors des dysfonctionnements manifestes. C'est au Royaume-Uni que la campagne médiatique et législative contre les abus d'un secteur mal connu ont été les premières et jusqu'à l'heure actuelle les plus virulentes. Le scandale des PPI (Product Payment Insurance) a éclaté en 2004, à la suite d'un article du Guardian suivi de plusieurs enquêtes de parlementaires et d'associations de consommateurs. Dès 2006 les autorités anglaises imposaient des amendes pour vente abusive à certains prêteurs. A partir de 2009 elles interdisaient certains types d'assurance-emprunteur. Aujourd'hui, les banques continuent de provisionner des milliards d'euros (entre 5 Md€ et 10 Md€ pour l'ensemble du secteur) pour dédommager les quelque trois millions de consommateurs trompés.

En France c'est l'UFC-Que choisir qui a mis l'assurance-emprunteur sur le devant de la scène médiatique. Dès avril 2007, l'association a accusé les banques d'avoir mis la main, au cours des dix dernières années, sur 16 Md€ destinés aux consommateurs au titre de la participation aux bénéfices liée aux assurances-emprunteurs. A ce titre, elle assigne en justice les groupes CNP, Caisse d'Épargne et Cofidis. C'est le point de départ des modifications législatives concernant l'assurance-emprunteur. Dans le cadre de cette action en justice, le 23 juillet 2012, le Conseil d'État confirme l'illégalité de l'article A 331-3 du code des assurances, dans sa version antérieure à 2007, qui excluait les contrats collectifs en cas de décès, et ainsi les assurances-emprunteurs, du calcul des bénéfices techniques et financiers à redistribuer aux assurés. Cette décision n'a pas pour autant ouvert la voie à un droit individuel à participer aux bénéfices. Depuis cet arrêt, ce sont les assurés de la compagnie en général, et pas uniquement les emprunteurs, qui bénéficient d'une partie des bénéfices issus des contrats d'assurance-emprunteur.

La convention AERAS (s'Assurer et Emprunter avec un Risque Aggravé de Santé), conclue le 6 juillet 2006 et entrée en vigueur le 7 janvier 2007, ainsi que la loi 2007-131 du 31 janvier 2007 relative à l'accès au crédit des personnes présentant un risque aggravé de santé, ont également permis d'aborder

certain points repris ensuite dans la « loi Lagarde ». En effet, c'est lors des négociations sur les risques aggravés que les acteurs du marché de l'assurance-emprunteur ont pour la première fois accepté le principe qu'un emprunteur peut ne pas souscrire au contrat bancaire de groupe et choisir de s'assurer auprès d'un tiers (déliation). De la même façon, le principe de garanties équivalentes a été développé dans le cadre de la convention AERAS.³

Parallèlement à ce constat, l'Autorité de Contrôle Prudentiel et de Résolution (ACPR) a souligné la mauvaise volonté des groupes bancaires installés sur le marché à mettre en œuvre les nouveaux droits accordés par le législateur aux emprunteurs. Après avoir réalisé un état des lieux des pratiques appliquées par les principaux acteurs bancaires lors de la mise en œuvre de ces dispositions, elle a constaté que l'information délivrée à la clientèle sur les conditions permettant de faire usage de ce droit était souvent insuffisante et que certaines modalités de traitement des demandes d'assurance externe pouvaient ralentir les démarches des demandeurs voire limiter l'exercice de leur droit⁴.

Le député Pierre-Alain Muet, vice-président de la commission des finances de l'Assemblée Nationale, soulignait dans des débats récents que *L'assurance emprunteur est depuis longtemps un secteur où, malgré les législations successives, les banques ont conservé une situation de quasi-monopole conduisant à des rentes considérables au détriment des emprunteurs⁵.*

De la même manière, le cabinet BAO souligne en 2017 que le mouvement législatif enclenché en 2010 a laissé croire aux acteurs alternatifs que le marché allait enfin s'ouvrir. *Ce fut tout sauf le cas, mais ces acteurs ont régulièrement fait évoluer leurs tarifs à la baisse en anticipation⁶.*

La dernière étape de l'émancipation des emprunteurs à l'égard des prêteurs de crédit immobilier résulte de l'extension du droit de résiliation annuelle prévue par l'article L. 113-12 du code des assurances au-delà des précédentes hypothèses (formation du contrat ou durant les douze premiers mois d'exécution), dans le cadre de Sapin 2, entrée en vigueur en 2017.

1.1.4.1 Loi Lagarde

La « loi Lagarde » instaure l'obligation d'inscrire dans toute offre de prêt la possibilité, pour l'emprunteur, de souscrire une assurance-crédit auprès de l'assureur de son choix.

Le rapport de l'inspection générale des finances précédemment cité donne la lecture suivante de cette loi :

La loi Lagarde a notamment pour but de renforcer la concurrence dans le secteur de l'assurance-emprunteur par le biais de différentes mesures inscrites dans les articles L 312-8 et L 312-9 du code de la consommation. L'article L 312-8 du code précité dispose que toute offre de prêt doit mentionner la possibilité, pour l'emprunteur, de souscrire une assurance-crédit auprès de l'assureur de son choix. Cette obligation d'information du consommateur est complétée par une série de dispositions venant encadrer les activités du prêteur et détaillées dans l'article L312-9 du même code.

Le prêteur a ainsi interdiction :

- *d'exiger la souscription à son propre contrat d'assurance ;*
- *de refuser une assurance tierce présentant un niveau de garantie équivalent au contrat proposé par le prêteur ;*
- *de modifier l'offre de prêt en fonction du contrat d'assurance choisi.*

Tout refus par le prêteur d'un contrat alternatif que lui aurait soumis l'emprunteur doit être motivé.

³ (Inspection générale des Finances, 2013)

⁴ (ACPR, 2017)

⁵ (Muet, 2016)

⁶ (BAO, 2017)

En plus de ces exigences concernant la période précédant immédiatement la signature du contrat de prêt, la « loi Lagarde » ajoute que l'assureur est tenu d'informer le prêteur du non-paiement par l'emprunteur de sa prime d'assurance ou de toute modification substantielle du contrat d'assurance.⁷

En facilitant la concurrence dans un marché de l'assurance-emprunteur alors dominé par les établissements bancaires, les nouvelles règles mises en place par la « loi Lagarde » visaient à permettre une baisse des prix sans remettre en cause le niveau global des garanties.

Les conséquences pratiques de cette loi et ses limites sont très bien décrites dans les différentes études de cabinets de conseil spécialisés dans le domaine. La vivacité du discours montre l'appétence des professionnels pour cette ligne de business, par exemple :

Depuis Septembre 2010, avec l'entrée en vigueur de la Loi Lagarde, tout emprunteur peut choisir « librement » son assurance de prêt immobilier dès lors que les garanties sont équivalentes à celles de la banque. La banque ne peut alors pas modifier les conditions de prêt du fait du choix d'une assurance externe, et doit justifier de tout refus d'assurance sur le plan de la non équivalence de garanties. Mais en réalité, l'offre de prêt n'est pas émise avant la souscription d'assurance puisqu'elle l'intègre dans son contenu. Or, avant l'émission de l'offre de prêt, la banque n'a aucun engagement formel quant à l'octroi du crédit et de ses conditions...Si la loi est une avancée significative qui affirme clairement la liberté de l'emprunteur dans son choix d'assurance, elle ne lui donne finalement aucun moyen opérationnel de la faire respecter. Il signe toujours une offre de prêt liée à une assurance, dans un contexte commercial où il a avant tout besoin d'un prêt...⁸

1.1.4.2 Loi Hamon

La loi Hamon, votée en janvier 2014, accorde un délai d'un an aux emprunteurs, à partir de la signature d'un prêt immobilier, pour changer d'assurance de prêt.

Les dispositions stipulées par ladite loi s'appliquent aux contrats d'assurance signés après le 26 juillet 2014. Les assurances concernées sont celles adossées aux prêts immobiliers :

- destinés à financer l'achat de biens neufs ou anciens à usage d'habitation ou à usage professionnel et d'habitation ;
- accordés pour le financement de travaux d'un montant supérieur à 75 000 euros ;
- octroyés pour la construction d'un bien immobilier (si l'emprunteur achète un terrain).

La loi donne donc un délai d'un an aux emprunteurs, à partir de la signature d'un prêt immobilier, pour changer d'assurance de prêt et encadre les délais de réponse par la banque : à compter de la date de réception de la demande de résiliation du contrat du client, les banques disposent d'un délai de 10 jours ouvrés pour y répondre. En cas de non-respect de ce délai, les établissements bancaires sont soumis à une amende de 3 000 euros.

La loi limite par ailleurs la pratique de facturation à laquelle les banques avaient recours pour décourager la délégation d'assurance : frais d'avenants à l'occasion d'une résiliation effectuée pendant la première année suivant la date de signature de l'offre de prêt, frais pour l'examen d'une assurance externe...

Enfin, la loi impose aux banques de se baser sur la grille du Comité Consultatif du Secteur Financier (CCSF) qui a été établie pour uniformiser la comparaison des contrats d'assurance des emprunteurs.

⁷ (Inspection générale des Finances, 2013)

⁸ (BAO, 2011)

Grâce à la loi Hamon, l'emprunteur qui désire résilier son contrat d'assurance emprunteur actuel pour trouver une offre plus avantageuse peut le faire plus aisément et plus rapidement. Pour cela, il doit respecter quelques conditions :

- la substitution de l'assurance doit être effectuée dans les 12 mois qui suivent la signature de l'offre de prêt ;
- le contrat de substitution doit présenter des garanties aux moins équivalentes à celles du contrat initial ;
- la demande de résiliation doit être effectuée par courrier recommandé avec accusé de réception, 15 jours avant la fin de la première année de signature l'offre de prêt au plus tard.

1.1.4.3 Loi Sapin II

L'article 10 de la loi Sapin II est venu amender la loi Hamon (via un amendement introduit par le sénateur Bourquin). Il permet de changer annuellement (à chaque date d'anniversaire) l'assurance emprunteur, sous réserve de respecter un délai de préavis de 2 mois. À date d'échéance annuelle du contrat, il est donc maintenant possible de résilier un contrat d'assurance de prêt pour lui en substituer un autre souscrit auprès d'un assureur différent et ce, pendant toute la durée du crédit couvert. Les garanties proposées par le nouveau contrat doivent, toutefois, être identiques ou supérieures à celles du contrat d'assurance emprunteur initial.

1.1.4.4 Risque aggravé de santé / Droit à l'oubli

L'accès au crédit représente un enjeu important de société. La question se pose avec une acuité particulière lorsqu'elle concerne des personnes candidates à l'emprunt placées en situation de risque de santé aggravé du fait d'une maladie ou d'un handicap, dans la mesure où l'assurance emprunteur est souvent une condition d'obtention des prêts. Une démarche conventionnelle engagée en 1991 et qui s'est poursuivie continuellement depuis a permis de faire progresser l'accès à l'assurance et au crédit pour les personnes présentant un risque aggravé de santé. L'instauration du « droit à l'oubli » et de la Grille de référence AERAS (GRA) constituent les dernières avancées significatives.

Progressivement :

- Les banques ont été tenues de rechercher toutes les garanties alternatives à cette assurance quand celle-ci était complexe à obtenir, d'autant que ces alternatives existent couramment au travers de la prévoyance des salariés
- La recherche d'une assurance auprès d'assureurs spécialisés a été facilitée, en interdisant aux banques de refuser ces solutions trouvées en dehors de leur intermédiation
- Les surprimes médicales des plus modestes ont été "écrêtées"
- Un système national a été mis en place, pour contrôler les refus d'assurance du marché et trouver les solutions qui n'auraient pas été investiguées par les acteurs. Ainsi, le 3ème niveau AERAS géré par le BCAC est en charge d'examiner tous les refus d'assurance.

Ces dispositions ont été retracées dans la convention AERAS de 2006, traduite en loi en janvier 2007, convention élargie en 2011, puis en 2015 avec le droit à l'oubli. Les informations issues du niveau 3 AERAS relatives au traitement de ces dossiers en risque aggravé de santé sont publiées annuellement auprès de l'ensemble des acteurs du marché.

1.1.4.5 Equivalence des garanties

Le préambule de l'avis du CCSF de 01/2015 décrit les objectifs recherchés via cette notion d'équivalence :

La notion d'équivalence du niveau de garantie a été introduite en assurance emprunteur par la loi Lagarde du 1er juillet 2010 : l'article L 312-9 du code de la Consommation est ainsi rédigé : « Le

prêteur ne peut pas refuser en garantie un autre contrat d'assurance dès lors que ce contrat présente un niveau de garantie équivalent au contrat d'assurance de groupe qu'il propose. Toute décision de refus doit être motivée. ». Les lois successives du 26 juillet 2013 et du 17 mars 2014 n'ont pas touché à la notion d'équivalence du niveau de garantie sur laquelle le CCSF s'est penché à plusieurs reprises. Dès 2006, il plaidait pour une comparaison pertinente entre assurance déléguée et assurance de groupe, intégrant les garanties au-delà du seul aspect tarifaire et son Avis du 18 décembre 2012 définit plusieurs principes et recommandations :

- *le dispositif d'évaluation de l'équivalence du niveau de garantie mis en place par l'établissement doit permettre l'analyse de bonne foi de toute proposition d'assurance apportée par l'emprunteur ;*
- *l'équivalence du niveau de garantie n'est pas l'identité des garanties mais découle d'une analyse globale => un refus d'assurance déléguée ne peut être motivé par leur non-identité ;*
- *le devoir de conseil du distributeur d'assurance est essentiel pour que les garanties soient analysées en fonction de la situation concrète de l'assuré ;*
- *sauf cas particuliers, les dossiers d'assurance déléguée devraient être analysés dans un délai maximal de huit jours, (porté à 10 jours par la loi bancaire du 26 juillet 2013)*
- *en cas de refus de l'assurance déléguée, la demande devrait être réexaminée par une personne ou par une structure autre que celle qui a pris la décision de refus ;*
- *les motivations de refus d'assurance déléguée doivent être explicites, datées et conformes aux meilleures pratiques constatées.*

Cet accord entre toutes les parties prenantes (banques, assurances, intermédiaires, consommateurs...) porte notamment sur la définition par le CCSF d'une liste de Place limitative des garanties minimales exigibles par les banques en cas de demande d'assurance alternative lors de la souscription d'un prêt ou dans le cadre d'une résiliation d'assurance emprunteur dans l'année qui suit la signature du prêt. Au sein de cette liste globale, chaque banque pourra choisir et rendra publics 11 critères au plus pour les garanties obligatoires et, le cas échéant, 4 critères supplémentaires au plus pour la garantie perte d'emploi.

L'accord prévoit que toute banque devra remettre au candidat au crédit une fiche personnalisée précisant la liste détaillée des critères exigés. Cette remise devra se faire, en fonction de la situation spécifique de l'emprunteur, suffisamment tôt, en amont de l'émission de l'offre de prêt, pour permettre au candidat à l'emprunt d'exercer sa liberté de choix en matière d'assurance emprunteur dans les conditions prévues par la loi.⁹

La liste des critères d'équivalence issus de cet accord a été présentée en Annexe 1.

1.2 Tarification et mesure de la rentabilité en assurance emprunteur

Dans ce chapitre, nous détaillons les principes actuariels de tarification des produits d'assurance emprunteur et les principes applicables à la mesure de la rentabilité de ces produits. Dans une optique de simplification de l'étude, les éléments de tarification et de mesure de rentabilité présentés ci-dessous portent sur la garantie Décès uniquement.

1.2.1 Principes de tarification

Dans cette partie, nous envisageons l'assurance proposée à un client dans le cadre d'une acquisition immobilière et utilisons les notations suivantes :

⁹ (Comité consultatif du secteur financier, 2015)

- x est l'âge de l'assuré au moment de la souscription,
- l'opération est financée par un prêt d'une durée de N années et de taux d'intérêt i_m . Le capital emprunté est noté K et le capital restant dû en début d'année k est noté CRD_k ,
- les notations actuarielles habituelles de probabilité de survie (et respectivement décès) entre x et $x+k$ pour une personne d'âge x sont utilisées : ${}_k p_x$ (respectivement : probabilité de décès ${}_k q_x$,
- r_k est la probabilité que l'assuré n'ait pas résilié son contrat après k années.

1.2.1.1 Prime pure

La prime pure du contrat d'assurance emprunteur résulte du calcul du montant que l'assureur doit prévoir en $t=0$ (date de début du contrat) pour fournir la garantie donnée.

La tarification repose sur les principes classiques de tarification : à la souscription du contrat, les valeurs actuelles probables des engagements respectifs de l'assureur - $VAP(A)$ - et de l'assuré - $VAP(a)$ - sont égales.

Le montant égal à la valeur actuelle probable des engagements futurs de l'assureur est appelé prime unique pure. L'assuré verse quant à lui une prime à échéances régulières, sa composante technique est appelée prime périodique pure.

La prime pure :

- correspond à la tarification des engagements de l'assureur envers le bénéficiaire ;
- n'intègre pas le financement des coûts commerciaux et des coûts de gestion supportés par l'assureur. Pour faire face à ces coûts, la prime pure sera majorée de chargements par l'assureur.

La commercialisation et la gestion des contrats comporte des frais pour les entreprises d'assurance :

- Frais d'acquisition : commercialisation, commissions, préparation des dossiers...
- Frais de gestion des contrats : encaissements des primes et comptabilisation, relation client, relevés annuels de situation, etc.
- Frais de gestion des sinistres : paiement des capitaux, des rentes, gestion des contentieux, etc.
- Frais de gestion financière : gestion des investissements (personnel, délégation), coûts de transaction, etc.
- Frais généraux / frais d'administration : direction générale, comptabilité générale, commissariat aux comptes, etc.

1.2.1.2 Segmentation tarifaire

Historiquement et jusqu'au début des années 2000, le marché, tenu par les établissements bancaires, reposait sur des contrats fortement mutualisés. La majorité des contrats collectifs bancaires proposait un tarif unique pour tous les emprunteurs (environ 0,42 % du capital initial par an) quelles que soient leurs caractéristiques personnelles ou celles de leur crédit.

L'entrée de courtiers et d'assureurs alternatifs, et notamment d'April, sur le marché de l'assurance-emprunteur à la fin des années 1990 a eu un impact significatif sur le secteur. Capables de proposer des tarifs très attractifs pour une partie des consommateurs (particulièrement les jeunes, les catégories socio-professionnelles les plus favorisées (CSP+) et les non-fumeurs) grâce à une segmentation plus fine des tarifs, ils ont obligé les distributeurs de contrats de groupe à modifier leurs offres. Le tarif unique pour tous les emprunteurs a alors été segmenté dès le début des années 2000. Les plus jeunes se sont alors vu offrir des tarifs sensiblement plus bas tandis que ceux proposés aux emprunteurs plus âgés augmentaient. Désormais, l'ensemble des contrats collectifs bancaires différencie le tarif en fonction de l'âge des emprunteurs. Certains d'entre eux modulent également le tarif de l'assurance en fonction de la durée du prêt. Les contrats alternatifs fixent quant à eux leurs tarifs en intégrant davantage d'informations sur

l'emprunteur. Le coût de l'assurance dépend, en plus de l'âge de l'emprunteur, de sa catégorie socioprofessionnelle et du fait d'être fumeur ou non-fumeur. Certains contrats mobilisent également des informations complémentaires sur les conditions d'exercice de sa profession (manipulation de charges lourdes ou dangereuses, fréquence des déplacements professionnels) ou sur son état de santé (comme son indice de masse corporelle) ou son lieu de résidence géographique.

1.2.1.3 Contraintes réglementaires

Trois éléments sont utilisés pour le calcul des tarifs :

- un taux technique,
- une table de mortalité,
- des chargements.

Les aspects techniques de la tarification (taux technique et table de mortalité) et du provisionnement sont encadrés par la réglementation française (art. A. 132-18 du code des assurances) :

Les tarifs pratiqués par les entreprises d'assurance sur la vie comprennent la rémunération de l'entreprise et sont établis d'après :

- un taux d'intérêt technique (...)
- une tables de mortalité (...)

En revanche, la réglementation française n'impose aucune contrainte directe sur le niveau des chargements, qui constitue donc un élément concurrentiel de différenciation.

1.2.1.4 Tarification en « capital initial »

Pour des raisons de simplicité et de lisibilité par le client, les acteurs bancaires ont quasi systématiquement opté pour des offres d'assurance emprunteur construites sur la base de tarifs constants sur toute la durée de couverture, et exprimées en % du Capital Initial ou montant emprunté.

La prime d'assurance mensuelle est constante et proportionnelle au capital emprunté : $T_{X_x}^{K,n} \times K$, où $T_{X_x}^{K,n}$ se calcule par égalisation des engagements de l'assureur et de l'assuré à $t=0$:

$$\text{Engagement}_{t=0}(\text{Assureur}) = \sum_{k=0}^{N \times 12 - 1} \text{CRD}_k \times \frac{1}{12} q_{x+\frac{k}{12}} \times \frac{k}{12} p_x \times r_{\frac{k}{12}} \times \frac{1}{(1+i_m)^{k+0,5}},$$

$$\text{Engagement}_{t=0}(\text{Assuré}) = T_{X_x}^{K,n} \times K \times \sum_{k=0}^{N \times 12 - 1} \frac{k}{12} p_x \times r_{\frac{k}{12}} \times \frac{1}{(1+i_m)^k},$$

$$T_{X_x}^{K,n} = \frac{\sum_{k=0}^{N \times 12 - 1} \text{CRD}_k \times \frac{1}{12} q_{x+\frac{k}{12}} \times \frac{k}{12} p_x \times r_{\frac{k}{12}} \times \frac{1}{(1+i_m)^{k+0,5}}}{K \times \sum_{k=0}^{N \times 12 - 1} \frac{k}{12} p_x \times r_{\frac{k}{12}} \times \frac{1}{(1+i_m)^k}}.$$

Par ailleurs, afin d'adapter leur structure de tarification à la concurrence des contrats individuels, les réseaux bancaires et leurs partenaires assureurs ont procédé, depuis une dizaine d'années, à une démutualisation partielle des tarifications, essentiellement en fonction de l'âge de l'assuré à la souscription, corrigeant une partie du différentiel entre contrats groupes et contrats individuels, en particulier sur les clients les plus jeunes.

Les contraintes informatiques et le respect par les acteurs bancaires de fortes contraintes liées à la chaîne de crédit et la nécessité d'adaptation rapide et régulière des conditions tarifaires aux évolutions du marché individuel ont cependant conduit les réseaux à une mise en œuvre très hétérogène de cette segmentation tarifaire. Certains réseaux, plus flexibles, ont privilégié l'adaptation régulière des grilles tarifaires « catalogue », alors que d'autres ont, au-delà des ajustements réalisés, privilégié une approche par dérogation tarifaire à partir de ces grilles « catalogue », pilotées le plus souvent de manière structurée et allouées en fonction des caractéristiques de chaque souscripteur.

1.2.1.5 Tarification en « capital restant dû »

La tarification en « capital initial » présente pour le client un prix qui semble élevé lorsqu'il somme les primes à payer sur la durée totale de son crédit sans prendre en compte sa probabilité de rachat anticipé du prêt et de résiliation de l'assurance. Pour offrir des contrats plus compétitifs selon cette mesure de primes à payer sur l'ensemble de la durée du prêt, les assureurs alternatifs ont développé des offres construites sur la base de tarifs variant sur la durée de couverture, calculés à partir de l'application d'une grille de tarification en pourcentage du capital sous risque selon l'âge de l'assuré. Cette tarification est dite en capital restant dû.

La prime mensuelle due par l'assuré au mois k est proportionnelle au capital restant dû en début de mois : $Tx_x^{CRD,n} \times CRD_k$, où $Tx_x^{CRD,n}$ se calcule par égalisation des engagements de l'assureur et de l'assuré à $t=0$:

$$Engagement_{t=0}(\text{Assureur}) = \sum_{k=0}^{N \times 12 - 1} CRD_k \times \frac{1}{12} q_{x+\frac{k}{12}} \times \frac{k}{12} p_x \times r_{\frac{k}{12}} \times \frac{1}{(1+i_m)^{k+0,5}},$$

$$Engagement_{t=0}(\text{Assuré}) = Tx_x^{CRD,n} \times \sum_{k=0}^{N \times 12 - 1} \frac{k}{12} p_x \times r_{\frac{k}{12}} \times \frac{1}{(1+i_m)^k} \times CRD_k,$$

$$Tx_x^{CRD,n} = \frac{\sum_{k=0}^{N \times 12 - 1} CRD_k \times \frac{1}{12} q_{x+\frac{k}{12}} \times \frac{k}{12} p_x \times r_{\frac{k}{12}} \times \frac{1}{(1+i_m)^{k+0,5}}}{\sum_{k=0}^{N \times 12 - 1} \frac{k}{12} p_x \times r_{\frac{k}{12}} \times \frac{1}{(1+i_m)^k} \times CRD_k}.$$

1.2.1.6 Tarification en « capital restant dû et âge atteint »

Pour proposer un profil de tarification différent, les assureurs ont également développé une tarification qui représente à chaque instant le risque encouru. Cette tarification est dite en capital restant dû et âge atteint.

La prime mensuelle due par l'assurée au mois k est proportionnelle au capital restant dû en début de mois : $Tx_x^{CRD,n}(p) \times CRD_k$ et le coefficient est dépendant de l'année p . La formule de calcul de la prime dans l'année p s'établit ainsi de la même manière que la tarification en capital restant dû, mais en faisant l'hypothèse d'un contrat annuel et en égalant les engagements de l'assureur et de l'assuré pour chaque année.

$$T_{X_x}^{\text{CRD},n}(p) = \frac{\sum_{k=p \times 12}^{(p+1) \times 12 - 1} \text{CRD}_k \times \frac{1}{12} q_{x+\frac{k}{12}} \times \frac{k}{12} p_x \times r_{\frac{k}{12}} \times \frac{1}{(1+i_m)^{k+0,5}}}{\sum_{k=p \times 12}^{(p+1) \times 12 - 1} \frac{k}{12} p_x \times r_{\frac{k}{12}} \times \frac{1}{(1+i_m)^k} \times \text{CRD}_k}$$

1.2.1.7 Matrices de déformation

De nombreux assureurs ont ajouté à leur tarification pure des stratégies de déformations de la tarification, réduisant (voire offrant) les primes à la fin du prêt et les augmentant en début. Cela a pour avantage :

- De diminuer le coût de l'assurance au cours du temps, pour des clients ayant de la visibilité sur leurs revenus actuels mais plus d'incertitude sur l'avenir,
- De différencier le tarif appliqué par rapport à celui des concurrents, sans impact sur la rentabilité,
- De présenter un coût total de l'assurance optimisé.

Figure 1 : Courbe de déformation constatée sur le marché (Actelior, 2017)

Année	% Population présent chaque année	% déformation de la prime				
		5 ans	10 ans	15 ans	20 ans	25 ans
0	100%	130%	130%	130%	130%	130%
1	96%	120%	130%	130%	130%	130%
2	92%	110%	130%	130%	130%	130%
3	87%	100%	120%	130%	130%	130%
4	81%	40%	110%	130%	130%	130%
5	75%		100%	120%	130%	130%
6	68%		70%	110%	125%	130%
7	61%		40%	100%	120%	130%
8	55%		30%	90%	115%	125%
9	49%		20%	80%	110%	120%
10	44%			60%	100%	115%
11	39%			40%	85%	110%
12	35%			20%	70%	105%
13	32%			20%	60%	100%
14	28%			20%	50%	90%
15	25%				40%	80%
16	23%				30%	70%
17	20%				20%	60%
18	18%				20%	50%
19	16%				20%	40%
20	15%					30%
21	13%					20%
22	12%					20%
23	11%					20%
24	9%					20%
25	8%					

1.2.1.8 Comparaison des tarifs

Les comparateurs d'assurance présentent en général le coût de l'assurance comme la somme algébrique des primes dues par l'assuré à l'assureur sur toute la durée du contrat. Comme présenté dans le paragraphe précédent, le caractère pluriannuel des garanties emprunteur, adossées à des produits de durée réelle 8 à 10 ans et de durée originelle moyenne de 18 à 20 ans permet à l'assureur de modifier la construction tarifaire de son offre pour optimiser sa rentabilité sans impact sur le coût total de l'assurance tel que présenté au client. En d'autres termes, la lecture de l'indicateur de coût total de l'assurance donne une vision biaisée du coût espéré de l'assurance pour le client. Pour présenter une vision plus transparente, certains comparateurs présentent également un coût de l'assurance sur 8 ou 10 ans.

1.2.2 Principes de mesure de la rentabilité

Le compte de résultat d'un portefeuille d'assurance emprunteur est facilement lisible et présente les items suivants :

- Les primes : sommes versées par les clients (taxes déduites)
- Les frais de distribution
- Les frais de gestion de l'assureur mobilisés à l'acquisition puis au cours de la vie du contrat
- La sinistralité : sommes versées aux clients en cas d'évènement garantis.
- La provision pour risque croissant (PRC) : provision règlementaire qui oblige l'assureur à provisionner lorsque l'engagement assureur est supérieur à l'engagement assuré
- Les produits financiers : correspondent aux intérêts des placements (actifs) relatifs aux contrats

Les comptes de résultat des portefeuilles d'assurance emprunteur actuellement commercialisés par les différentes compagnies d'assurance sont en revanche complexes à interpréter pour un observateur externe compte tenu de stratégies de tarification, de comptabilisation des frais de commercialisation, et d'ampleur de la sélection médicale. De plus, l'analyse de la rentabilité d'un produit d'assurance emprunteur doit être faite en tenant compte de sa durée de vie totale, en projetant tant les éléments du compte de résultat que les éléments de capital.

Figure 2 : Exemple du compte de résultat d'un portefeuille d'assurance Emprunteur en run-off (en K€)

	1	2	3	4	5	6	7	8	9	10
Primes HT	100	98	92	82	65	52	42	33	27	22
Frais de distribution	-22	-22	-21	-18	-15	-12	-9	-7	-6	-5
Frais de gestion	-12	-12	-11	-10	-8	-6	-5	-4	-3	-3
Sinistres yc APRC	-23	-37	-49	-58	-47	-37	-30	-24	-19	-15
Produits financiers	0,0	0,2	0,2	0,3	0,3	0,4	0,3	0,2	0,1	0,1
Résultat	43	27	11	-4	-5	-3	-2	-2	-1	-1

Les compagnies d'assurance ont musclé dans les dernières années leurs principes de gouvernance des nouveaux produits, qui s'appuie en général sur des seuils de rentabilité que doivent respecter les trois indicateurs suivants :

- Le Taux de Rendement Interne
- La New Business Value
- La durée de retour sur investissement

Les politiques de souscription des compagnies d'assurance sont abondamment commentées dans leurs rapport Embedded Value

1.2.2.1 Taux de rendement interne (TRI)

Le premier indicateur examiné par les compagnies d'assurance pour mesurer la rentabilité d'un business d'assurance emprunteur est le taux de rendement interne. La dimension mesurée par cet indicateur est la rémunération du capital à mobiliser pour le business proposé. Ce taux est calculé de telle sorte que la valeur du résultat distribuable soit nulle sous une actualisation à ce taux. L'équation s'écrit alors :

$$\sum_{t=0}^T \frac{FCF_t}{(1 + TRI)^t} = 0,$$

où les FCF_t représentent les free-cashflows (flux de résultats distribuables) périodiques futurs de l'investissement considéré et T la durée du projet.

Le seuil fixé pour valider un nouveau produit est en général déterminé en fonction du taux des emprunts à 10 ans augmenté d'un premium.

1.2.2.2 New Business Value

Le second indicateur examiné par les compagnies d'assurance pour mesurer la rentabilité d'un business d'assurance emprunteur est la New Business Value. La dimension mesurée par cet indicateur est le niveau de profit généré sur la durée pour le business proposé.

La NBV se décompose comme suit :

$$NBV = Strain + VIF ,$$

avec :

- **Strain** : Résultat de l'année 0 incluant les coûts et commissions d'acquisition, généralement négatif
- **Value Of Inforce (VIF)** : Valeur actuelle des résultats futurs distribuables à l'actionnaire générés par le New Business de l'exercice. Soit :

$$VIF = PVFP - \text{coût du capital} - NHR ,$$

- **Coût du capital** : Coût d'immobilisation de la Marge de Solvabilité
- **Present Value of Future Profits (PVFP)** : Valeur actuelle des résultats statutaires nets d'impôts hors prise en compte du coût du capital
- **Non Hedgeable Risks (NHR)** : Correspond à la valeur actuelle nette du coût des risques non-couvrables

La NBV est homogène à un montant en euros. Pour la ramener à une grandeur adimensionnée et pouvoir comparer la performance de produits à un chiffre d'affaires différent, on peut la rapporter à différentes métriques. On considère les ratios suivants :

- **Ratio 1** : NBV/APE , où l'APE (*Annualized Premium Equivalent*) désigne le montant de versement annuel moyen des 10 premières années (dans le cadre de l'assurance emprunteur, l'APE est égale à la prime de première année).
- **Ratio 2** : $NBV/PVNBP$, où la PVNBP (*Present Value of New Business Premium*) est la valeur actualisée des primes

Le seuil fixé pour valider un nouveau produit est en général déterminé en considérant l'un de ces deux ratios. Le seuil fixé par les compagnies d'assurance est publié de manière plus ou moins précise dans leurs rapports annuels, en général 10-20% pour le ratio 1 ou 0-4% pour le ratio 2.

1.2.2.3 Payback Period (PbP)

Le troisième indicateur examiné par les compagnies d'assurance pour mesurer la rentabilité d'un business d'assurance emprunteur est la durée de retour sur investissement. La dimension mesurée par cet indicateur est la capacité du business à générer du profit. Cette durée est calculée comme la période nécessaire à ce que la somme des profits industriels (sans prise en compte du capital à mobiliser) soit positive. Le seuil fixé pour valider un nouveau produit est en général déterminé en considérant le ratio de PBP sur duration des contrats. Le plafond fixé par les compagnies d'assurance est en général égal à 60-75%.

2 MODÉLISATION D'UN PORTEFEUILLE D'ASSURANCE EMPRUNTEUR DANS LE PARADIGME ACTUEL

Dans le cadre de ce mémoire l'enjeu principal est de modéliser et optimiser la rentabilité d'un portefeuille d'assurance emprunteur dans deux scénarios bien distincts pour pouvoir juger vers lequel il est préférable de se tourner :

- le premier scénario étant l'impossibilité pour l'assuré de transférer son ancien contrat d'assurance en cas de renégociation ou de changement de bien (à iso capital assuré), soit le *paradigme actuel*,
- le second est le scénario où l'assuré dispose justement de cette faculté, soit le *nouveau paradigme*.

Pour pouvoir modéliser cette rentabilité il est donc primordial de disposer des hypothèses techniques nécessaires au calcul du résultat de l'assureur, en particulier les lois de décès et de résiliation qui ont un impact déterminant. C'est l'objectif de cette partie qui s'organise de la manière suivante :

- Dans un premier chapitre, nous caractérisons la base de données utilisée en détaillant le périmètre couvert, les traitements effectués et en présentant des statistiques descriptives.
- Dans un deuxième chapitre, nous présentons quelques notions de modèles de durée utiles pour les modélisations qui suivent.
- Les troisième et quatrième chapitres ont pour objectif de modéliser respectivement la loi de décès et la loi de chute du portefeuille.

2.1 Présentation de la base de données utilisée

Ce chapitre a pour objectif d'exposer la base de données utilisée, nous y précisons le périmètre de clients couverts, les traitements de préparation et autres retraitements effectués. Enfin nous donnons des statistiques descriptives des données sélectionnées pour l'étude.

2.1.1 Périmètre

Nous allons étudier une partie du portefeuille d'assurance emprunteur de Generali en observant le comportement de 3 produits distribués par Generali depuis 2010 au travers de partenariats de distribution avec des courtiers ou des banques. Ces produits sont gérés sur des outils séparés et nous disposons donc de trois sources de données que nous agrégeons pour obtenir un volume suffisant à l'analyse. Nous retraitions de notre étude les contrats ayant fait l'objet d'une surprime sur la garantie décès, de telle sorte que les lois de durée et de mortalité estimées soient valables pour les clients non surprimés.

Nous nous appuyons sur les données arrêtées à 12/2019 et vues à 08/2020 pour avoir suffisamment de recul. En effet cela nous permet notamment d'obtenir une base de données retraitée au maximum des annulations précoces de contrats pour des motifs divers mais qui ne sont pas des résiliations à proprement parlé. Nous prendrons soin dans la section suivante de détailler les différentes raisons qui peuvent mettre fin au contrat et la manière dont l'évènement est retranscrit dans la base de données.

2.1.2 Préparation des données

La préparation des données s'effectue sous R (R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>), privilégié pour le traitement de bases de données volumineuses et reconnu dans le domaine actuariel pour la modélisation statistique.

Nous cherchons à rassembler les informations permettant d'effectuer deux modélisations techniques nécessaires pour notre étude : la modélisation d'une loi de résiliation des contrats et d'une loi de décès. Pour cela nous allons chercher à constituer une base donnant pour chaque client ses caractéristiques propres, les caractéristiques du prêt couvert, celles du contrat d'assurance et sa sinistralité (date du décès éventuel).

2.1.2.1 Données clients et contrats

Nous disposons de trois bases, correspondant aux trois produits commercialisés, avec les caractéristiques des clients (nom, prénom, date de naissance...) et des contrats souscrits (durée, capital emprunté...).

En raison de sources différentes un travail d'homogénéisation a dû être effectué en amont afin de pouvoir traiter et interpréter les différentes informations nécessaires de la même manière et ainsi ne pas biaiser notre étude. Nous ne détaillerons pas ces traitements mais ces derniers ont été longs et fastidieux.

Les bases de données sont telles que nous observons autant de lignes que d'assurés mais également autant de lignes que de prêts. Par exemple si un assuré pour une même opération financière effectue un prêt à taux zéro et un prêt amortissable classique, on observera 2 lignes distinctes dans la base de données détaillant les caractéristiques de chaque prêt pour cet assuré.

Voici ci-dessous les étapes que nous effectuons :

- Nous agrégeons les 3 bases de données homogénéisées,
- Nous supprimons les contrats annulés ou sans suite,
- Nous ne gardons que les contrats présents dans le portefeuille pendant toute ou une partie de la période d'observation (du plus ancien contrat jusqu'au 31/12/2019)

Nous obtenons à l'issue de ce retraitement 575 922 lignes.

La base contrats ainsi établie a à présent deux utilités (loi de résiliation et de décès) qui nécessitent des traitements supplémentaires différents que nous détaillerons dans les sections correspondantes.

2.1.2.2 Données de sinistre

Nous disposons d'une base des sinistres réglés, contenant l'historique des sinistres décès, incapacité et invalidité, tous produits confondus, survenus depuis leur création.

Nous ne gardons que les sinistres correspondant aux personnes ayant souscrit à l'un des 3 produits retenus pour l'étude et uniquement les sinistres décès conformément au périmètre de notre étude. Nous dénombrons alors 781 décès entre 2012 et 2019.

La base contrats est alors complétée de cette base sinistres grâce aux nom, prénom et date de naissance de l'assuré qui servent de clé de liaison entre les deux bases.

2.1.3 Définition des variables

A l'issue de ces traitements nous avons en notre possession un certain nombre de données que nous pouvons classer en deux groupes : les variables quantitatives et les variables qualitatives.

2.1.3.1 Variables quantitatives

Les variables quantitatives sont caractérisées par l'ensemble des informations « numériques » en lien avec l'assuré et son prêt. Nous disposons ainsi des variables suivantes :

- **Date_naissance** : Date de naissance de l'assuré
- **BegAge** : Age de l'assuré à l'adhésion
- **EndAge** : Age de l'assuré à la date de sortie du contrat (déduit à partir de la date de naissance)
- **Date_effet** : Date de prise d'effet des garanties du contrat
- **Date_fin** : Date de sortie du contrat pour cause de décès, résiliation ou date d'échéance atteinte (si aucune de ces causes n'est réalisée la date de fin est enregistrée au 31/12/2019, date de fin de l'observation)
- **Duree** : Durée théorique du contrat calculée à partir des date d'effet et date de fin théorique
- **Anciennete** : Durée réelle du contrat calculée à partir des date d'effet et date de fin observée
- **Capital_assure** : Somme des capitaux sous risque pour un même assuré (montant du prêt × quotité assurée)

2.1.3.2 Variables qualitatives

Les variables qualitatives dont nous disposons sont les suivantes :

- **Cle_client** : Nom et prénom de l'assuré
- **Fumeur** : Caractère tabagique Fumeur ou Non Fumeur de l'assuré (présente pour une partie de la base seulement avec les modalités *Oui* ou *Non*)
- **Gender** : Sexe de l'assuré
- **CSP** : Classe socio professionnelle de l'assuré classée en trois catégories. Les cadres et professions intellectuelles sont la classe 1, les employés et professions intermédiaires la classe 2 et les autres professions ou employés avec travail en hauteur ou déplacements nombreux la classe 3 (variable présente pour une petite partie de la base seulement)
- **Distribution** : Mode de distribution du contrat (courtage, web, réseaux bancaires)
- **Appartenance** : Nom du distributeur (chaque distributeur opère selon un mode de distribution uniquement)
- **Garanties** : Cumul des garanties de l'assuré classé en 2 sous-groupes (0 : Décès seul, 1 : Décès + ITT etc...)
- **Censure_resil** : Variable qui spécifie si un contrat a été résilié ou non (1 si résilié, 0 sinon)
- **Censure_deces** : Variable qui spécifie si l'assuré est décédé ou non (1 si décédé, 0 sinon)
- **Etat_contrat** : Motif de sortie du contrat (décès, détail cause de résiliation, arrivé à échéance)
- **Statut** : Variable de trois modalités spécifiant le statut du contrat (en vigueur, résilié, terminé). Notons qu'un contrat terminé est un contrat qui a pris fin car arrivé à échéance ou parce que l'assuré est décédé.
- **Surprime** : Définit si l'assuré a fait l'objet ou non d'une surprime (modalité à *Oui* ou *Non*). Notons que les contrats supprimés sont retraités dans le cadre de la modélisation de la loi de décès

2.1.4 Statistiques descriptives

La base de données issue des informations sur les contrats et sinistres est la base que nous allons utiliser pour l'ensemble des travaux qui vont suivre dans ce mémoire.

Nous détaillons dans cette partie quelques statistiques descriptives à partir des variables décrites précédemment. Notre but étant dans un premier temps de construire une table de mortalité et une loi de résiliation, il est intéressant d'analyser et comprendre les variables qui pourraient être explicatives dans notre modélisation.

2.1.4.1 Mode de distribution

La base de données que nous avons constituée tire ses informations des trois produits phares commercialisés par Generali. Chacun des produits est vendu via son propre mode de distribution. Soit via les banques, via le courtage ou via Internet.

Voici ci-dessous comment se présentent les contrats selon ces trois grands modes de distribution :

Figure 3 - Répartition des contrats selon le mode de distribution

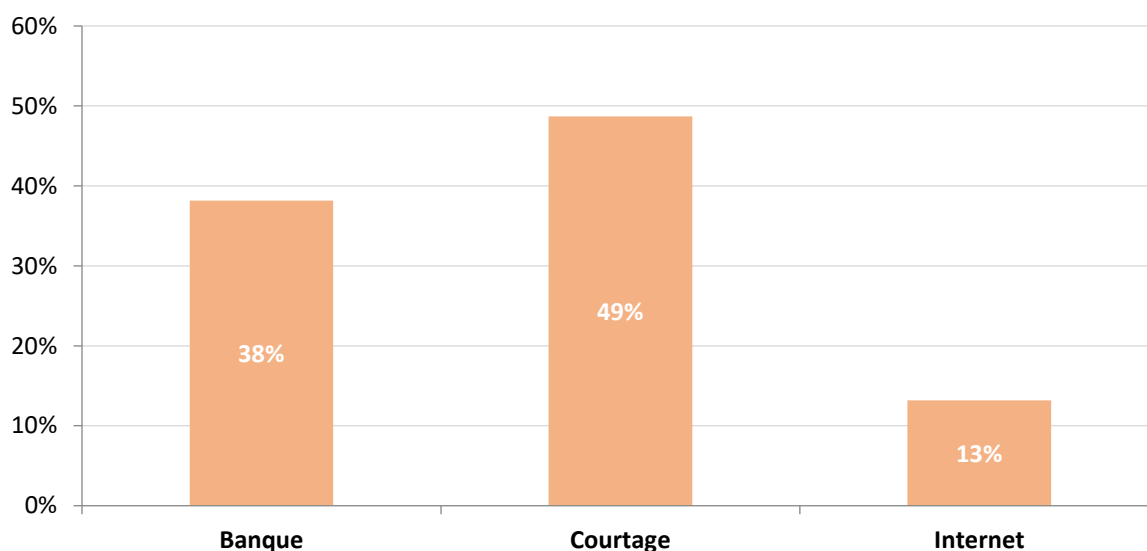


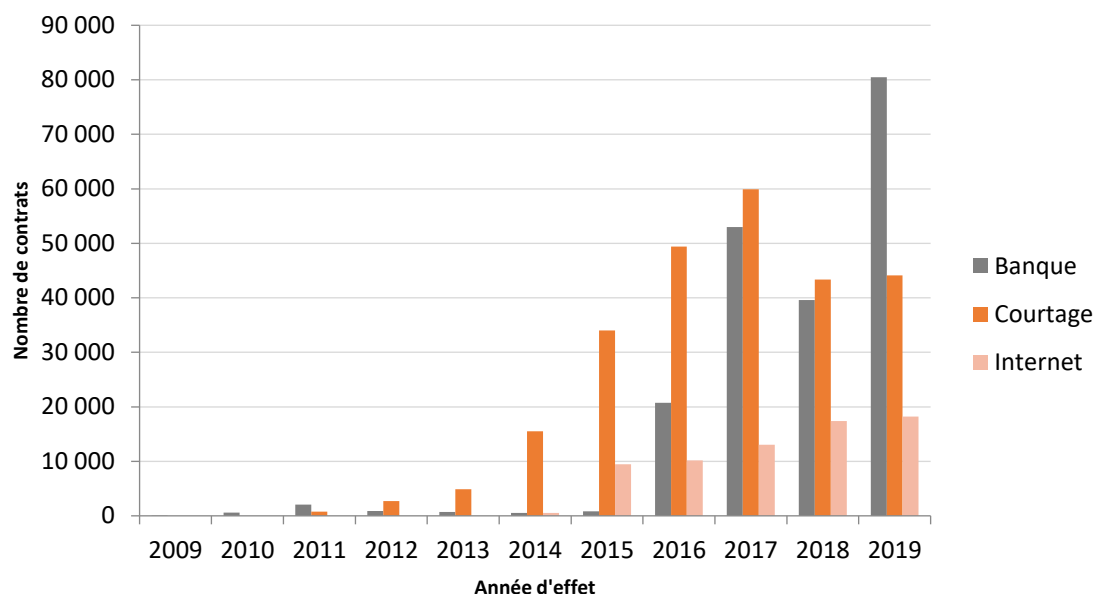
Figure 4 - Nombre de contrats par mode de distribution

Banque	Courtage	Internet	Total
199 438	254 609	68 813	522 860

Le mode de distribution via Courtage est le plus représentatif de notre base de données, soit 49% du portefeuille constitué. Ceci s'explique par le positionnement historique privilégié de Generali sur le marché du courtage (cette répartition n'est pas représentative du marché, qui est capté en premier lieu par les banques, chez lesquelles Generali a des positions plus réduites). Enfin, le mode de distribution online est surreprésenté chez Generali par rapport à la part de marché qu'il représente : sa montée en puissance a été permise par l'essor des reprises de contrat via les lois Hamon et Bourquin, pour lesquelles Generali a mis en place des processus rapides.

Le portefeuille étudié est un portefeuille récent et en forte croissance, comme le montre le graphique ci-dessous.

Figure 5 - Nombre de contrats en fonction de l'année d'effet



2.1.4.2 Sexe de l'assuré

Il est notoire que le sexe de l'assuré est une variable déterminante dans l'explication de la mortalité. Il a déjà été démontré par une multitude d'études que l'espérance de vie des femmes est plus longue que celle des hommes. Il s'agit d'un des principaux enjeux d'un assureur de savoir mesurer cette hétérogénéité dans son portefeuille afin d'en estimer au mieux la mortalité.

Regardons ainsi comment se répartie notre portefeuille selon cette variable :

Figure 6 - Répartition du portefeuille selon le sexe et le mode de distribution

	Banque	Courtage	Internet	Total
Femme	44,3%	47,0%	24,9%	43,1%
Homme	55,7%	53,0%	35,3%	51,7%
Unknown	0,0%	0,0%	39,9%	5,2%

Nous constatons que la répartition entre hommes et femmes est relativement équitable quel que soit le mode de distribution Banque ou Courtage, bien que la proportion d'hommes soit légèrement plus élevée.

Lors de l'établissement de la base de données nous avons dû faire face à une contrainte sur les contrats commercialisés selon le mode de distribution online : le sexe de l'assuré n'étant pas une caractéristique tarifante, celle-ci n'est pas toujours récupérée. Dans le cadre de l'étude, nous créons donc une troisième modalité pour la variable sexe des contrats distribués online : *Unknown*.

2.1.4.3 Age à la souscription

L'âge de l'assuré à la souscription est une variable essentielle dans la modélisation de la table de mortalité car au même titre que le sexe, elle est connue pour être discriminante.

Voici ci-dessous les quelques statistiques élémentaires de cette variable :

Figure 7 - Fréquence des âges à la souscription

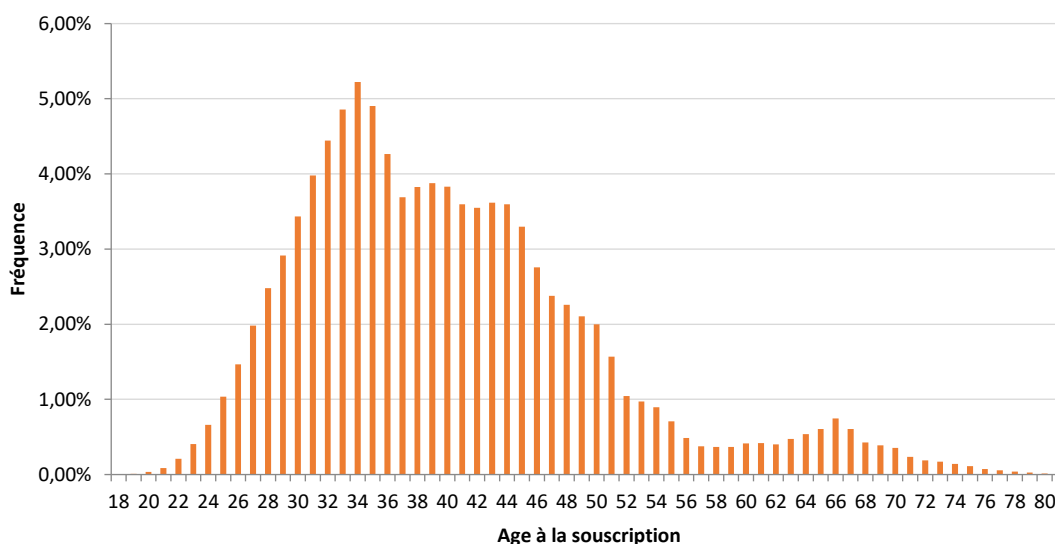


Figure 8 - Statistiques descriptives de l'âge à la souscription

Moyenne	Ecart-type	Médiane	Min	Max
40,2	10,5	39	18	81

	Banque	Courtage	Internet
Age moyen Femme	37,3	41,2	38,5
Age moyen Homme	38,1	42,8	41,5
Age moyen Unknown	-	-	41,6
Age moyen Total	37,7	42,1	40,8

L'âge moyen selon le sexe est relativement stable. Rappelons que l'information sur le genre de l'assuré n'est pas toujours renseignée pour les données issues du mode de distribution Internet.

Nous remarquons que l'âge moyen du Courtage est légèrement plus élevé, ce qui peut s'expliquer par une spécialisation de certains courtiers du portefeuille sur l'assurance octroyée dans le cadre de regroupements de crédit.

Néanmoins avec un âge moyen global de 40,2 ans, nous sommes légèrement au-dessus du profil de l'emprunteur moyen du marché du crédit immobilier (36,1 ans en 2019 – source www.meilleurtaux.com). Ceci s'explique par une prépondérance des contrats courtés qui historiquement se positionnent sur une population plus âgée.

Figure 9 - Age moyen par statuts de contrat

	En vigueur	Résilié	Terminé	Total
Age moyen	40,1	41,3	42,0	40,2

Concernant l'âge moyen par statut de contrat, il ne semble pas vraiment y avoir de disparité selon cette variable. Cela nous laisse penser intuitivement que l'âge n'a que très peu d'influence sur l'acte de résiliation que nous étudions.

2.1.4.4 Profil fumeur

Le profil fumeur est une caractéristique que définit l'assuré au moment de la souscription au contrat d'assurance. En fonction de la version du produit ce paramètre peut être tarifant, le profil tabagique étant réputé pour avoir un impact sur l'espérance de vie.

De la même manière que pour le sexe de l'assuré la variable fumeur n'est pas toujours correctement renseignée, de ce fait nous avons recréé une modalité *Unknown*.

Figure 10 - Profil fumeur par modes de distribution

	Banque	Courtage	Internet	Total
Fumeur	6,0%	3,8%	0,0%	4,1%
Non fumeur	64,6%	29,5%	0,0%	39,0%
Unknown	29,4%	66,8%	100,0%	56,9%

Mais si nous nous concentrons uniquement sur le sous-ensemble pour lequel la modalité est connue voici ce que nous obtenons :

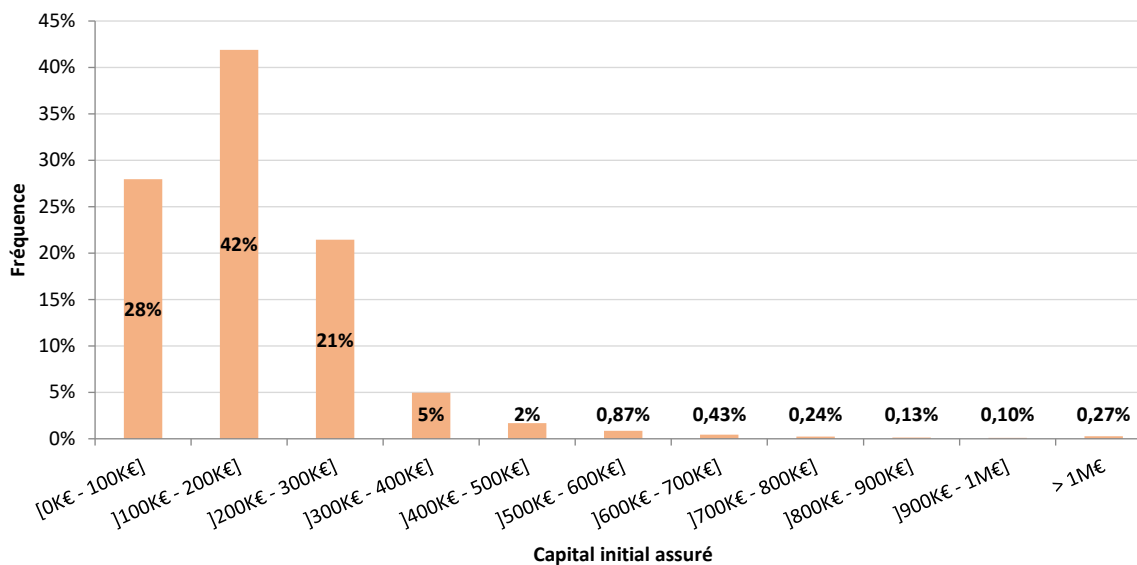
	Banque	Courtage	Total
Fumeur	8,4%	11,3%	9,5%
Non fumeur	91,6%	88,7%	90,5%

Le taux de fumeur à 9,5% est relativement bas comparé à la moyenne française de 24% en 2019 (*source : Fédération Française de la Cardiologie*). Ceci peut s'expliquer par le fait que le portefeuille Generali soit historiquement composé en majorité de Cadres. En effet d'après Santé Publique France la catégorie socioprofessionnelle qui recense le plus de fumeurs est la catégorie « Ouvriers ».

2.1.4.5 Capital initial assuré

Nous nous intéressons maintenant à la variable capital assuré qu'il ne faut pas confondre avec capital emprunté. Effectivement au moment de la souscription du contrat un assuré peut choisir de se couvrir sur tout ou partie de son prêt en choisissant la quotité à assurer.

Figure 11 - Répartition des capitaux assurés



Nous constatons que l'essentielle des affaires souscrites couvrent des capitaux entre 100 K€ et 200 K€, soit 42% du portefeuille étudié. Sachant que le capital moyen assuré du marché de l'assurance emprunteur se situe autour de 180 K€, cette répartition semble cohérente.

Figure 12 - Capitaux moyens assurés par âges

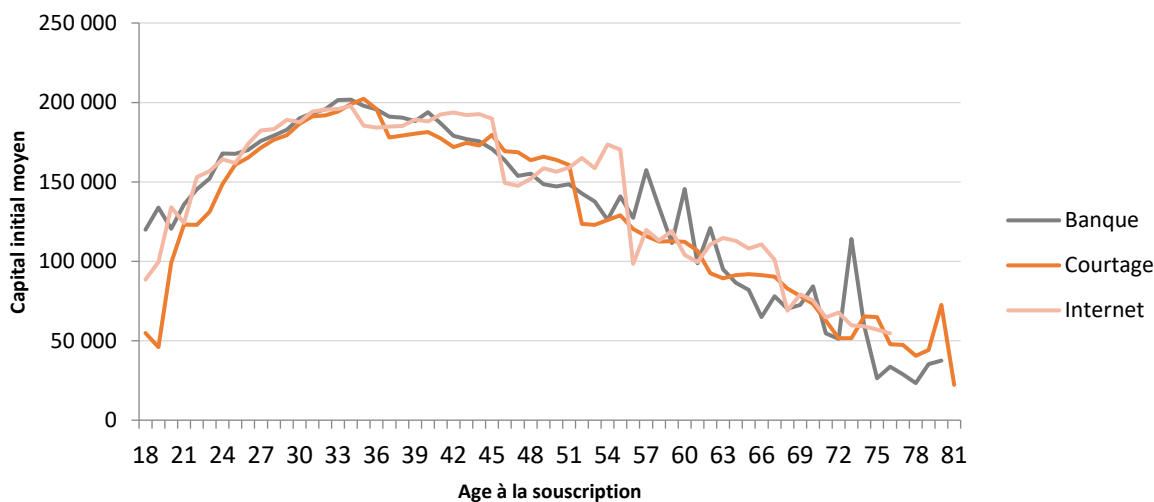


Figure 13 - Capitaux moyens par modes de distribution et statuts de contrat

	En vigueur	Résilié	Terminé	Total
Banque	180 413	208 762	162 071	179 858
Courtage	168 438	151 957	73 299	165 884
Internet	177 773	166 021	95 372	176 307
Total	173 617	157 472	145 414	171 690

Aux vues des graphique et tableau ci-dessus, le canal de distribution n'influe que très peu sur les capitaux moyens assurés. Toutefois nous ne pouvons pas en dire autant de la variable statut du contrat, cette dernière soulignant quelques dissemblances.

En effet les capitaux moyens des contrats terminés apparaissent plus faibles, d'autant plus pour réseaux de Courtage et Internet. Rappelons qu'un contrat terminé est un contrat qui a pris fin soit par ce que l'assuré a fini de rembourser son emprunt soit par ce qu'il ou son/sa conjoint(e) est décédé(e).

Généralement les petits emprunts se remboursent plus rapidement et comme les contrats sinistrés ne sont que très peu représentés cela explique ce résultat.

2.1.4.6 Durée du prêt

La durée du prêt est définie avec la banque lors de l'établissement du contrat de prêt. Il s'agit de la durée théorique de remboursement des échéances de l'emprunt.

Regardons ci-dessous comment se ventilent ces durées de prêt dans notre portefeuille étudié :

Figure 14 - Fréquence des durées de prêt

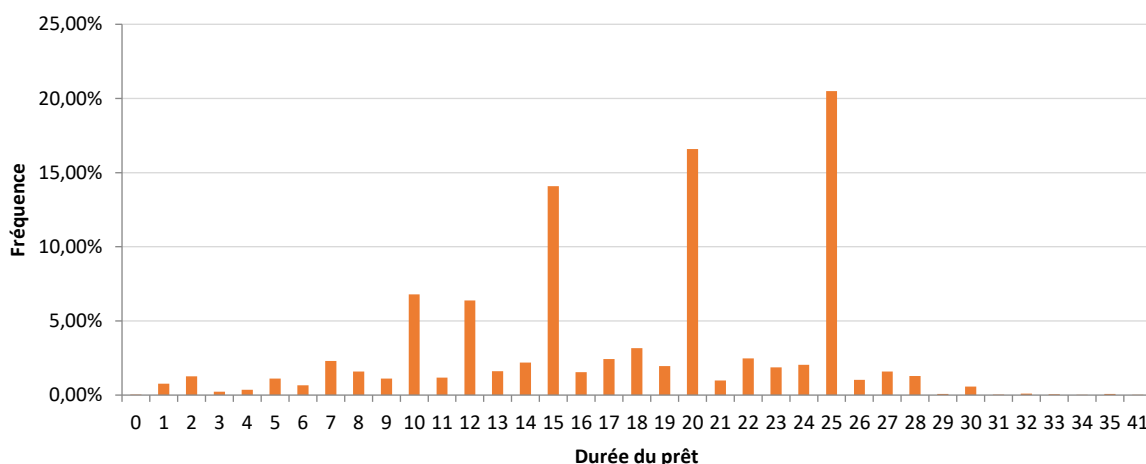


Figure 15 - Durées moyennes selon le mode de distribution

	Banque	Courtage	Internet	Total
Durée moyenne	17,7	17,4	19,3	17,8

Nous observons des pics de fréquence aux durées 15, 20 et 25 ans, ce qui est totalement en lien avec les offres par défaut proposées par les banques et donc les durées de prêts le plus souvent octroyées.

De plus la durée moyenne constatée est en accord avec ce qui se fait sur la place (18,6 ans en 2019).

2.2 Modèles de durée : fondements théoriques

Pour les modélisations présentées dans la partie suivante, nous nous fonderons sur les connaissances classiques en matière de modèle de durée que nous appliquerons tant pour modéliser la durée des contrats (durée entre la souscription et la fin du contrat quelle qu'en soit la cause) que pour modéliser la durée de vie des assurés (durée entre la naissance et la mort). Ce chapitre a ainsi pour objectif de rappeler des notions générales sur les modèles de durée, puis d'exposer les lois et méthodes utilisées pour les modélisations et les tests d'adéquation effectués par la suite.

Nous renvoyons aux nombreux articles scientifiques et cours sur le sujet, en particulier le cours de F. Planchet donné à l'ISFA¹⁰. Nous en reprenons ici seulement quelques notions qui ont un rôle prépondérant dans notre analyse.

Une originalité du mémoire est l'utilisation d'un arbre de décision pour la loi de durée, selon une méthode développée par (Larocque & Bou-Hamad, 2009) (cf. partie 2.2.5.2).

2.2.1 Censures et troncatures

Les données observées pour nos modélisations sont censurées et tronquées.

- **La censure à droite** concerne tant les données de durée des contrats que les données de durée de vie des assurés dans la mesure où notre observation s'arrête au moment de la rédaction de ce mémoire, avant extinction totale du portefeuille.
- **La troncature à gauche** concerne également les données de durée des contrats et les données de durée de vie des assurés, dans la mesure où l'observation débute à un instant donné (avant lequel des clients sont déjà en portefeuille) et s'arrête lors de la résiliation des contrats, qui a parfois pour cause le décès de l'assuré, mais plus souvent une résiliation du contrat pour cause de remboursement anticipé du prêt.

2.2.2 Un modèle non paramétrique : l'estimateur de Kaplan-Meier de la fonction de survie

Le modèle de Kaplan-Meier est l'un des plus employés dans les modèles de durée et utilise l'estimateur suivant de la fonction de survie :

$$\hat{S}(t) = \prod_{T_i \leq t} \left(1 - \frac{d_i}{n_i}\right),$$

où d_i et n_i sont respectivement le nombre de décès au temps T_i et le nombre d'individus juste avant le temps T_i .

Il a les propriétés suivantes : il est convergent et asymptotiquement gaussien.

C'est l'estimateur utilisé dans le package *survival* de R auquel nous avons recours pour l'estimation des lois de durée de base dans ce mémoire.

2.2.3 Un modèle paramétrique : le modèle de Cox

Le modèle de Cox est le modèle classique d'estimation de l'impact des différentes covariables sur la fonction de hasard, sans que cette dernière n'ait besoin d'être définie. Pour sa justification théorique, nous renvoyons à (Planchet & Thérond, Modélisation statistique des phénomènes de durée : Applications actuarielles, 2011).

¹⁰ (Planchet, Modèles de durée, 2019)

Dans le cadre de ce modèle, la fonction de hasard s'écrit :

$$h(x|z; \theta) = \exp(-z' \theta) h_0(x).$$

Le modèle proposé par Cox permet de déterminer les paramètres via une vraisemblance partielle ne faisant pas intervenir de contrainte sur la fonction de hasard de base.

La validation du modèle de Cox nécessite le test de deux hypothèses :

- (i) la validation de l'hypothèse de hasard proportionnel, par exemple via le test des résidus de Schoenfeld (en moyenne nuls si validé).
- (ii) la validation de la non nullité globale des coefficients, par exemple via un test de Wald ou un test du score.

Dans la suite, nous utiliserons un modèle de Cox pour déterminer l'impact de covariables sur la mortalité et la résiliation de nos assurés.

2.2.4 Méthodologie de positionnement et critères de validation

Le groupe de travail Mortalité de l'Institut des Actuaire s'est penché sur la question de la construction de tables de mortalité prospectives et a proposé un arsenal méthodologique complet¹¹. Nous utiliserons cet arsenal dans la suite pour déterminer la fonction de hasard de base nécessaire à la construction d'une table de mortalité.

Quatre méthodes de positionnement de la mortalité par rapport à une table de référence sont envisagées :

- Un positionnement par coefficient de réduction appliqué aux forces de mortalité de la référence qui permet d'ajuster le SMR (*Standardized Mortality Ratio*),
- Une variante semi paramétrique du modèle de Brass dans laquelle le logit de la mortalité est une fonction affine du logit de la mortalité de référence,
- Un modèle linéaire généralisé de Poisson basé sur la mortalité de référence et des interactions avec l'âge,
- Une quatrième méthode experte.

Les critères de validation du positionnement se basent sur :

- Une analyse locale, graphique, comparant la mortalité observée à l'ajustement, et montrant les résidus,
- Une analyse globale, quantitative, étudiant la qualité du lissage et la distance entre les observations et l'ajustement.

2.2.5 Théorie des arbres de survie à temps discret

Les études impliquant des données temporelles sont nombreuses et concernent beaucoup de domaines de recherche. Le modèle de régression à risque proportionnel de Cox et ses extensions sont très souvent utilisés pour étudier les variables de survie avec censure. Ces modèles paramétriques (et semi-paramétriques) sont très utiles, car ils permettent d'interpréter de façon simple des effets de covariables et peuvent facilement être utilisés pour l'inférence (tests d'hypothèses et plus encore). Cependant, de tels modèles imposent un lien spécifique entre les covariables et la variable à expliquer. Même si les interactions entre les covariables peuvent être incorporées dans le modèle, elles doivent être précisées par l'analyste en amont.

¹¹ (Groupe de travail Mortalité de l'Institut des Actuaire, 2016)

Il existe justement des approches plus flexibles lorsque l'analyste ne souhaite pas imposer une fonction de lien. Les arbres et les forêts de survie sont des alternatives non-paramétriques populaires aux modèles (semi) paramétriques car ils offrent une grande flexibilité et peuvent détecter automatiquement certains types d'interactions sans avoir à les préciser au préalable. En effet un arbre peut aisément regrouper les sujets selon leur comportement de survie à l'aide de leurs covariables.

Ainsi c'est pour ces raisons que nous avons choisi d'appliquer cette méthode à nos données de résiliations (ou durées de vie des contrats). Mais avant d'expliquer plus en détail l'approche mathématique des arbres de survie, intéressons-nous dans un premier temps aux fondamentaux d'un arbre de décision.

2.2.5.1 Introduction aux arbres de décision

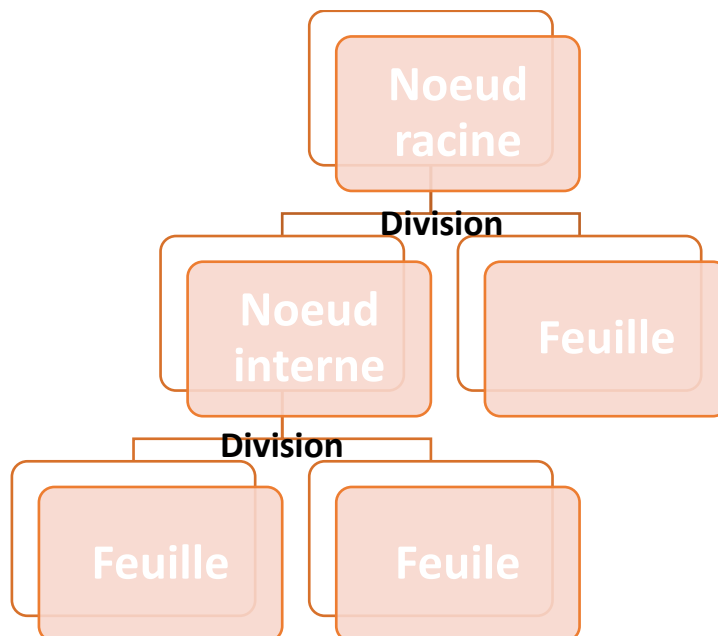
Initialement, les arbres ont été développés pour modéliser des variables discrètes ou continues en utilisant un ensemble de covariables à partir d'un échantillon de données. Ils ont été introduits par Morgan et Sonquist (1963) mais sont vraiment devenus populaires dans les années 80 grâce essentiellement au développement du modèle CART (*Classification And Regression Tree*) décrit dans la monographie de Breiman (Breiman, Friedman, Stone, & Olshen, 1984).

Le principe de base d'un arbre est de partitionner l'espace des covariables de manière récursive pour former des groupes (nœuds dans l'arbre) de sujets qui sont similaires jusqu'à un critère d'arrêt défini.

L'ensemble des nœuds se divise en trois catégories :

- **Nœud racine** (l'accès à l'arbre se fait par ce nœud),
- **Nœuds internes** : les nœuds qui ont des descendants (ou enfants), qui sont à leur tour des nœuds,
- **Nœuds terminaux** (ou feuilles) : nœuds qui n'ont pas de descendant.

Figure 16 - Exemple d'arbre de décision



L'approche de base se concentre sur des divisions binaires en utilisant une seule covariable :

- Pour une covariable continue X , une division potentielle a la forme $X \leq c$ où c est une constante,

- Pour une covariable catégorielle X , une division potentielle a la forme $X \in \{c_1, \dots, c_k\}$ où $\{c_1, \dots, c_k\}$ sont des valeurs possibles de X .

L'algorithme commence au nœud racine avec toutes les observations, ensuite effectue une recherche de toutes les divisions binaires potentielles avec les covariables et enfin sélectionne la meilleure selon un critère de division.

Notons que dans l'approche CART, le processus est répété de manière récursive sur les nœuds enfants jusqu'à ce que le critère d'arrêt soit satisfait. Il en résulte de cet algorithme un grand arbre maximal beaucoup trop ajusté aux données. On appelle ce phénomène le sur-apprentissage (*overfitting*). En d'autres termes, quand l'arbre s'adapte trop aux spécificités de la base de données il devient alors difficile à généraliser à un autre jeu de données. Une méthode d'élagage est alors appliquée afin de trouver un sous-arbre approprié.

Nous allons donc voir dans la partie suivante que dans le cas de données de survie, l'estimateur de Kaplan-Meier peut être un bon indicateur de similitude au sein d'un même nœud.

2.2.5.2 Application à des données de survie

À l'origine, les méthodes d'arbres ont été développées pour modéliser une variable réponse catégorielle univariée (arbre de classification) ou continue (arbre de régression). Cependant, elles ont été étendues à une large variété de situations dont les données longitudinales (Segal, 1992), les données multivariées (Zhang, 1998) et les données censurées (Gordon & Olshen, 1985). Cette dernière extension est connue sous le nom d'arbre de survie.

En général, un arbre de survie sert à regrouper les individus (ou unités statistiques) en classes homogènes par rapport au comportement de survie. Ce regroupement peut être utilisé afin d'identifier les facteurs de pronostic. En outre, un arbre de survie peut fournir des estimations des probabilités de survie et de la fonction de risque en fonction des covariables.

Les méthodes existantes sont conçues pour traiter la variable représentant le temps de survie sur une échelle continue. Cependant, dans la pratique on rencontre souvent des variables de survie mesurées sur une échelle discrète. Lorsque le nombre de temps observés est important et qu'il n'existe pas un nombre trop important de liens, les traiter comme continus serait une approche raisonnable. Cependant, lorsque le nombre de temps observés est petit, certaines méthodes plus adaptées sont nécessaires. Larocque et Bou-Hamad (2009) ont introduit une nouvelle méthode de construction d'arbres spécifiquement adaptée à de telles variables de survie en temps discret dont le critère de séparation est basé sur l'approche de maximum de vraisemblance. C'est cette approche que nous retenons dans le cadre de notre mémoire et pour ce faire nous utilisons le package R *DStree* (*Discret Time Survival Tree*) développé par D. Larocque et qui s'appuie sur ces méthodologies.

Les méthodes basées sur des arbres partitionnent l'espace des covariables de manière récursive avec des règles basées sur ces mêmes covariables. Deux composantes sont nécessaires dans un algorithme de construction d'arbre :

1. une règle de division,
2. une méthode pour sélectionner un arbre à partir de la séquence d'arbres générée par la règle de division (élagage).

2.2.5.2.1 Règle de division

Pour la méthode introduite par Larocque et Bou-Hamad, nous exploiterons la vraisemblance définie ci-dessous pour chaque feuille :

$$L(t) = \prod_{i=1}^{N(t)} \pi_i^{\delta_i}(\tau_i) S_i^{1-\delta_i}(\tau_i),$$

Où :

Pour $U \in \{1, 2, \dots\}$ une variable aléatoire discrète représentant le temps et $\tau = 1, 2, \dots$

- $t \in \mathbb{N}$ est un nœud de l'arbre,
- $i = 1, \dots, N(t)$ est l'ensemble des individus au nœud t ,
- Chaque individu i est observé entre τ_i^0 et τ_i , avec δ_i l'indicateur de censure à la fin d'observation qui vaut 0 si le temps d'observation de l'individu i est censuré à droite et qui vaut 1 si l'évènement est observé,
- $S(\tau) = P(U > \tau)$ est la probabilité de survie au-delà de τ estimée avec l'estimateur de Kaplan-Meier appliqué à la population du nœud considéré,
- $\pi(\tau) = P(U = \tau)$ est la probabilité de l'évènement observé déduite de l'estimateur de Kaplan-Meier de survie appliqué à la population du nœud considéré.

La division choisie est celle qui maximise le produit des vraisemblances (ou la somme des log-vraisemblances) des deux feuilles issues de la division.

Comme pour la méthode CART, seules les divisions binaires sur une covariable sont prises en compte.

2.2.5.2.2 Elagage

La croissance d'un arbre peut en principe continuer tant qu'il y a suffisamment d'observations dans un nœud pour lui permettre d'être encore divisé. Dans la pratique, les arbres sans limitations sont confrontés aux problèmes de sur-apprentissage, c'est pourquoi un critère d'arrêt est nécessaire. Cette étape s'appelle l'élagage.

Définissons l'erreur en termes de coût-complexité comme suit :

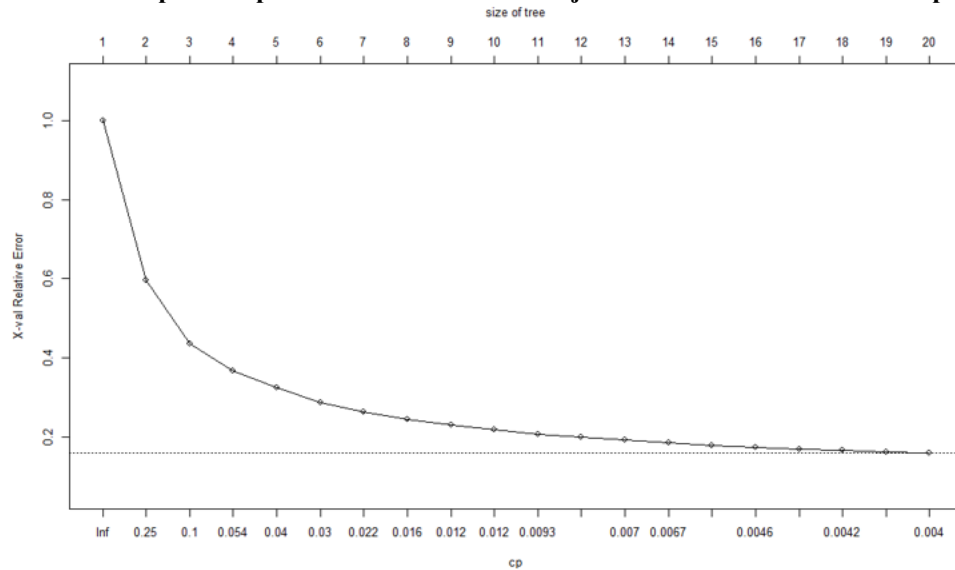
$$Err_\alpha(A) = Err(A) + \alpha |W(A)|,$$

Où pour un arbre A donné :

- $Err(A)$ est la somme des erreurs d'ajustement (Brier Score) à chaque nœud de l'arbre A ,
- $|W(A)|$ est le nombre de nœuds non terminaux de l'arbre A (excepté les feuilles),
- α est un terme de pénalité supérieur ou égal à 0, aussi appelé complexité (cp).

La méthode d'élagage consiste donc à trouver le sous-arbre A_α qui minimise la grandeur $Err_\alpha(A)$ pour une valeur de α donnée.

Figure 17 - Exemple de représentation de l'erreur d'ajustement en fonction de la complexité α



La taille de l'arbre (exprimée ici en nombre de feuille) évolue selon α :

- Si α est petit, la complexité recherchée est élevée, et le sous-arbre est grand,
- Inversement, si α tend vers 0, le sous arbre qui minimise l'erreur est la racine.

2.2.6 Lissage des taux bruts par Splines

Lors de la construction d'une loi de survéance, la première étape consiste en l'estimation de taux bruts ; cette étape est incontournable, que l'approche retenue soit une approche paramétrique ou une approche non paramétrique.

Les valeurs ainsi obtenues présentent une certaine irrégularité, dont il est légitime de penser qu'elle ne reflète pas le phénomène sous-jacent que l'on cherche à mesurer, mais qu'elle est le résultat de l'imperfection des conditions de l'expérience ; les fluctuations d'échantillonnage induisent ainsi une variabilité « parasite » dans les valeurs estimées.

On souhaite alors « ajuster » ou « lisser » les valeurs brutes pour représenter de manière plus fidèle la loi (inconnue) que l'on veut estimer.

L'idée du lissage par splines est de découper la plage de la fonction à ajuster en sous-intervalles, puis d'ajuster sur chaque sous-intervalle une fonction simple, en prenant des précautions pour le raccordement aux points de jonction. Un découpage bien choisi doit en effet permettre d'utiliser sur chaque sous-intervalle une fonction sensiblement plus simple que la fonction qu'il aurait fallu ajuster globalement (Planchet, Modèles de durée, 2019).

Les polynômes sont des fonctions simples et peuvent à ce titre être utilisés pour construire des lissages par spline ; en pratique, on considère généralement des polynômes de degré 3 qui permettent de construire des splines cubiques. Le raccordement de ces arcs se fait en imposant aux points de jonction la continuité ainsi que l'égalité des pentes et des courbures.

Dans le cas où l'on effectue un découpage en 2 parties de la plage de variation des âges (ou anciennetés) pour simplifier les explications, on pose donc :

$$q_x = \begin{cases} p_0(x), & x_0 \leq x \leq x_1 \\ p_1(x), & x_1 \leq x \leq x_2 \end{cases}$$

$p_i(x)$ étant un polynôme de degré 3, avec les contraintes au point de jonction :

$$p_0(x_1) = p_1(x_1) , \quad \frac{d}{dx}p_0(x_1) = \frac{d}{dx}p_1(x_1) , \quad \frac{d^2}{dx^2}p_0(x_1) = \frac{d^2}{dx^2}p_1(x_1) .$$

Cela conduit à poser $p_0(x) = c_1 + c_2x + c_3x^2 + c_4x^3$ et $p_1(x) = p_0(x) + c_5(x - x_1)^3$. Le problème comporte donc 5 inconnues (les 8 coefficients des polynômes moins les 3 contraintes de régularité). On utilise pour le résoudre un critère de moindres carrés pondérés, sur la base de poids (w_x), qui conduit à rechercher les paramètres qui minimisent :

$$M = \sum_{x=x_0}^{x_2} w_x (q_x - \hat{q}_x)^2 .$$

On décompose la somme intervenant dans le critère M en deux sommes puis on écrit les équations normales en annulant les dérivées par rapport aux paramètres. Le problème revient alors à résoudre :

$$\frac{\partial M}{\partial c_i} = 0 .$$

Le choix d'un lissage des données brutes fait intervenir deux types de contraintes qui doivent être prises en considération de manière conjointe :

- la précision (ou fidélité) : il est naturel d'attendre des taux révisés qu'ils soient proches des taux initiaux ;
- la régularité : la suite des taux ajustés sera recherchée aussi régulière que possible.

Dans le cadre de ce mémoire nous nous appuyons sur la méthode des *Cubic Smoothing Spline* documentée dans (Denoeux, 2016), une extension des splines cubique qui justement incorpore dans l'erreur de prédiction des moindres carrés M un critère de régularité en plus de celui de fidélité (fonction *smooth.spline* de R), de sorte à avoir :

$$M_\lambda = \sum_{x=x_0}^{x_2} w_x \left[(q_x - f(x))^2 + \lambda \int (f''(x))^2 dx \right] ,$$

Où

- $f(x) = \hat{q}_x$ est la fonction de spline,
- λ est le paramètre de régularité.

Le paramètre λ définit le niveau de régularité souhaité. Si $\lambda = 0$ aucune contrainte de régularité n'est imposée à f , si $\lambda \rightarrow \infty$ alors f doit nécessairement être linéaire. Choisir λ est équivalent à choisir le degré de liberté (trace de la matrice de lissage, fonction de λ) qui varie entre 0 et le nombre de points à ajuster.

En effet si on note K la matrice diagonalisable telle que :

$$M_\lambda = f K f^T ,$$

Avec $f = (f(x_1), f(x_2))$, alors S_λ la matrice de transformation linéaire s'écrit :

$$\hat{q}_x = (I + \lambda K)^{-1} q_x = S_\lambda q_x ,$$

alors le degré de liberté est défini par :

$$df = Tr(S_\lambda) = \sum_i \frac{1}{1 + \lambda d_i},$$

où les d_i sont les valeurs propres de la matrice K (Cantoni & Hastie, 2000).

Cette approche est celle que nous choisissons pour le lissage de taux bruts car le degré de liberté (df) étant borné cela facilite son utilisation et son interprétation. Par construction plus λ est grand, plus df est petit et donc plus la fonction d'ajustement est « lisse ».

Comme lors de toute démarche statistique, la pertinence de la procédure d'ajustement doit être validée par des tests. Très souvent le test d'adéquation du Khi-deux ou encore le test des changements de signes.

2.2.7 Test d'adéquation du Khi-deux

Le test d'adéquation du χ^2 permet de mesurer la qualité de l'ajustement des différents modèles. Par exemple, pour une loi de mortalité, on effectue le test sur un intervalle d'âge donné :

$$\chi^2 = \sum_{(x,t)} \frac{(D_{x,t} - E_{x,t} \tilde{q}_x(t))^2}{E_{x,t} \tilde{q}_x(t) (1 - \tilde{q}_x(t))},$$

Où

- $D_{x,t}$ est le nombre de décès observés d'individus d'âge x à l'instant t ,
- $E_{x,t}$ est le nombre d'individus d'âge x à l'instant t (exposition),
- $\tilde{q}_x(t)$ est la loi de mortalité estimée.

On recherche la valeur la plus faible pour le test, inférieure à un seuil fixé en fonction du nombre de degrés de liberté du modèle (χ^2 suit une loi de Khi deux de degré de liberté égal au nombre d'observations diminué du nombre de contraintes).

2.2.8 Test des changements de signe

Il s'agit d'un test non-paramétrique qui examine la fréquence des changements de signes de la différence entre les q_x observées et ajustées. Sous l'hypothèse nulle H_0 , la médiane entre les signes positifs et négatifs de cette différence est nulle. Soit le nombre n_+ de signes positifs et n_- de négatifs, avec $n = n_+ + n_-$, la statistique du test des signes S , s'écrit :

$$S = \frac{|n_+ - n_-| - 1}{\sqrt{n}}.$$

Si H_0 est vraie, cette statistique suit une loi Normale de centrée réduite,

$$S \sim N(0,1).$$

Ainsi, l'hypothèse nulle H_0 sera rejetée si :

$$|S| > N_{1-\frac{\alpha}{2}}(0,1),$$

où $N_{1-\alpha/2}(0,1)$ est le $(1 - \alpha/2)$ quantile de la distribution Normale centrée réduite. La p -valeur est la plus petite valeur du risque de première espèce (α) pour laquelle on rejette le test. On privilégiera donc le modèle ayant la p -valeur la plus proche de 1 :

$$p - \text{valeur} = \mathbb{P} \left[N_{1-\frac{\alpha}{2}}(0,1) > |S| \right] = 2 \times \left(1 - F_{N(0,1)}(|S|) \right).$$

2.3 Modélisation de la loi de décès actuelle

L'objectif de ce chapitre est la modélisation de la loi de décès du portefeuille. Pour cela, nous souhaitons calculer les taux annuels de mortalité pour chaque individu en fonction de ses caractéristiques propres et des caractéristiques du prêt.

Pour modéliser la loi de décès de notre portefeuille, nous avons décidé de nous appuyer sur les méthodes classiques, à la fois éprouvées dans de nombreux travaux théoriques et mises en pratique de manière généralisée dans les services de provisionnement, à savoir :

- L'utilisation d'un modèle à risques proportionnels de Cox
- La modélisation de la fonction de hasard de base par un positionnement par rapport à une table de référence.

Avant de nous lancer dans une modélisation détaillée des différents effets, nous visualisons l'estimation Kaplan-Meier de la mortalité du portefeuille.

Figure 18 : Loi de survie observée sur le portefeuille Emprunteur étudié

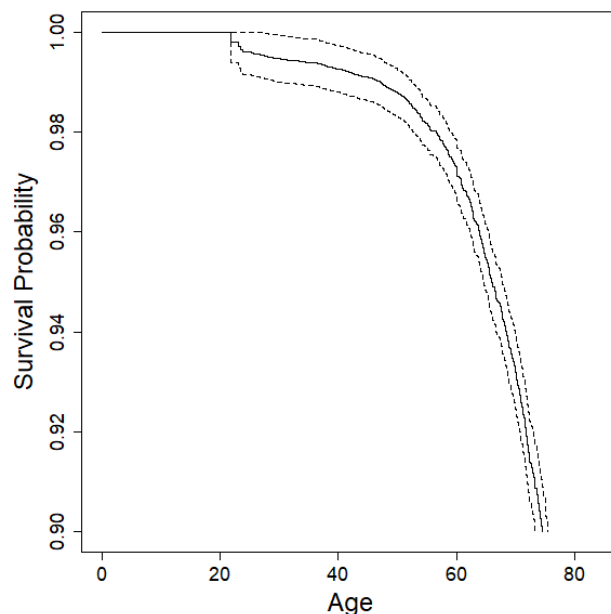
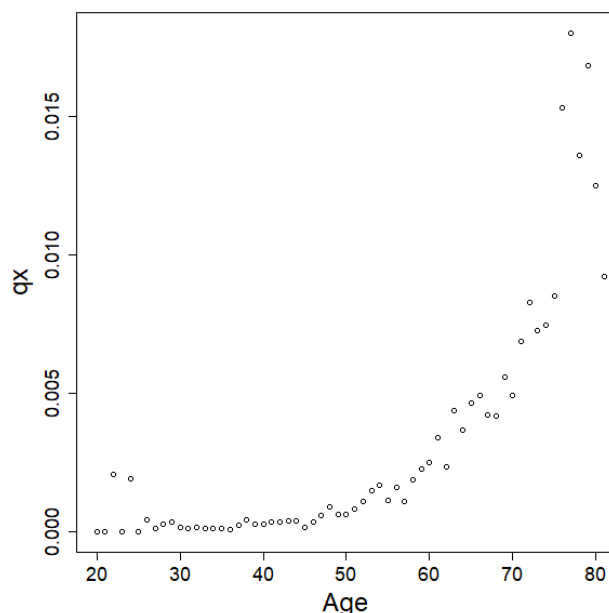


Figure 19 : Loi de mortalité observée sur le portefeuille Emprunteur étudié



Les lois de survie et de mortalité du portefeuille Emprunteur étudié présentent une allure classique, avec quelques sauts liés au nombre limité d'observations. Pour estimer plus précisément la mortalité du portefeuille, il est tentant de modéliser l'impact des différentes caractéristiques connues des clients sur leurs statistiques de mortalité. Aussi allons-nous nous appuyer sur un modèle de Cox dans cet objectif.

2.3.1 Utilisation d'un modèle à risques proportionnels de Cox

Le modèle à risques proportionnels de Cox est une approche classique pour la modélisation du risque de mortalité au sein d'une population présentant des hétérogénéités (classe socio-professionnelle, usage de la cigarette...) et qui présente l'avantage d'être facilement exploitable en tarification, tant d'un point de vue informatique que d'un point de vue pédagogique (du point de vue des réseaux de distribution, il est facile à justifier aux clients). Il permet enfin d'obtenir une mesure de l'effet de la sélection médicale. Tout d'abord nous visualisons graphiquement l'effet de différentes variables.

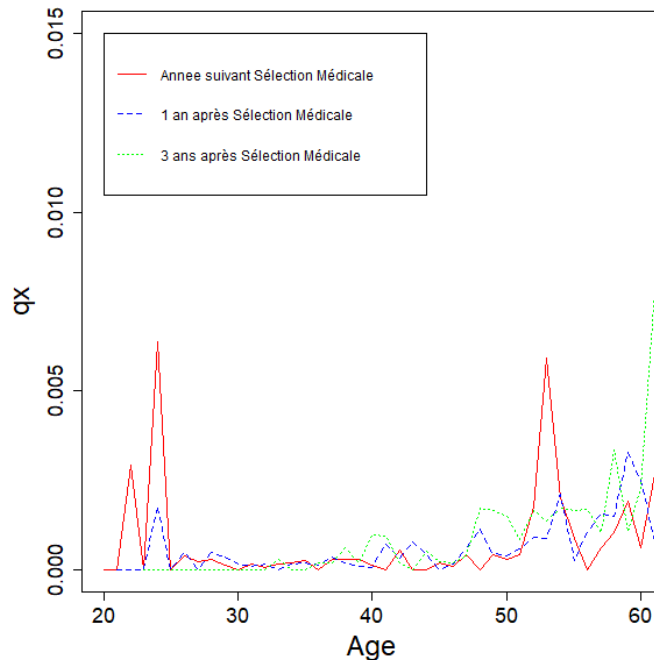
2.3.1.1 Visualisation de l'effet des différentes variables

2.3.1.1.1 Effet de la sélection médicale

La première variable dont nous souhaitons visualiser l'effet est la sélection médicale : nous voulons mesurer l'impact de la sélection médicale sur la mortalité observée dans les années suivant celle-ci (autrement dit : nous souhaitons calculer le taux d'abattement de la loi de mortalité pour les années suivant immédiatement la sélection médicale). Pour cela, nous partitionnons le temps et découpons chaque tête en autant de lignes que d'années restées en portefeuille, tout en ajoutant une nouvelle variable correspondant au temps écoulé entre l'année de la souscription (et donc de la sélection médicale) et l'année observée (cette variable vaut 0 pour la première année de la tête dans le portefeuille et i pour la i -ème année de la tête dans le portefeuille). Pour visualiser l'effet de la sélection médicale, nous estimons par une approche Kaplan-Meier la loi de survie en stratifiant sur cette nouvelle variable correspondant à l'antériorité de la sélection médicale (cf. Annexe 8).

Le graphique n'est pas aisé à lire dans la mesure où les différentes courbes se croisent. Intuitivement, on s'attendrait à visualiser des courbes avec une pente décroissante en fonction de l'antériorité de la sélection médicale. Pour mieux visualiser cet effet, nous allons représenter les courbes de hasard (q_x) issues de l'estimateur Kaplan-Meier de la survie.

Figure 20 : observation de l'effet de la sélection médicale via la comparaison des lois de mortalité constatées sur le portefeuille selon l'antériorité de la sélection médicale.



Cette visualisation correspond davantage à l'effet attendu : si l'on retire les pics et valeurs observés aux âges très jeunes ou très élevés, peu présents en portefeuille, on constate :

- Au sein d'une strate d'antériorité de la sélection médicale : un taux de mortalité croissant avec l'âge
- A un âge donné fixé : un taux de mortalité croissant avec l'antériorité de la sélection médicale (moins élevé dans l'année suivant la sélection médicale et croissant au fur et à mesure).

Autrement dit, ces courbes confirment totalement notre intuition selon laquelle la sélection médicale a un effet bénéfique sur le taux de mortalité.

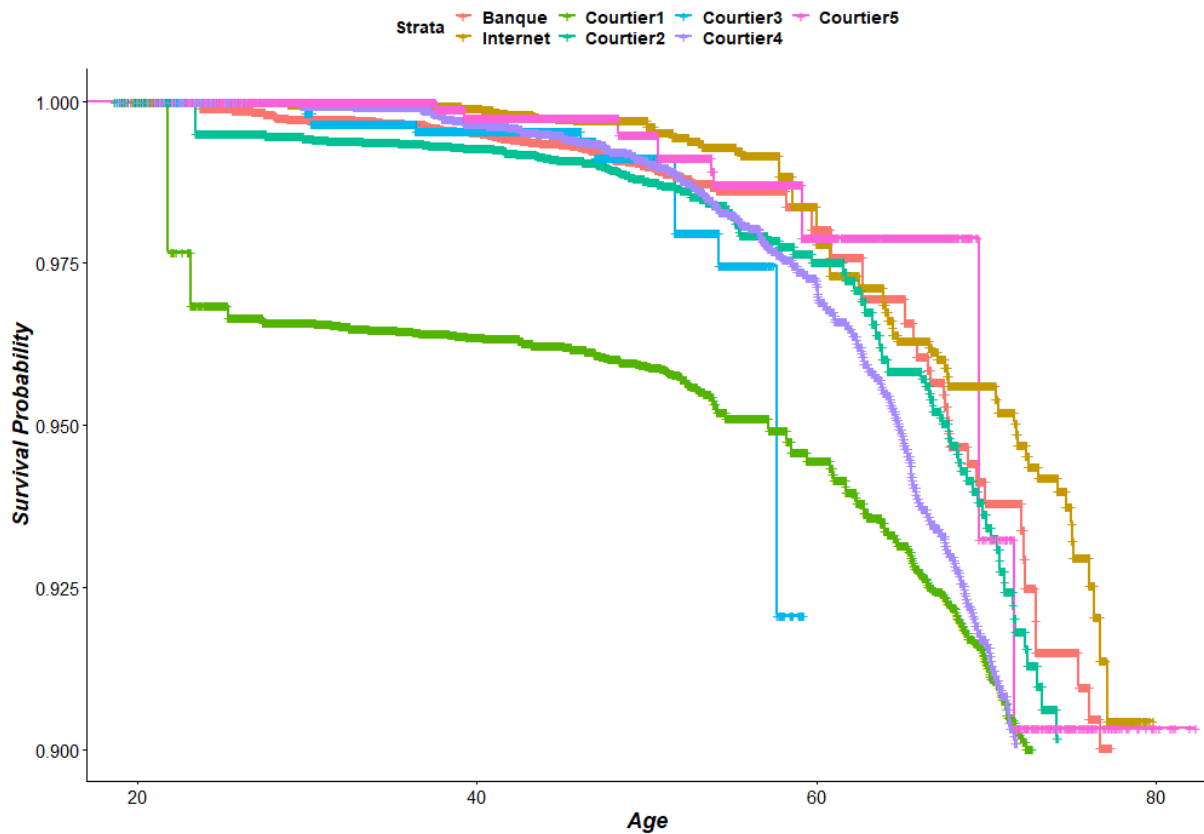
2.3.1.1.2 Effet du mode de distribution et du réseau de distribution

La seconde variable dont l'on souhaite visualiser l'effet est le mode de distribution. Notre portefeuille étant composé de contrats vendus selon des modes très différents, nous souhaitons observer leur effet sur la mortalité. Intuitivement, différents effets contradictoires peuvent jouer :

- la distribution bancaire devrait être moins antisélective que la distribution courtée dans la mesure où les clients prennent cette solution en plus grand nombre, mais...
- la distribution courtée est choisie par des populations plus éduquées financièrement, capables d'organiser la déliaison de leur contrat d'assurance, ce qui est en général corrélé avec un niveau socio-économique plus élevé et un état de santé meilleur. Nous disposons par ailleurs du nom du courtier apporteur, nous pouvons donc vérifier si l'apporteur a un impact sur la loi de mortalité.

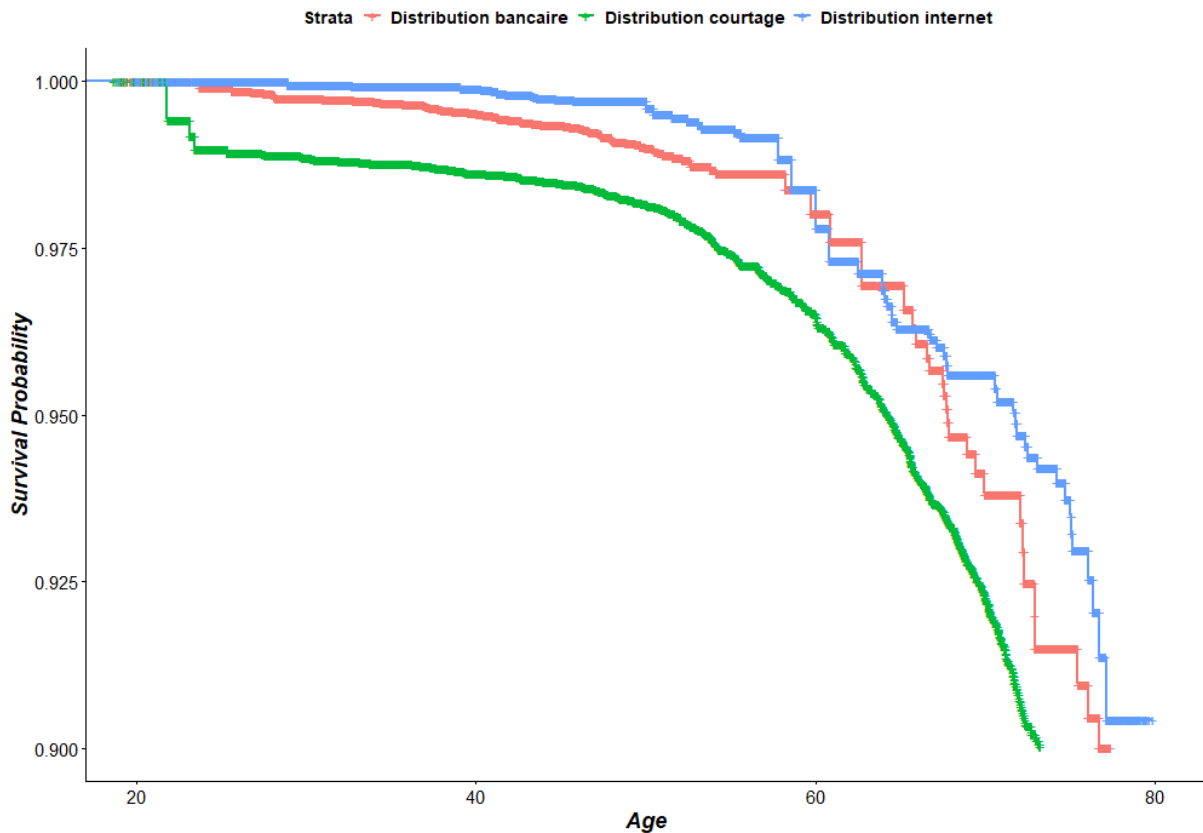
On visualise tout d'abord les différentes courbes de survie estimées selon le mode de distribution, avec un détail par courtier apporteur pour le mode de distribution courtage. Aucune tendance ne se lit facilement sur le graphique.

Figure 21 : Visualisation de l'effet du distributeur sur la loi de survie du portefeuille Emprunteur étudié



Pour tenter de visualiser plus facilement une tendance, on observe les différentes courbes de survie selon le mode de distribution (avec un regroupement de l'ensemble des courtiers apporteurs au sein du mode de distribution Courtage).

Figure 22: Visualisation de l'effet du mode de distribution sur la loi de survie du portefeuille Emprunteur étudié



Les effets contradictoires exposés ci-dessus se constatent sur la visualisation dont on peine de nouveau à dégager une conclusion générale. Il sera intéressant dans la modélisation de confirmer ou d'infirmer cette intuition d'absence de conclusion générale liée au mode de distribution.

2.3.1.2 Modélisation de l'effet de l'ensemble des variables

Après avoir visualisé les différents effets, nous allons maintenant chercher à modéliser l'effet des variables par le modèle de Cox.

Dans une première tentative, on cherche à mesurer l'effet de toutes les variables dont on dispose : sexe, usage de la cigarette (oui/non), niveau de capital emprunté, réseau de distribution via une simulation dans R avec la fonction `coxph`.

Figure 23 : Sortie R de la simulation du modèle de Cox appliqué au portefeuille Emprunteur étudié, avec l'ensemble des variables

Call:

```
coxph(formula = Surv(BegAge, EndAge, Status, type = "counting") ~
      Gender + Fumeur + seg_cap + SM + Appartenance, data = CoxPortfolio)
```

n= 1269125, number of events= 781

	coef	exp(coef)	se(coef)	z	Pr(> z)	
GenderMale	0.75226	2.12180	0.08282	9.083	2,00E-16	***
GenderUnknown	0.45616	1.57800	0.27948	1.632	0.102640	
Fumeur	0.37280	1.45179	0.08100	4.603	4.17e-06	***
seg_cap	-0.11304	0.89312	0.03237	-3.492	0.000479	***
SM	0.13997	1.15024	0.02379	5.884	4.01e-09	***
Appartenance Internet	-0.64897	0.52258	0.21585	-3.007	0.002642	**
Appartenance COURTIER1	-0.03573	0.96490	0.12620	-0.283	0.777101	
Appartenance COURTIER2	-0.21625	0.80553	0.13840	-1.563	0.118156	
Appartenance COURTIER3	0.55743	1.74618	0.28699	1.942	0.052095	.
Appartenance COURTIER4	0.17456	1.19073	0.12837	1.360	0.173888	
Appartenance COURTIER5	-0.61583	0.54019	0.35062	-1.756	0.079016	.

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

	exp(coef)	exp(-coef)	lower.95	upper.95
GenderMale	2.1218	0.4713	1.8039	2.4958
GenderUnknown	1.5780	0.6337	0.9125	2.7290
Fumeur	1.4518	0.6888	1.2387	1.7016
seg_cap	0.8931	1.1197	0.8382	0.9516
SM	1.1502	0.8694	1.0978	1.2051
Appartenance Internet	0.5226	1.9136	0.3423	0.7978
Appartenance COURTIER1	0.9649	1.0364	0.7535	1.2357
Appartenance COURTIER2	0.8055	1.2414	0.6142	1.0565
Appartenance COURTIER3	1.7462	0.5727	0.9950	3.0646
Appartenance COURTIER4	1.1907	0.8398	0.9259	1.5314
Appartenance COURTIER5	0.5402	1.8512	0.2717	1.0740

Concordance= 0.633 (se = 0.013)

Likelihood ratio test= 210 on 11 df, p=<2e-16

Wald test = 208.8 on 11 df, p=<2e-16

Score (logrank) test = 216.1 on 11 df, p=<2e-16

Les sorties du modèle de Cox présentent des résultats intéressants :

- certains réseaux de distribution présentent un impact statistiquement significatif (statistique de Wald élevée, p-value faible) sur le taux de mortalité
- Les trois variables ayant le plus d'effet sont le sexe de l'assuré, son caractère fumeur / non fumeur, le niveau de capital emprunté.
- Le niveau global de significativité du modèle est excellent, avec un test du score présentant une p-value très faible.

Néanmoins, le test de validation de l'hypothèse de proportionnalité des risques présente des résultats tranchés :

Figure 24 : sortie R du test de validation de l'hypothèse de proportionnalité des risques

	chisq	df	p
Gender	0.0545	2	0.97311
Fumeur	12.6809	1	0.00037
seg_cap	0.0615	1	0.80418
SM	0.0242	1	0.87630
Appartenance	8.1925	6	0.22434
GLOBAL	14.9573	11	0.18446

Le test ne permet pas de conclure à la proportionnalité des risques au global, en particulier en raison de l'effet du caractère fumeur / non fumeur d'une part, et de l'effet du réseau de distribution d'autre part. Le graphique du test de Schoenfeld permet de visualiser ces effets.

Figure 25 : sortie R des résidus du test de Schoenfeld sur la proportionnalité des risques dans le modèle de Cox appliqué à l'ensemble des variables (visualisation sans intervalles de confiance)

Global Schoenfeld Test p: 0.1845

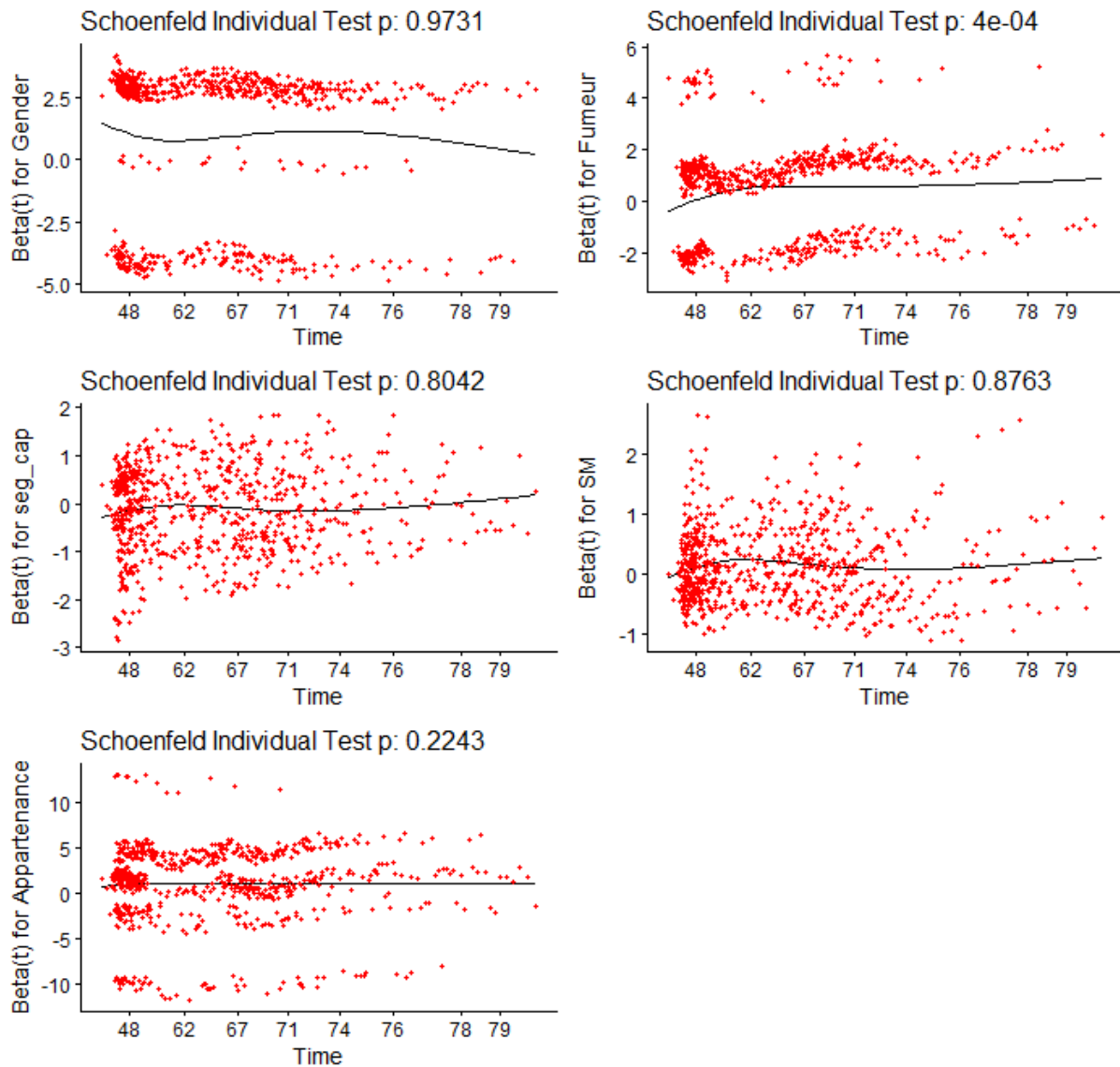
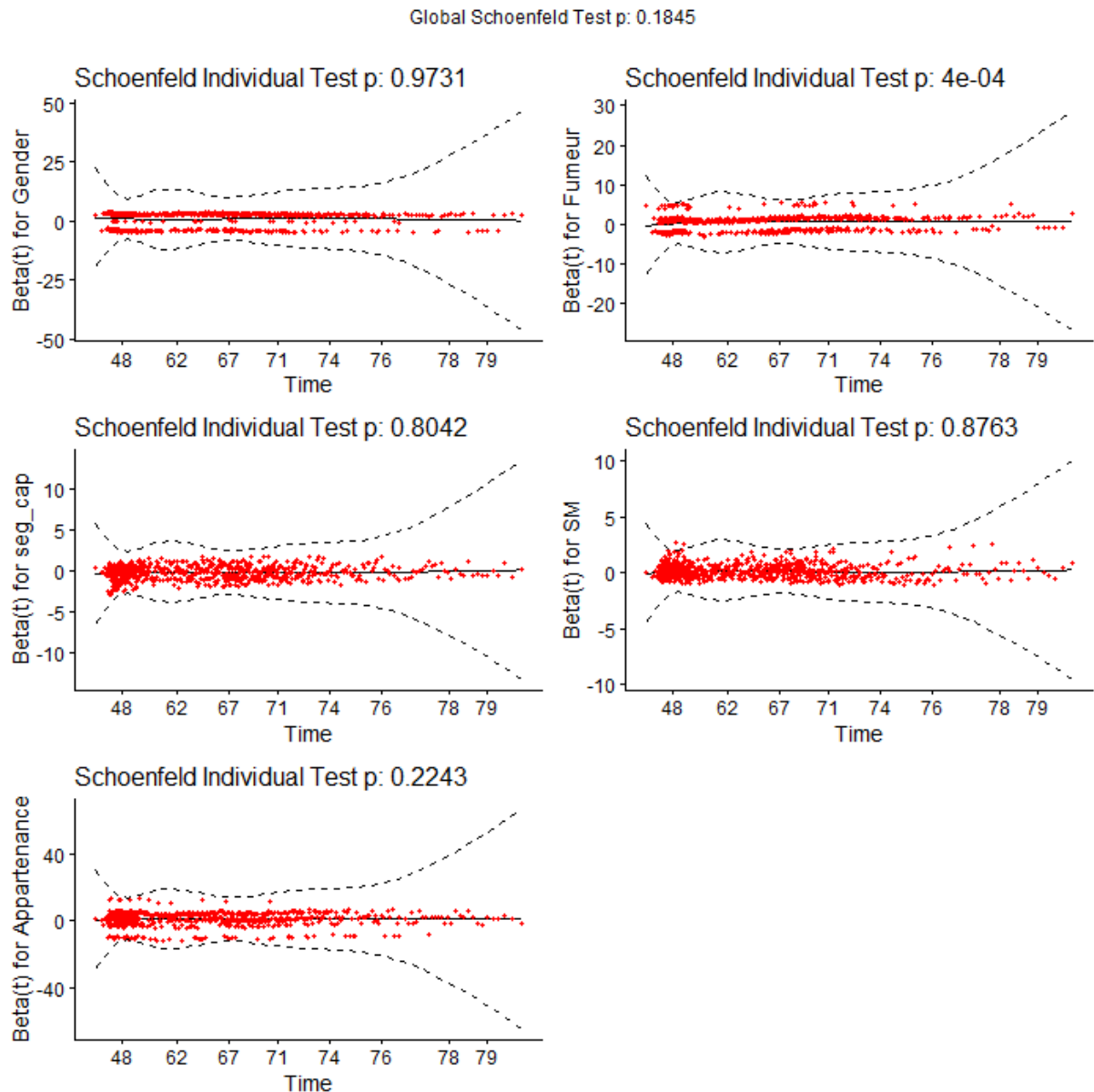


Figure 26 : sortie R des résidus du test de Schoenfeld sur la proportionnalité des risques dans le modèle de Cox appliqué à l'ensemble des variables (visualisation avec intervalles de confiance)



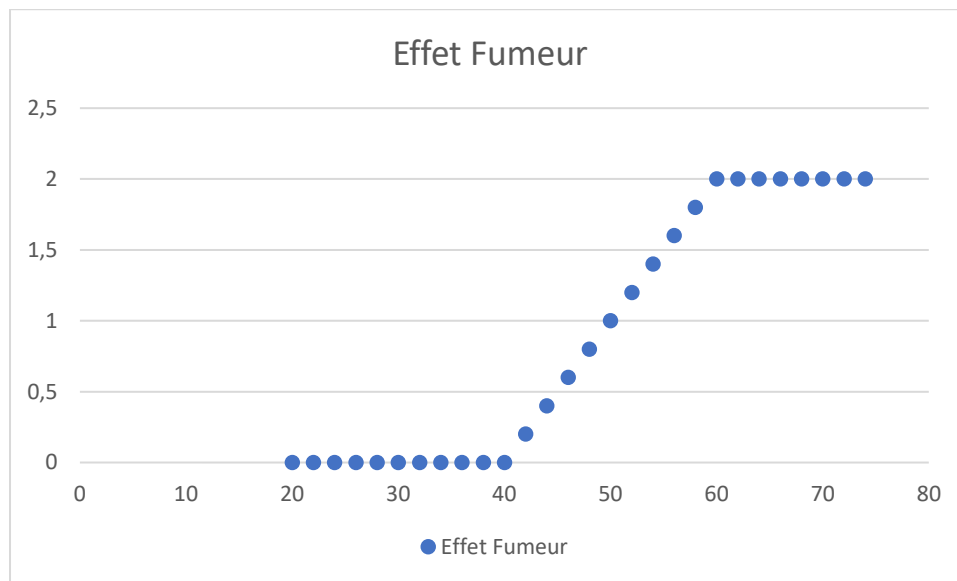
La visualisation de ces graphiques conforte deux intuitions :

- L'effet du réseau de distribution sur le taux de mortalité n'est probablement pas simple à appréhender car il conjugue des effets contraires et est corrélé avec d'autres facteurs,
- L'usage de la cigarette a peu de répercussions à un âge jeune et davantage à un âge plus avancé (on observe sur le graphique des résidus de Schoenfeld la tendance temporelle suivante : nombreux résidus du même signe pour les âges plus élevés)

Nous cherchons les moyens de modéliser de manière plus robuste l'effet du caractère fumeur / non fumeur de chaque assuré. Pour cela, nous allons partitionner le temps et introduire une nouvelle variable synthétique qui vaut 0 pour les âges jeunes quel que soit le comportement face à la cigarette, puis évolue à partir de 40 ans : restant à 0 si la personne est non fumeuse, augmentant vers 2 si la personne est fumeuse :

```
CoxPortfolio$Fumeur.age=((CoxPortfolio$BegAge>40)*(CoxPortfolio$BegAge<60)*(CoxPortfolio$BegAge-40)/(60-40)+(CoxPortfolio$BegAge>=60))*as.numeric(CoxPortfolio$Fumeur)
```

Figure 27 : évolution de la variable synthétique Fumeur.age pour un fumeur selon son âge



Nous allons également tenter d'améliorer la modélisation de l'effet de la sélection médicale. L'expérience de souscripteurs médicaux laisse penser que l'effet de la sélection médicale n'est plus visible sur la sinistralité au-delà de 3 ans, aussi nous capons cette variable à 3.

```
CoxPortfolio$SM[CoxPortfolio$SM>3]<-3
```

Concernant l'effet du réseau de distribution, qui n'est pas une variable significative et ne respecte pas l'hypothèse de proportionnalité, nous tentons une modélisation en considérant le mode de distribution seulement (c'est-à-dire en regroupant les différents courtiers en un mode de distribution Courtage).

Figure 28 : Sortie R de la nouvelle simulation du modèle de Cox appliqué au portefeuille Emprunteur étudié, avec une variable mode de distribution

```
Call:
coxph(formula = Surv(BegAge, EndAge, Status, type = "counting") ~
      Gender + Fumeur.age + seg_cap + SM + Distribution, data = CoxPortfolio)

n= 1269125, number of events= 781
```

	coef	exp(coef)	se(coef)	z	Pr(> z)	
GenderMale	0.74895	2.11478	0.08277	9.048	2,00E-16	***
GenderUnknown	0.42322	1.52686	0.27960	1.514	0.130113	
Fumeur.age	0.62105	1.86088	0.08657	7.174	7.29e-13	***
seg_cap	-0.10571	0.89969	0.03187	-3.317	0.000909	***
SM	0.20046	1.22197	0.03562	5.628	1.82e-08	***
DistributionCourtage	0.02468	1.02499	0.10814	0.228	0.819444	
DistributionInternet	-0.66640	0.51355	0.21299	-3.129	0.001755	**

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

      exp(coef)  exp(-coef)  lower .95  upper .95
GenderMale      2.1148      0.4729      1.7981      2.4873
GenderUnknown   1.5269      0.6549      0.8827      2.6412
Fumeur.age      1.8609      0.5374      1.5705      2.2050
seg_cap         0.8997      1.1115      0.8452      0.9577
SM              1.2220      0.8184      1.1396      1.3103
DistributionCourtage 1.0250      0.9756      0.8292      1.2670
DistributionInternet 0.5136      1.9472      0.3383      0.7796
```


Concordance= 0.642 (se = 0.013)
 Likelihood ratio test= 206.2 on 7 df, p=<2e-16
 Wald test = 202 on 7 df, p=<2e-16
 Score (logrank) test = 207.5 on 7 df, p=<2e-16

Figure 29 : sortie R du test de validation de l'hypothèse de proportionnalité des risques, avec une variable mode de distribution

	chisq	df	p
Gender	0.22995	2	0.89
Fumeur.age	0.00109	1	0.97
seg_cap	1.17926	1	0.28
SM	0.08436	1	0.77
Distribution	1.37864	2	0.50
GLOBAL	2.67101	7	0.91

La variable mode de distribution ne respecte toujours pas l'hypothèse de proportionnalité du modèle de Cox. Nous renonçons donc à utiliser cette variable dans notre simulation.

La modélisation ci-dessus montre par ailleurs que l'effet du capital emprunté n'est pas significatif. Nous retirons donc cette variable.

Enfin, pour tenir compte de l'absence de données fiable sur le caractère fumeur/non fumeur au sein du portefeuille, nous fixons de manière arbitraire le paramètre fumeur à 1 pour les personnes pour lesquelles l'information n'est pas connue.

La modélisation finale donne enfin le résultat recherché :

Figure 30 : Sortie R de la simulation finale du modèle de Cox appliqué au portefeuille Emprunteur étudié, avec l'ensemble des variables retenues

```
Call:
coxph(formula = Surv(BegAge, EndAge, Status, type = "counting") ~
      Gender + Fumeur.age + SM, data = CoxPortfolio)

n= 1269125, number of events= 781

      coef      exp(coef)    se(coef)      z      Pr(>|z|)
GenderMale  0.69767      2.00907    0.08225     8.483  2.00E-16 ***
GenderUnknown -0.30967      0.73369    0.21114    -1.467  0.142
Fumeur.age  0.63040      1.87836    0.08481     7.433  1.06E-13 ***
SM          0.20277      1.22480    0.03525     5.752  8.81E-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

      exp(coef)  exp(-coef)  lower .95  upper .95
GenderMale    2.0091    0.4977    1.7100    2.361
GenderUnknown 0.7337    1.3630    0.4851    1.110
Fumeur.age    1.8784    0.5324    1.5907    2.218
SM            1.2248    0.8165    1.1430    1.312
```

Concordance= 0.629 (se = 0.013)
 Likelihood ratio test= 175.4 on 4 df, p=<2e-16
 Wald test = 170.9 on 4 df, p=<2e-16
 Score (logrank) test = 175.2 on 4 df, p=<2e-16

Le test de validation de l'hypothèse de proportionnalité des risques est concluant :

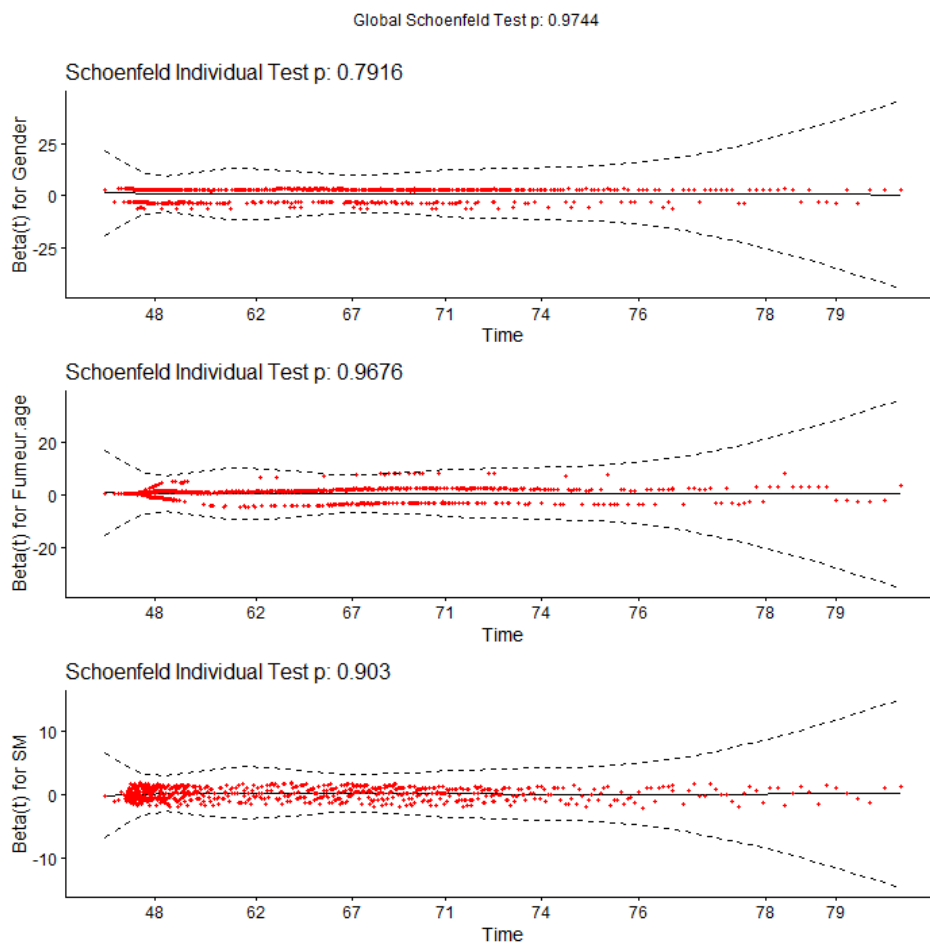
Figure 31 : sortie R du test de validation de l'hypothèse de proportionnalité des risques

	chisq	df	p
Gender	0.46746	2	0.79
Fumeur.age	0.00165	1	0.97
SM	0.01484	1	0.90
GLOBAL	0.49093	4	0.97

Ceci est confirmé par l'observation des graphiques des résidus de Schoenfeld : les résidus représentés ne présentent pas de tendance temporelle décelable à l'œil nu, et la p-value calculée vient confirmer cela.

Pour tenir compte des volumes présents dans le portefeuille, nous choisirons dans les parties suivantes de déterminer la fonction de hasard de base sur le périmètre des hommes non fumeurs dans l'année suivant la sélection médicale. Concrètement, cela revient à utiliser les coefficients issus de la sortie R ci-dessus (cf. Figure 30), à l'exception du coefficient Femme qui est déduit par inversion du coefficient Homme.

Figure 32 : sortie R des résidus du test de Schoenfeld sur la proportionnalité des risques dans le modèle de Cox appliqué aux variables sélectionnées



2.3.2 Positionnement par rapport à la table TH

2.3.2.1 Présentation de la modélisation : principe et périmètre

Dans cette seconde étape, nous allons chercher à modéliser la fonction de hasard de base, qui correspond à la mortalité du périmètre de portefeuille correspondant au premier niveau des variables retenues. Pour tenir compte des volumes présents dans le portefeuille, nous choisissons que la fonction de hasard de base (à laquelle les coefficients calculés par le modèle de Cox seront appliqués) sera basée sur le périmètre des hommes non fumeurs dans l'année suivant la sélection médicale.

Nous commençons par observer ce périmètre graphiquement en représentant son effectif, les nombres de décès observés et les taux de mortalité observés.

Figure 33 : Effectifs et nombre de morts sur le périmètre de portefeuille utilisé pour la modélisation de la fonction de hasard de base

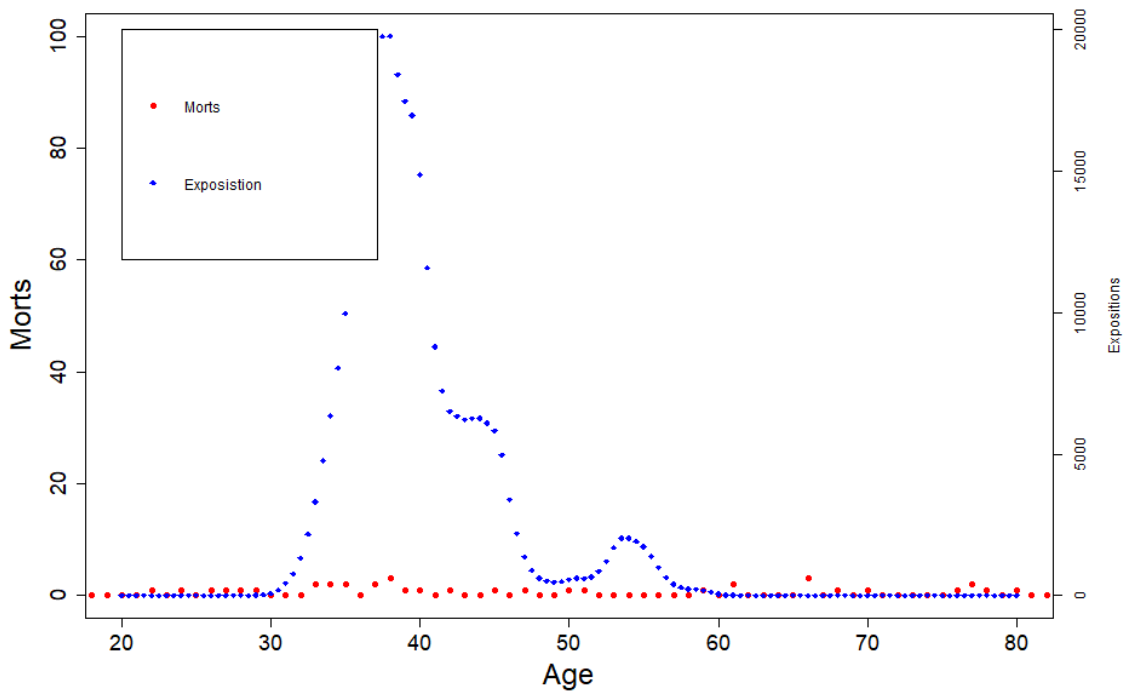
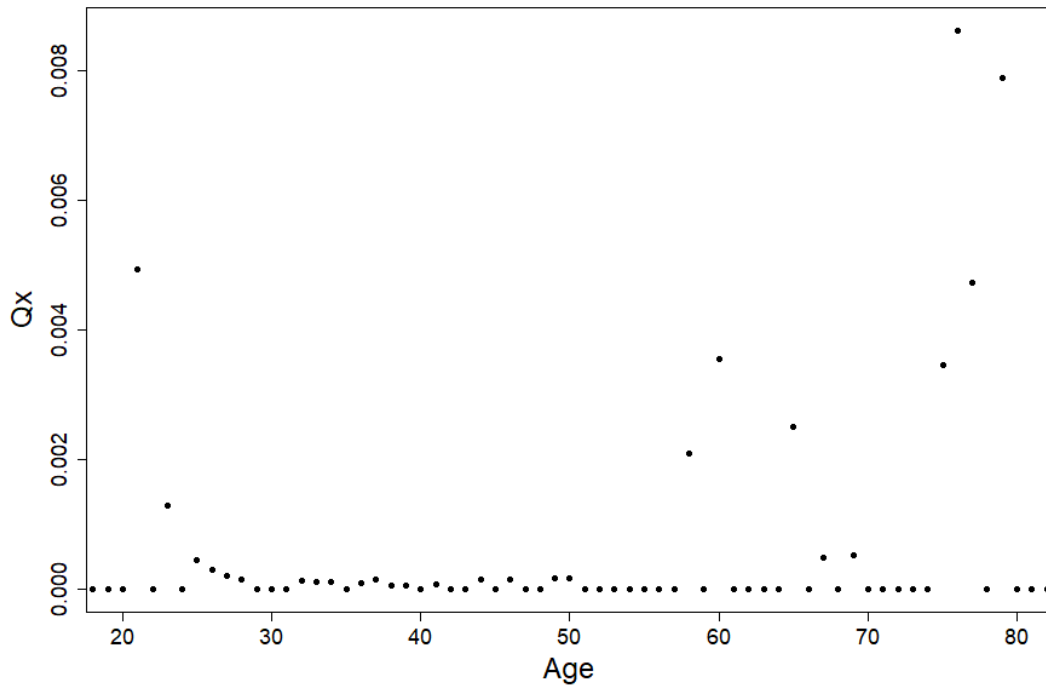


Figure 34 : probabilités de décès sur le périmètre de portefeuille utilisé pour la modélisation de la fonction de hasard de base



L'échantillon en question est trop peu représentatif pour l'utiliser comme seule base de travail d'un lissage, aussi nous allons :

- Etendre l'échantillon en cherchant à modéliser globalement la mortalité des hommes du portefeuille étudié. Cette approche est prudente dans la mesure où le portefeuille total comprend les personnes dont la mortalité est plus élevée (plus grande antériorité de la sélection médicale, caractère fumeur) mais pas si éloignée de la mortalité de référence, compte tenu de la jeunesse du portefeuille et de sa composition,
- Positionner la mortalité du portefeuille étudié par rapport à une table de référence.

Pour identifier à quel point cette approche est prudente, nous observons les mêmes représentations graphiques sur le périmètre élargi.

Figure 35 : Effectifs et nombre de morts sur le périmètre de portefeuille utilisé pour la modélisation de la fonction de hasard de base.

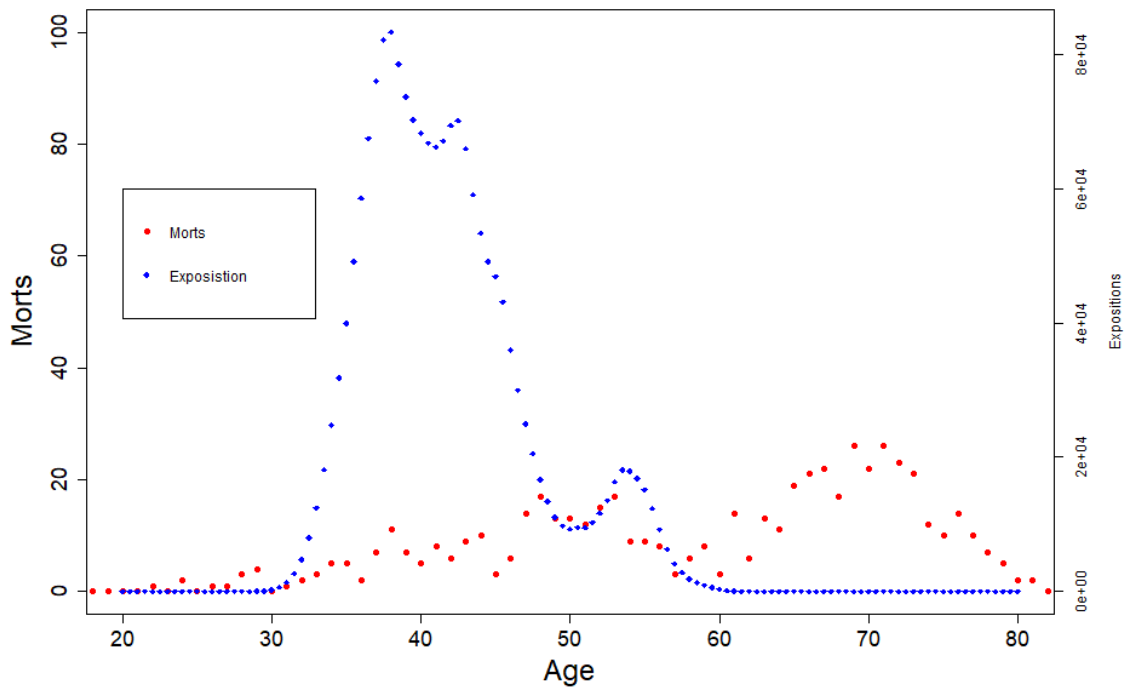
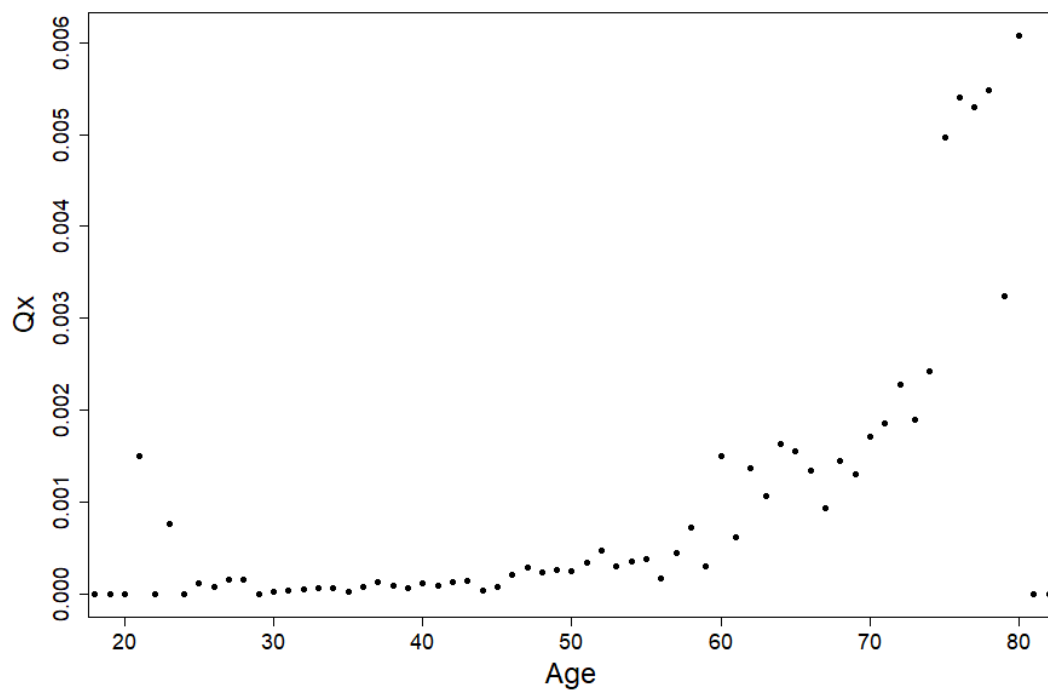


Figure 36 : probabilités de décès sur le périmètre de portefeuille utilisé pour la modélisation de la fonction de hasard de base



La représentation graphique de la probabilité de décès observée sur le portefeuille élargi présente une allure beaucoup plus conforme à ce à quoi l'on s'attend, avec en particulier une diminution des artefacts causés par quelques décès locaux.

Par ailleurs, on va vérifier la proximité entre le portefeuille observé et le périmètre restreint d'observation via le calcul des moyennes des variables alimentant le modèle de Cox (on s'attend à cette proximité en raison de la jeunesse du portefeuille et du profil des clients) :

BasePortfolio\$Fumeur.age vaut 0 dans la fonction de hasard de base et évolue au sein du portefeuille entre 0 et 2 selon l'âge de la personne. Son calcul sur le portefeuille vaut :

```
[1] "mean(BasePortfolio$Fumeur.age[BasePortfolio$Gender==Male]) = 0.155354250420078"
```

On peut donc estimer que l'utilisation du portefeuille total correspond à une prudence de $\exp(0.1553542) \sim 115\%$ (surestimation de la fonction de décès de 15%).

BasePortfolio\$SM vaut 0 pour la fonction de base et évolue au sein du portefeuille entre 0 et 3 selon l'ancienneté de la sélection médicale. Les valeurs moyennes sur le portefeuille sont :

```
[1] "mean(BasePortfolio$SM[BasePortfolio$Gender==Male]) = 1.30762164751686"
```

Cela signifie donc que l'estimation de la mortalité sera très prudente.

Pour positionner la table, nous nous appuyons sur les éléments publiés par le groupe de travail Mortalité de l'Institut des Actuaraires (cf. notes de travail II1291-11 à 14) qui comprennent notamment des méthodologies de positionnement et les critères de validation adéquats. Pour les justifications théoriques de ces éléments, nous renvoyons aux articles classiques, dont Tomas, J. , Planchet, F. , *Prospective mortality tables and portfolio experience*, in Computational Actuarial Science, with R ; Arthur Charpentier Editor, Chapman, 2014 et les différents modules de cours dispensés dans les écoles d'actuariat. Les modélisations sont faites en utilisant le package R *ELT* développé par Julien Tomas, Frederic Planchet, Wassim Youssef pour mettre en application les méthodologies développées dans leurs articles.

Nous décidons de positionner par rapport à la table de référence la fonction de hasard de base, qui correspond à la mortalité des hommes non fumeurs dans l'année suivant la sélection médicale. Le positionnement est effectué par rapport à la table TH-002D et avec les 4 méthodes proposées par la note de l'institut.

Les résultats issus des différentes méthodes sont ensuite comparés au travers des tests classiques quantitatifs et visuels.

2.3.2.2 Positionnement par rapport à la table de référence TH-002

La première méthode proposée par le groupe de travail est une approche à un paramètre avec le SMR (application d'un simple coefficient de réduction / majoration, appliqué de manière multiplicative aux probabilités conditionnelles de décès de la référence) :

$$\tilde{q}_x(t) = SMR \times q_x^{ref}(t), \quad \text{avec } SMR = \frac{\sum_{(x,t)} D_{x,t}}{\sum_{(x,t)} E_{x,t} q_x^{ref}(t)}.$$

Comme on s'y attend, le SMR est très faible. La sortie R suivante montre un coefficient SMR de 5,1% de la loi de mortalité de référence :

Figure 37 : Sortie R de l'exécution de la méthode 1

```
[1] "Execute method 1 ..."  
[1] "Compute the SMR ..."  
[1] "QxtFittedMale = 0.0514691956386396 * QxtRefMale"
```

Ce coefficient très bas est cohérent avec le fait connu qu'en assurance emprunteur la mortalité est très réduite en comparaison de la table réglementaire, ce à quoi d'ajoute que notre estimation porte sur une mortalité peu après la sélection médicale, qui a un effet sur celle-ci.

La seconde méthode utilisée est une variante semi-paramétrique du modèle de Brass. Le positionnement est effectué à l'aide de la fonction logistique,

$$\text{logit } \tilde{q}_{x^*}(t^*) = \alpha + \beta \text{logit } q_{x^*}^{ref}(t^*) + e_{x^*,t^*},$$

où x^* et t^* correspondent respectivement à la plage d'âges et période d'observation communes avec la référence de mortalité, e_{x^*,t^*} est le terme d'erreur et $q_{x^*}^{ref}(t^*)$ est la référence de mortalité. L'estimation est effectuée par une minimisation de la distance entre les décès estimés et observés :

$$\sum |E_{x,t}(\tilde{q}_x(t) - \hat{q}_x(t))|.$$

La sortie R suivante donne les coefficients d'interpolation $\alpha = -1,54$ et $\beta = 1,29$ du logit de la fonction de mortalité de référence pour obtenir le logit de la mortalité estimée :

Figure 38 : Sortie R de l'exécution de la méthode 2

```
[1] "Execute method 2 ..."  
[1] "Compute the parameters of the semi-parametric relational model ..."  
[1] "logit (QxtFittedMale) = -1.53782885381019 + 1.29021634742726 * logit (QxtRefMale)"
```

La troisième méthode utilisée est un modèle linéaire généralisé de Poisson, dans lequel on inclut la mortalité de référence et on permet des interactions avec l'âge. On suppose que le nombre de décès de l'organisme à l'âge x^* et l'année calendaire t^* est déterminé par :

$$\mathcal{D}_{x^*,t^*} \sim \mathcal{P}(E_{x^*,t^*} \cdot \mu_{x^*}(t^*)),$$

avec

$$\mu_{x^*}(t^*) = \beta_0 + \beta_1 \log \tilde{q}_{x^*}^{ref}(t^*) + \beta_2 x^*.$$

La sortie R suivante donne les coefficients d'interpolation $\beta_0 = 10,9$, $\beta_1 = 2,5$ et $\beta_3 = -0,12$ du logit de la fonction de mortalité de référence et de l'âge pour obtenir la mortalité estimée :

Figure 39 : Sortie R de l'exécution de la méthode 3

```
[1] "Execute method 3 ..."  
[1] "Compute the parameters of the Poisson model ..."  
[1] "Summary of the Poisson model, Male population"
```

Call:

```
glm(formula = D_i ~ as.numeric(log(mu_i)) + as.numeric(Age),  
     family = poisson, data = data.frame(DB), offset = log(DB[,  
     "E_i"]))
```

Min	1Q	Median	3Q	Max
-2.7075	-0.6243	-0.3307	-0.0260	3.4946

```

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  10.92108   7.52949   1.450  0.14694
as.numeric(log(mu_i)) 2.52459   0.81625   3.093  0.00198 **
as.numeric(Age)    -0.11704   0.06507  -1.799  0.07205 .

```

```

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

(Dispersion parameter for poisson family taken to be 1)

```

Null deviance: 387.10 on 327 degrees of freedom
Residual deviance: 258.89 on 325 degrees of freedom
AIC: 558.45

```

Number of Fisher Scoring iterations: 6

Le modèle linéarisé ne donne pas des coefficients très probants (la p-value des coefficients est respectivement de 15%, 2%, 7%).

La méthode 4 est un lissage non paramétrique de la table de mortalité de référence sur lequel on applique une dérive :

$$D_{x^*,t^*} \sim \mathcal{P}(E_x q_{x^*}^{ref} \exp(f(x))),$$

Avec

$E_x q_{x^*}^{ref}$ le nombre de décès espérés de la table TH-002

f une fonction lisse de l'âge atteint (polynôme évalué localement sous une méthode de vraisemblance).

Figure 40 : Sortie R de l'exécution de la méthode 4

```

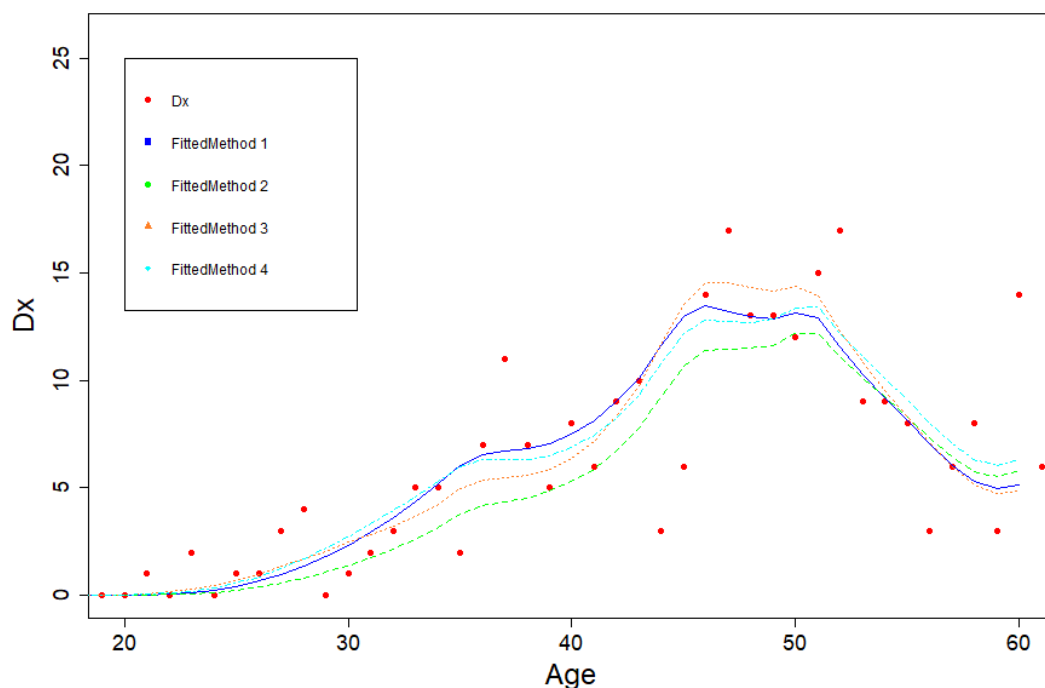
[1] "Execute method 4 Second Part - Female ..."
[1] "Non-parametric smoothing of the periodic table ..."
[1] "Optimal smoothing parameters:"
[1] "Female population:"
[1] "Bandwidth: 29 observations"
[1] "Polynomial degree: 1"
[1] "Values of the AIC: 62.5182"
[1] "Fitted DF: 2.6832"
[1] "Execute method 4 Second Part - Male ..."
[1] "Non-parametric smoothing of the periodic table ..."
[1] "Optimal smoothing parameters:"
[1] "Male population:"
[1] "Bandwidth: 33 observations"
[1] "Polynomial degree: 1"
[1] "Values of the AIC: 63.4122"
[1] "Fitted DF: 2.4314"
[1] "Create graphics of the fits before completion in ../Results/Graphics/Completion ..."

```

Ces 4 méthodes font l'objet de tests similaires pour déterminer laquelle retenir.

Une première analyse locale, graphique, compare la mortalité observée à l'ajustement.

Figure 41 : résultat des modélisations des nombres de décès Homme en fonction de l'âge par les quatre méthodes proposées par le GT Mortalité de l'Institut des Actuaires.



Cette observation graphique montre un bon ajustement et ne permet pas d'écarter au premier abord l'une des méthodes. 3 méthodes donnent des résultats très similaires, la méthode 2 s'écarte légèrement.

2.3.2.3 Validation de la modélisation

Une seconde analyse globale, quantitative, est menée pour valider la modélisation et sélectionner la méthode de positionnement à retenir via l'étude de la distance entre les observations et l'ajustement, d'une part, et l'étude de la qualité du lissage d'autre part.

Ecart entre ajustement et observation :

- On mesure d'abord la qualité de l'ajustement par la comparaison des quantités suivantes : nombre de résidus standardisés supérieurs à 2 et 3, critère du χ^2 utilisé par Forfar et al. (1988) ainsi que le *Mean Average Percentage Error* (MAPE) et R2 utilisés dans Felipe et al. (2002).
- Ensuite, on réalise des tests sur le ratio SMR et la déviance.

Qualité du lissage : on utilise le nombre de signes positifs et négatifs des résidus de la réponse, la valeur du test des signes correspondant, le nombre de runs et la valeur du test des runs correspondant, proposés par Forfar et al. (1988) et Debón et al. (2006).

Un bon ajustement, marqué par une répartition homogène des signes positifs et négatifs des résidus de la réponse et un grand nombre de runs ne doit pas résulter d'un écart trop important avec la mortalité passée, ou réciproquement. En conséquence, on cherchera toujours à équilibrer ces deux aspects complémentaires.

2.3.2.3.1 Mesure de l'écart entre ajustement et observation

a) Observation de quantités

Les quantités mesurées sont :

- le nombre de résidus standardisés supérieurs à 2 et 3, que l'on va chercher à avoir le plus faible possible,
- le $\chi^2 = \sum_{(x,t)} \frac{(D_{x,t} - E_{x,t} \tilde{q}_x(t))^2}{E_{x,t} \tilde{q}_x(t) (1 - \tilde{q}_x(t))}$, que l'on va chercher à avoir le plus faible possible,
- le *Mean Average Percentage Error MAPE* = $100 \times \frac{\sum_{(x,t)} |(D_{x,t}/E_{x,t} - \tilde{q}_x(t)) / (D_{x,t}/E_{x,t})|}{\sum_{(x,t)} D_{x,t}}$, moyenne des écarts en valeur absolue par rapport aux valeurs observées, que l'on va chercher à avoir le plus faible possible,
- le $R^2 = 1 - \left(\frac{\sum_{(x,t)} (D_{x,t}/E_{x,t} - \tilde{q}_x(t))^2}{\left(\sum_{(x,t)} (D_{x,t}/E_{x,t} - (\sum_{(x,t)} (D_{x,t}/E_{x,t})/n) \right)^2} \right)$, part de variance expliquée par rapport à la variance totale, que l'on va chercher à avoir le plus proche de 1 possible,

b) Test de la Déviance

Le premier test réalisé est celui du rapport de vraisemblance, qui va mesurer la déviance de l'ajustement par rapport aux observations. Sous l'hypothèse des décès suivant une loi de Poisson :

$$D_{x,t} \sim \mathcal{P}(E_{x,t} q_x(t)),$$

la déviance totale :

$$\sum_{(x,t)} \left(D_{x,t} \ln \frac{D_{x,t}}{E_{x,t} \tilde{q}_x(t)} - (D_{x,t} - E_{x,t} \tilde{q}_x(t)) \right),$$

suit une loi de χ^2 avec un nombre de degrés de liberté égal au nombre d'observations. On cherche à avoir la statistique de test la plus basse possible et une p-value proche de 1. Dans le tableau récapitulatif des résultats des différents tests, ce test est noté *Likelihood ratio test*.

c) Test du SMR

Le second test réalisé est celui du rapport entre le nombre de décès observés et le nombre de décès ajustés, mesuré par le ratio :

$$SMR \sim \frac{\sum_{(x,t)} D_{x,t}}{\sum_{(x,t)} E_{x,t} \tilde{q}_x(t)}.$$

Si ce ratio est différent de 1, on suit la méthodologie proposée par Liddell (1984) pour construire une statistique de test

$$\xi^{SMR} = 3 \times D^{1/2} \left(1 - (9D)^{-1} - \left(\frac{D}{E} \right)^{\frac{1}{3}} \right),$$

où $D = \sum_{(x,t)} D_{x,t}$ et $E = \sum_{(x,t)} E_{x,t} \tilde{q}_x(t)$.

ξ^{SMR} suit une loi normale centrée réduite sous l'hypothèse \mathcal{H}_0 . On cherche à obtenir la p-value la plus proche de 1.

Les résultats de ces tests et de la mesure de ces quantités sont présentés ci-dessous.

Figure 42 : Quantités mesurées et test de validation globale des modélisations par les 4 méthodes envisagées

<i>Likelihood ratio test</i>	Method1	Method2	Method3	Method4
Xi	254.64	3317	251.75	245.92
Threshold	371.23	371.23	371.23	371.23
Hyp	H0	H1	H0	H0
p.val	0.999	0	0.9994	0.9998

SMR test	Method1	Method2	Method3	Method4
SMR	1	4327189.0451	1	0.9638
Xi	0.0337	29.4826	0.0673	0.3003
Threshold	1.6449	1.6449	1.6449	1.6449
Hyp	H0	H1	H0	H0

<i>Standardized residuals</i>	Method1	Method2	Method3	Method4
Std. Res. > 2	17	66	18	15
Std. Res. > 3	5	66	5	4

<i>Quantities</i>	Method1	Method2	Method3	Method4
Chi2	556,42	2,47E+11	515,46	414,34
R2	0,0141	-0,0472	0,0144	0,0082
MAPE	55,16	100	56,39	54,71
Deviance	509,28	6634,01	503,49	491,83

Seule la méthode 2 (ajustement logistique) présente des résultats très éloignés de l'observation. Les quantités et tests menés confirment la bonne adéquation des méthodes 1,3,4 et leur proximité entre elles par l'observation de p-value proche de 1 pour le test de la déviance, un SMR très proche de 1, et un nombre de résidus standardisés supérieurs à 2 ou 3 très similaire entre les trois méthodes.

Il faut toutefois noter que les méthodes 1 et 3 sont très similaires dans leurs résultats tels que résumés par les quantités et statistiques de tests. La méthode 4 présente quant à elle une déviance et une statistique de χ^2 plus faible, mais un SMR plus bas (signe d'une sous-estimation de la mortalité).

2.3.2.3.2 Tests de la qualité du lissage

La qualité du lissage est mesurée par le test des signes et le test des runs.

a) *Test des signes*

La fréquence des changements de signes de la différence entre la probabilité de décès observée et ajustée est mesurée par la statistique de test $\xi^{SIG} = \frac{|n^+ - n^-| - 1}{\sqrt{n}}$, qui suit sous l'hypothèse \mathcal{H}_0 une loi normale centrée réduite. On privilégie le modèle pour lequel la p -valeur est la plus proche de 1.

b) *Test des runs*

Ce test s'applique à la séquence des signes des différences entre les probabilités de décès observée et ajustée : il cherche à identifier si des runs (suite de signes identiques) non indépendants sont présents dans la séquence obtenue. Sous l'hypothèse \mathcal{H}_0 , sachant le nombre de signes positifs n^+ et négatifs n^- , le nombre de runs suit une loi normale. La statistique :

$$\xi^{RUN} = \frac{\text{Nombre de runs} - \left(\frac{2n^+n^-}{n^+ + n^-} + 1 \right)}{\sqrt{\frac{2n^+n^-(2n^+n^- - (n^+ + n^-))}{(n^+ + n^-)^2(n^+ + n^- - 1)}}},$$

suit une loi normale centrée réduite. On privilégiera donc le modèle avec la p -valeur la plus proche de 1.

Figure 43 : Tests de validation globale de la qualité du lissage par les 4 méthodes envisagées

<i>Signs test</i>	Method1	Method2	Method3	Method4
Signs (+)	81	85	80	79
Signs (-)	247	243	248	249
Xi	9.1106	8.6689	9.221	9.3315
Threshold	1.96	1.96	1.96	1.96
Hyp	H1	H1	H1	H1
p.val	0	0	0	0

<i>Runs test</i>	Method1	Method2	Method3	Method4
Nber of runs	104	110	106	108
Signs (-)	247	243	248	249
Signs (+)	81	85	80	79
Xi (abs)	2.8271	2.4427	2.3979	1.9598
Threshold	1.96	1.96	1.96	1.96
Hyp	H1	H1	H1	H0
p.val	0.0047	0.0146	0.0165	0.05

Aucune des méthodes ne propose de résultat concluant mesuré par ces tests, toutes présentent des résultats très proches. Le nombre de signes négatifs (probabilité de décès observée < probabilité de décès estimée) est très supérieur au nombre de signes positifs (probabilité de décès observée > probabilité de décès estimée), ce que, dans le cadre où le SMR est proche de 1, on peut interpréter de la façon suivante : le nombre de décès globalement estimé est bon, ce qui est la somme entre de nombreuses et légères surestimation, et de moins nombreuses mais plus lourdes sous-estimations.

Visuellement, cette interprétation peut s'observer via le tracé des résidus présenté en Annexe 9.

2.3.2.4 Modélisation retenue

Compte tenu des résultats obtenus et en l'absence d'élément conduisant à privilégier l'une des méthodes plus complexes au détriment de la première méthode de modélisation testée, nous retenons comme modélisation la méthode 1, c'est-à-dire le positionnement par un paramètre de SMR.

La loi de décès de base de notre modèle de Cox sera donc issue de la TH-002, multipliée par un coefficient de 0,0515, qui correspond au coefficient SMR issu de l'application de la méthode 1 (cf. Figure 37).

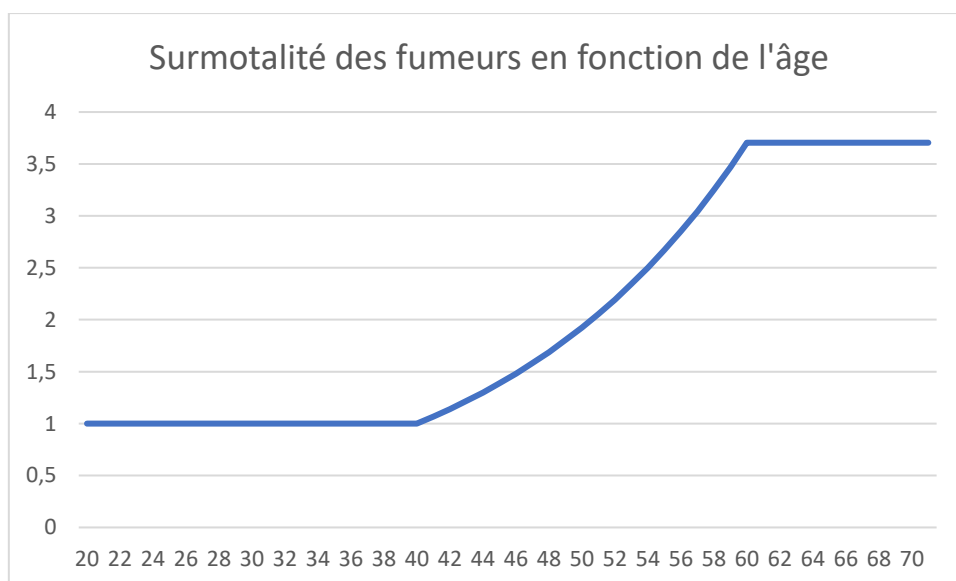
2.3.3 Présentation de la loi de décès simulée

Au total, nous avons donc modélisé une loi de décès constituée d'une fonction de hasard de base (positionnée en référence à la table TH002-D) et des coefficients de majoration, correspondant aux caractéristiques de l'assuré ou du prêt couvert.

Les résultats en sont les suivants :

- La mortalité choisie comme base est celle des hommes non fumeurs dans l'année suivant la sélection médicale : elle est égale à la mortalité prévue par la TH002-D multipliée par un coefficient de 0,0515.
- La mortalité des autres individus se déduit par l'application des coefficients suivants :
 - o La mortalité des Femmes est obtenue en appliquant un coefficient de 0.4 à la mortalité des Hommes.
 - o L'effet de l'habitude de fumer sur la mortalité est calculé conformément à la variable synthétique présentée plus haut : avant 40 ans, la mortalité des fumeurs et des non fumeurs ne présente pas de différence, au-delà de 60 ans, la mortalité des fumeurs est obtenue en appliquant un coefficient de 3,7. Entre 40 ans et 60 ans, l'évolution du coefficient est exponentielle.
 - o L'effet de la sélection médicale sur la mortalité s'estompe au cours du temps, avec une augmentation de la mortalité de 22% par année jusqu'à la troisième année (+83% de mortalité à partir de la troisième année suivant la sélection médicale). Présenté autrement, cela signifie que la sélection médicale a pour effet de réduire la mortalité de 45% en année 1, 33% en année 2 et 18% en année 3.

Figure 44 : Modélisation de la surmortalité des fumeurs en comparaison des non fumeurs en fonction de l'âge



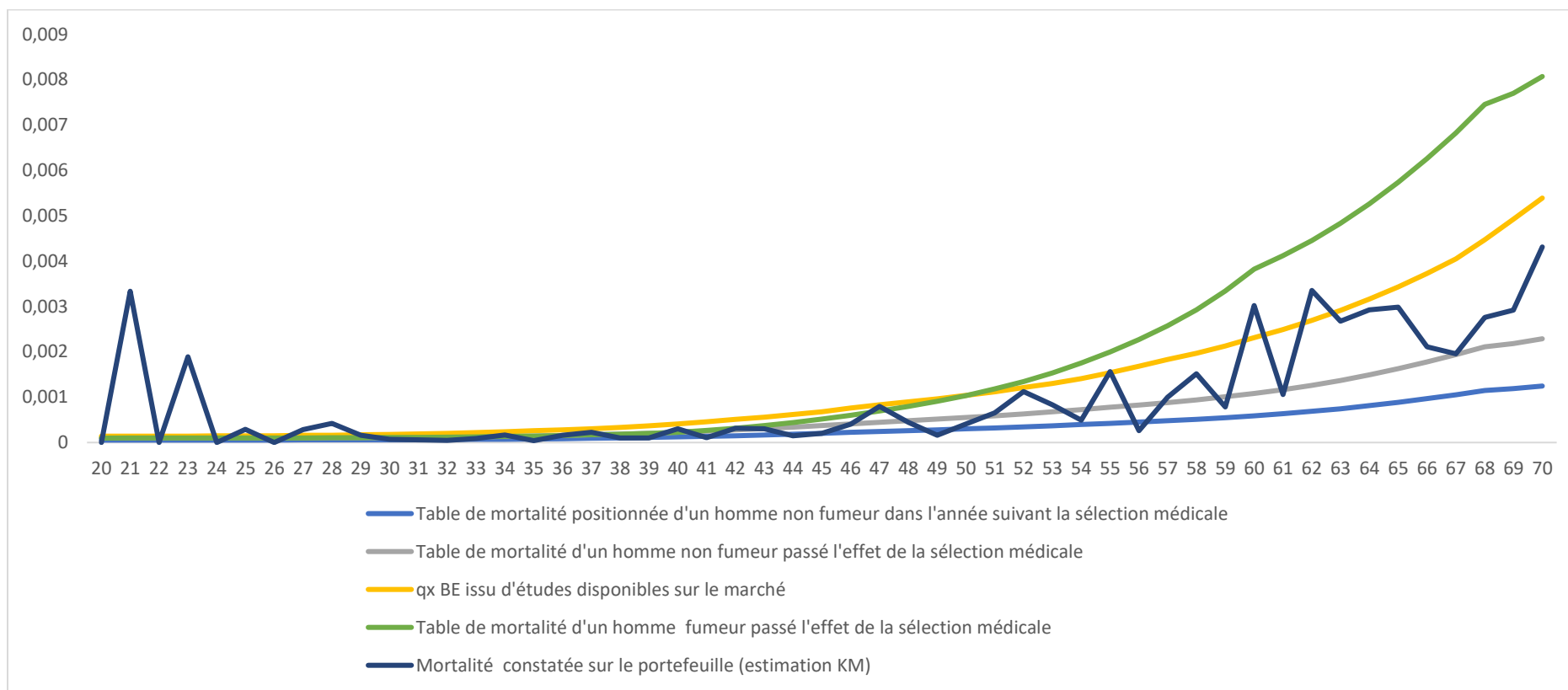
Pour vérifier les résultats, nous traçons les mortalités simulées et les comparons :

- A une mortalité *Best Estimate* (BE) issue de bases communément utilisées des consultants intervenants sur le marché de l'Emprunteur,
- A la mortalité moyenne observée sur le portefeuille Emprunteur objet de l'étude.

Cette vérification graphique montre une bonne adéquation de notre modélisation :

- Le q_x BE du marché se situe dans le range des lois de décès modélisées selon le profil des clients,
- La mortalité moyenne du portefeuille est située entre la mortalité de première année et la mortalité passé l'effet de la sélection médicale, avec une position plus proche de la mortalité de première année, ce qui est cohérent avec la taille de notre portefeuille et sa composition.

Figure 45 : modélisation de la loi de décès du portefeuille selon les caractéristiques des individus



2.4 Modélisation de la loi de résiliation actuelle

L'objectif de cette partie est la modélisation de la loi de résiliation du portefeuille. Pour cela, nous souhaitons calculer les taux annuels de chute pour chaque individu en fonction de ses caractéristiques propres et des caractéristiques du prêt.

Cette modélisation fait écho à la modélisation précédente et s'appuie sur les mêmes bases théoriques des modèles de durée (dans cette partie, l'évènement observé n'est plus le décès de l'assuré, mais l'acte de résiliation du contrat).

Nous procéderons de la manière suivante :

- Nous analyserons d'abord de manière qualitative les causes de la résiliation d'un contrat,
- Nous estimerons à l'aide d'un estimateur de Kaplan-Meier la loi de durée des contrats et la loi de résiliation,
- Nous étudierons graphiquement les variables expliquant l'hétérogénéité des observations,
- Nous utiliserons deux modèles : d'abord un modèle de Cox, puis un modèle par arbre.

2.4.1 Analyse des motifs de résiliation

Dans notre portefeuille, la variable renseignant sur la résiliation est le **statut** du contrat, pour laquelle trois modalités existent :

- **En vigueur** : le contrat de l'assuré est toujours en cours et ce dernier est couvert en cas de sinistre pour les garanties souscrites,
- **Sans suite** : une souscription a été entamée mais n'a jamais été aboutie pour diverses raisons ou encore annulée suite à la renonciation de l'assuré,
- **Résilié** : l'assuré ou l'assureur a mis fin au contrat d'assurance.

Dans le cadre de nos travaux, c'est ce dernier statut qui nous intéresse. Nous disposons dans nos bases d'information fréquente sur la cause de la résiliation : le tableau suivant répertorie les différents motifs de résiliation renseignés pour les 73 622 contrats notés comme « résilié » dans notre base.

Figure 46 - Motifs des contrats résiliés

	Nb contrats	% contrats
Remboursement anticipé	46 386	63%
Prêt arrivé au terme	10 857	15%
Impayé	7 886	11%
Motif non renseigné	3 791	5%
Changement d'assurance	3 738	5%
Sinistre décès	964	1%
TOTAL	73 622	100%

Nous constatons que la principale cause de résiliation est le remboursement anticipé du crédit. Le remboursement anticipé signifie que l'assuré solde son emprunt auprès de sa banque avant la date initiale prévue au contrat de prêt. Le remboursement anticipé a trois origines possibles :

- Le remboursement du prêt via des liquidités dont dispose l'emprunteur,
- Le remboursement du prêt dans le cadre d'une renégociation (en particulier dans les périodes de baisse des taux), financé par la souscription d'un nouveau prêt auprès d'un autre organisme prêteur,

- Le remboursement du prêt dans le cadre d'un changement de bien immobilier, financé par la vente du bien. Cet évènement est éventuellement suivi de la souscription d'un nouveau prêt auprès du même organisme ou d'un autre organisme, pour financer l'achat d'un nouveau bien. Dans ce cadre, un nouveau contrat d'assurance est également souscrit concomitamment à l'obtention du nouveau prêt.

Ces deux derniers évènements font l'objet de notre étude : la transférabilité du contrat d'assurance permettrait d'éviter sa résiliation, la couverture étant maintenue dans le cadre du nouveau prêt.

Deux autres motifs de résiliation sont prépondérants :

- Les prêts arrivés à terme : ce sont les prêts qui ont été entièrement remboursés selon les conditions prévues au contrat. Ce motif est considéré comme une résiliation dans notre base portefeuille, mais n'en est pas une à proprement parler. Nous notons ces contrats plutôt comme « terminés »,
- Les contrats résiliés pour sinistre décès : la totalité du capital restant dû a été versé au bénéficiaire, le contrat prend évidemment fin.

Dans le cadre de la construction de la loi de résiliation, nous faisons le choix de ne pas considérer les décès et fins de prêts comme des chutes de contrat puisque nous les modélisons par ailleurs (respectivement via la loi de mortalité et via la prise en compte de la durée des prêts dans nos modélisation de compte de résultat). Cependant nous utilisons ces contrats dans notre modélisation en les assimilant à des données censurées quand cela est pertinent.

De plus nous retirons complètement de la base de données les contrats sans suite (date de fin égale à la date d'effet) et supprimons les lignes pour lesquelles la date de fin est antérieure à la date d'effet.

Ainsi nous obtenons une base portefeuille qui se décompose comme suit :

Figure 47 - Nombre de contrats par statuts

	Nb contrats	% contrats
En vigueur	451 497	86%
Résilié	59 948	11%
Terminé	11 415	2%
TOTAL	522 860	100%

2.4.2 Estimation des taux bruts

L'évènement que nous cherchons à observer est la résiliation du contrat d'assurance pour chaque année d'ancienneté. L'ancienneté correspondant au temps écoulé entre la date d'entrée et la date de fin d'observation du contrat dans le portefeuille.

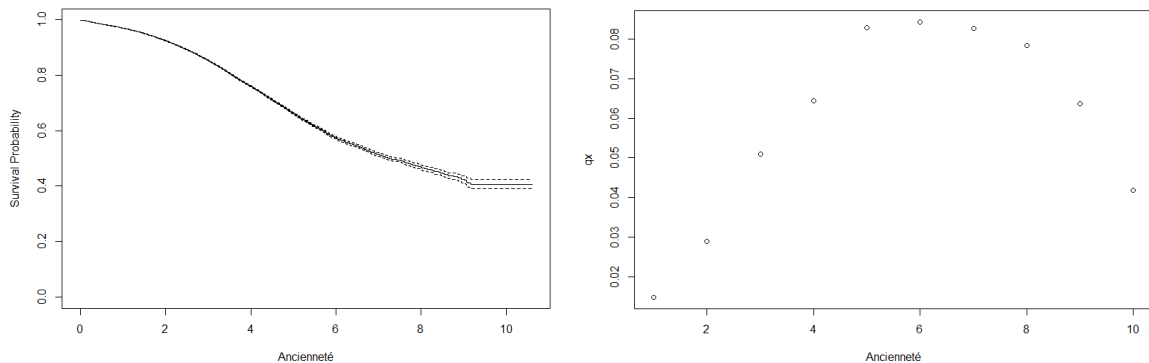
De manière analogue à une table de mortalité en fonction de l'âge, nous estimons cette fois une loi de survie du contrat en fonction de son ancienneté. Ainsi nous choisissons d'utiliser de nouveau l'estimateur de Kaplan-Meier pour calculer les taux bruts de résiliation.

Sur la période d'observation 2009-2019 l'évènement de résiliation ne s'est pas réalisé pour tous les contrats. En effet sur 522 860 contrats nous n'avons observé que 59 948 résiliations. Ceux pour lesquels l'évènement ne s'est pas produit sont soit terminés soit toujours en cours. Dans le cas des contrats terminés, ces derniers sont sortis du portefeuille avant que l'évènement se produise tandis que ceux en cours n'ont pas été résiliés avant la date de fin d'observation. Dans la théorie des modèles de durée ce phénomène est défini comme la censure à droite.

Nous créons alors dans notre base de données une nouvelle variable nommée « Censure » qui vaut 1 pour les contrats résiliés et 0 pour les contrats terminés ou en vigueur. Cette variable est essentielle à notre estimation et sert d'input à la fonction R *survfit* qui calcule les taux bruts par la méthode Kaplan-Meier.

Regardons graphiquement à quoi ressemblent la fonction de survie et les taux bruts de résiliations estimés par la méthode de Kaplan-Meier renvoyés par la fonction *survfit* :

Figure 48 - Fonction de survie et taux bruts de résiliation



Notons qu'en ligne pleine nous avons la loi de survie estimée par la méthode Kaplan-Meier et en pointillé l'intervalle de confiance associé.

A l'inverse de l'étude des taux de moralité, nous constatons ici que les taux de résiliation ne sont pas toujours croissants avec l'âge du contrat. En effet à partir de la 6^{ème} année d'ancienneté on commence à observer une baisse notable des résiliations. Cette courbe n'est pas forcément intuitive car on aurait pu penser qu'avec la mise en place de la loi Hamon les résiliations en première année seraient plus importantes or ce n'est pas le cas. Il est vrai que l'entrée en vigueur de cette loi est assez récente par rapport à l'ancienneté de notre portefeuille et par construction de l'estimateur de Kaplan-Meier l'exposition des contrats en 1^{ère} année confond toutes les générations de souscription. De ce fait l'impact est difficilement dissociable.

Le résultat de l'estimation de la loi de survie d'un contrat en fonction de son ancienneté a été obtenu sans segmentation particulière sur les caractéristiques de l'assuré ou du prêt. Cependant nous avons à notre disposition un ensemble d'informations qui peuvent justement nous aider à affiner cette estimation. Ainsi c'est ce que nous tâchons d'étudier dans la section suivante afin d'identifier les variables qui pourraient être explicatives de la résiliation.

2.4.3 Recherche graphique des paramètres d'hétérogénéité

Comme présenté dans la partie 2.1, nous disposons d'un certain nombre d'informations sur l'assuré et son/ses prêt(s). L'identification d'une hétérogénéité dans la population nous semble être une étape essentielle pour la poursuite de nos travaux, dans la mesure où elle permettra de segmenter finement notre approche tarifaire pour optimiser notre rentabilité.

Les différentes intuitions que nous avons quant à l'influence des variables sur la loi de résiliation sont les suivantes :

- Le sexe de l'assuré ne devrait pas jouer significativement : on ne voit pas de raison poussant une femme à résilier davantage qu'un homme ou vice versa.
- Le caractère fumeur de l'assuré pourrait avoir un impact sur sa loi de résiliation si un changement de comportement (arrêt de la cigarette) lui donne accès à des tarifs plus avantageux dans le cadre de la résiliation Hamon ou Bourquin. Compte tenu de la nouveauté de ces

dispositifs, nous ne pouvons pas observer cela et nous ne nous attendons donc pas à identifier d'impact de cette variable.

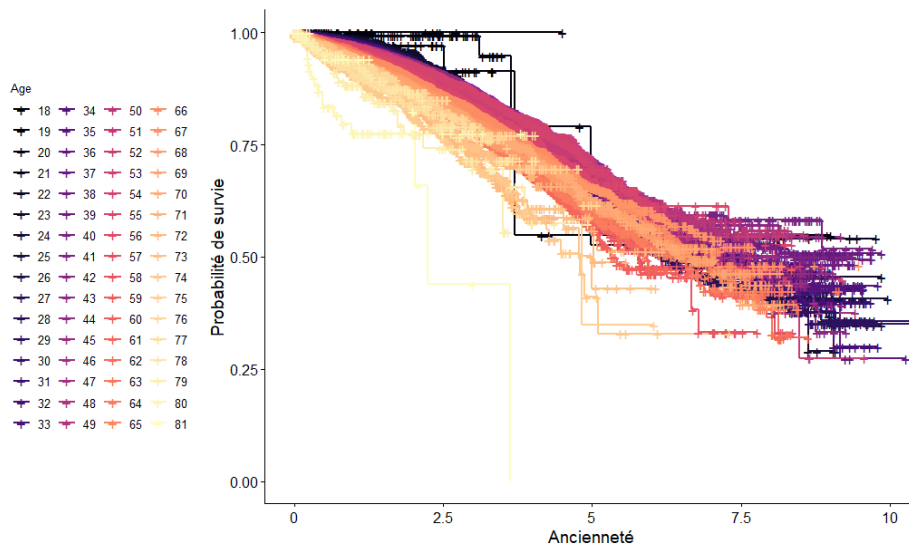
- Le tarif et le reste à payer pourraient avoir un impact sur l'acte de résiliation, cependant faute de données exhaustives sur ces variables cela serait difficile à estimer. En revanche le type de tarif (en fonction du capital initial ou du capital restant dû) peut avoir un impact non négligeable, c'est pourquoi à défaut d'avoir les éléments nécessaires nous étudions les modes de distribution qui diffèrent justement sur le type de tarif (en CRD pour la banque et en CI pour le courtage/Internet)
- Les variables corrélées à l'occurrence d'un changement de vie (changement familial, professionnel, enrichissement...) devraient également jouer : âge (ces changements ne sont pas équi-répartis sur tous les âges), CSP (variable dont nous ne disposons malheureusement pas, mais peut-être corrélée au mode de distribution, variable dont nous disposons), capital assuré.

Notre première approche est d'établir des sous populations selon ces covariables puis d'en observer graphiquement les lois de survies correspondantes par la méthode d'estimation de Kaplan-Meier.

2.4.3.1 Âge à l'adhésion

L'âge à l'adhésion est un paramètre qu'il nous a semblé pertinent d'étudier. Indéniablement essentiel dans l'analyse de la mortalité, qu'en est-il pour la résiliation ? Pour pouvoir répondre à cette question regardons comment se comportent les courbes de survie selon les différents âges.

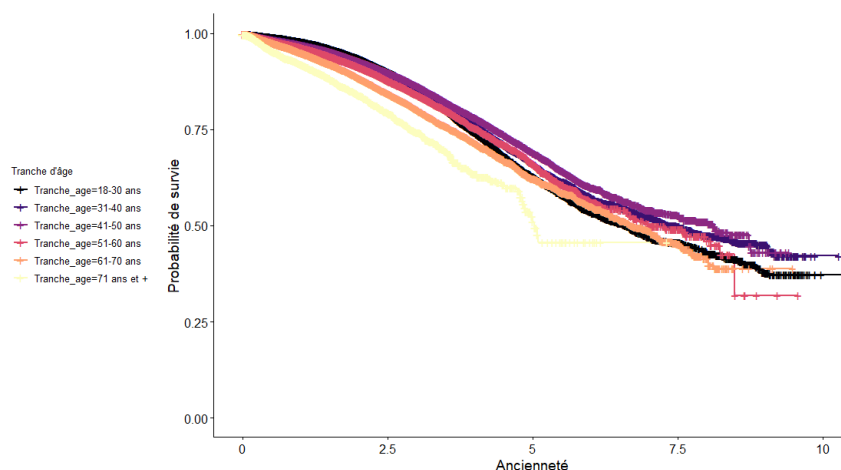
Figure 49 - Courbes de survie des contrats par âges



Visuellement il n'est pas évident de différencier les courbes de survie des différents âges mais on peut tout de même apercevoir que les courbes de survie des contrats des assurés de plus de 70 ans sont plus basses que pour les autres âges.

Pour en avoir le cœur net et essayer d'expliquer d'autres tendances nous subdivisons notre population en 6 tranches d'âges :

Figure 50 - Courbes de survie des contrats par tranches d'âges



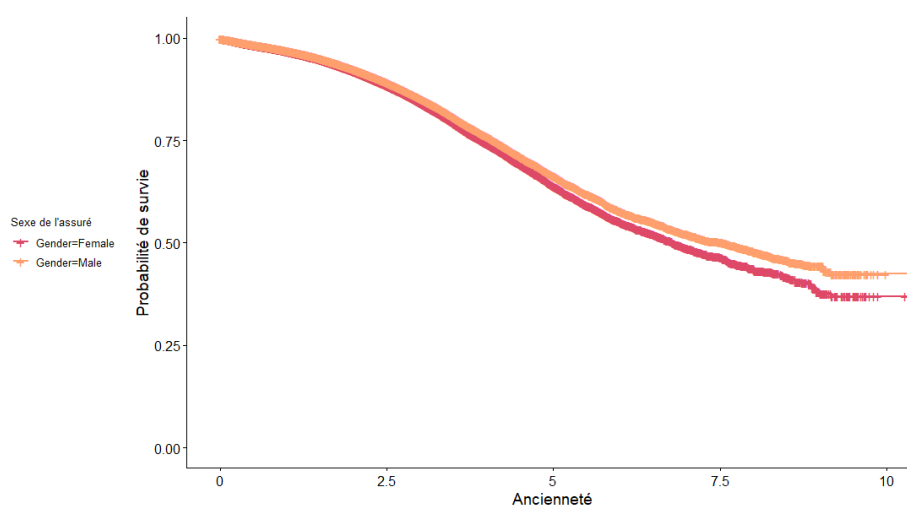
Effectivement nous pouvons affirmer que la courbe de survie des individus de plus de 70 ans est plus basse que les autres tranches d'âge mais cette différence n'est pas assez marquée pour en arriver à une conclusion satisfaisante. Les autres tranches apparaissent presque confondues et une décroissance avec l'âge n'est pas démontrée non plus. Par exemple la courbe de la tranche des 18-30 ans qui semble légèrement au-dessus les 2 premières années redescend très vite et fait partie des plus basses au bout de 10 ans.

Par conséquent, la dissemblance entre les tranches d'âges n'étant pas suffisamment marquée nous ne retiendrons pas l'âge à l'adhésion comme une variable explicative de la résiliation.

2.4.3.2 Sexe de l'assuré

Rappelons que le sexe de l'assuré est utilisé comme variable discriminante de la mortalité. Regardons ce qu'il en est pour la résiliation :

Figure 51 - Courbes de survie des contrats selon le sexe de l'assuré



Les courbes de survie en fonction du sexe sont très similaires, voir même indissociables les 5 premières années. L'écart que l'on observe à partir de la 5^{ème} année est léger et nous fait conclure que le sexe de

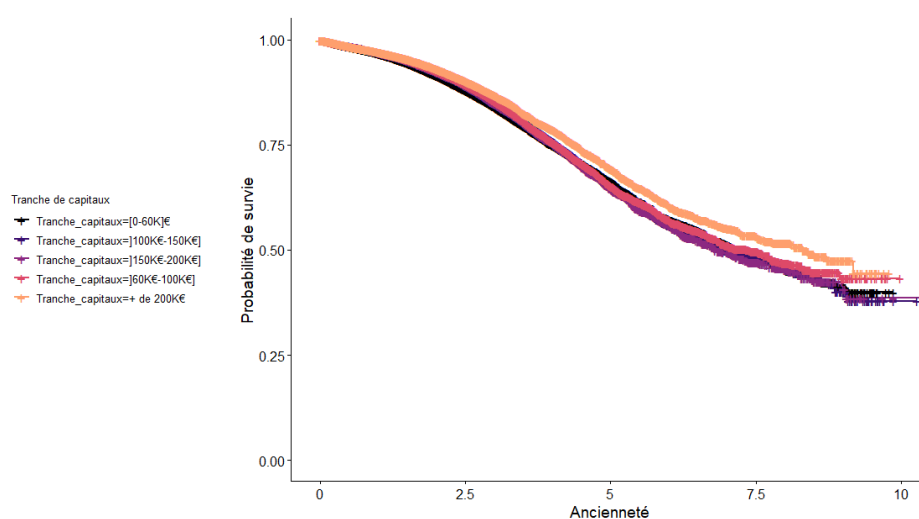
l'assuré n'est pas une variable suffisamment segmentante pour la résiliation. Nous décidons donc de ne pas garder le sexe comme variable discriminante pour notre modélisation.

2.4.3.3 Capital assuré

Les montants de capitaux qu'un assuré peut emprunter sont très vastes. La variable Capital Assuré est donc plutôt une variable continue dont les courbes de survie seraient difficilement observables en l'état. Ainsi comme pour les âges nous créons des tranches de capitaux assurés. Pour diviser notre portefeuille de sorte à avoir des tranches à peu près homogènes en nombre de contrat nous nous aidons des quantiles sur les montants assurés.

Regardons donc comment se comportent les courbes de survie des contrats en fonction de ces tranches :

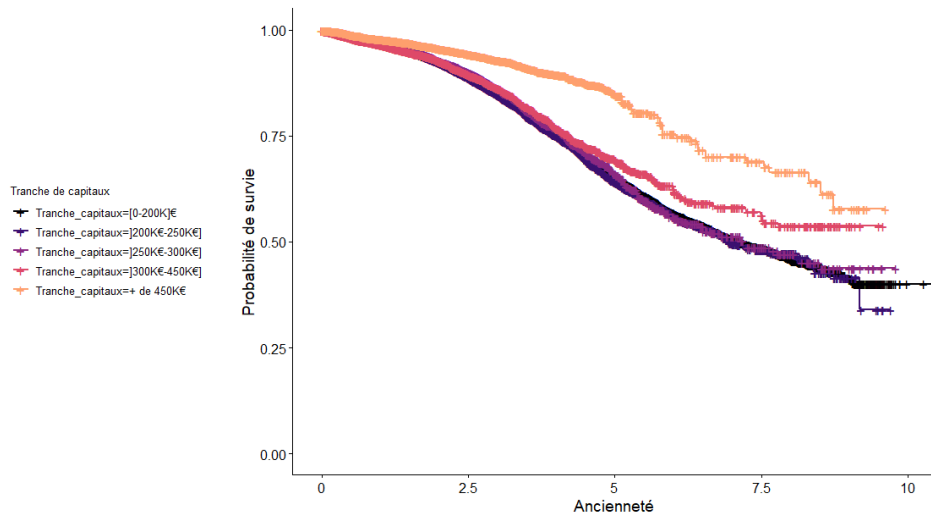
Figure 52 - Courbes de survie des contrats par tranches de capitaux (segmentation n°1)



Visuellement les capitaux supérieurs à 200 K€ semblent sortir un peu du lot tandis que les autres tranches sont confondues. Y aurait-il hétérogénéité de la population à ce seuil de capital ?

Nous décidons alors de resegmenter la tranche des capitaux supérieurs à 200 K€ afin d'explicitier une tendance que nous aurions manquée lors de notre première segmentation.

Figure 53 - Courbes de survie des contrats par tranches de capitaux (segmentation n°2)

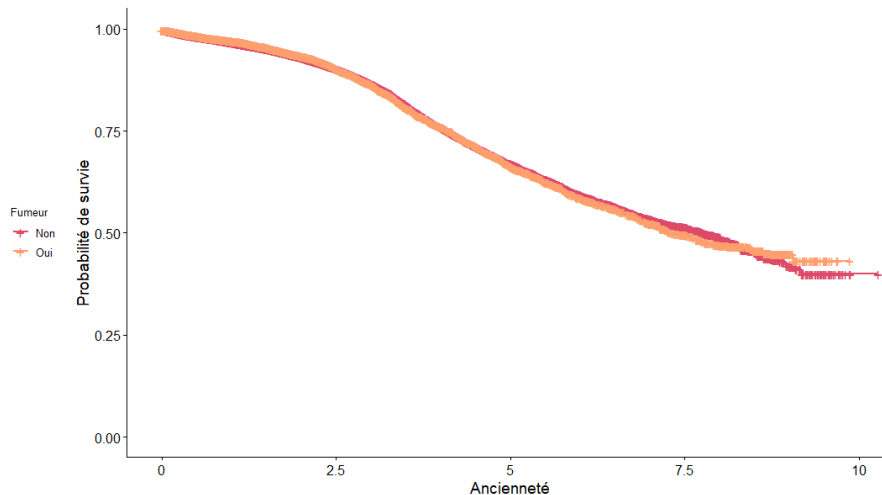


Effectivement comme pressenti on observe une disparité des courbes, mais plutôt à partir de 300 K€ et relativement prononcée au-delà de 450 K€. Pour l'instant nous retenons le capital assuré comme variable explicative de la résiliation avec les 3 tranches les plus marquées : [0-300K€],]300K€-450K€] et > 450K€. Nous tenterons de justifier par différents tests statistiques que cette segmentation est pertinente.

2.4.3.4 Profil fumeur

Nous savons que la significativité de la variable fumeur est avérée pour la mortalité. Par ailleurs rappelons que nous ne disposons pas d'une information fiable pour tous les produits sur ce paramètre, par conséquent nous restreignons l'observation uniquement aux individus dont on connaît le profil tabagique pour la résiliation.

Figure 54 - Courbes de survie des contrats selon le profil fumeur



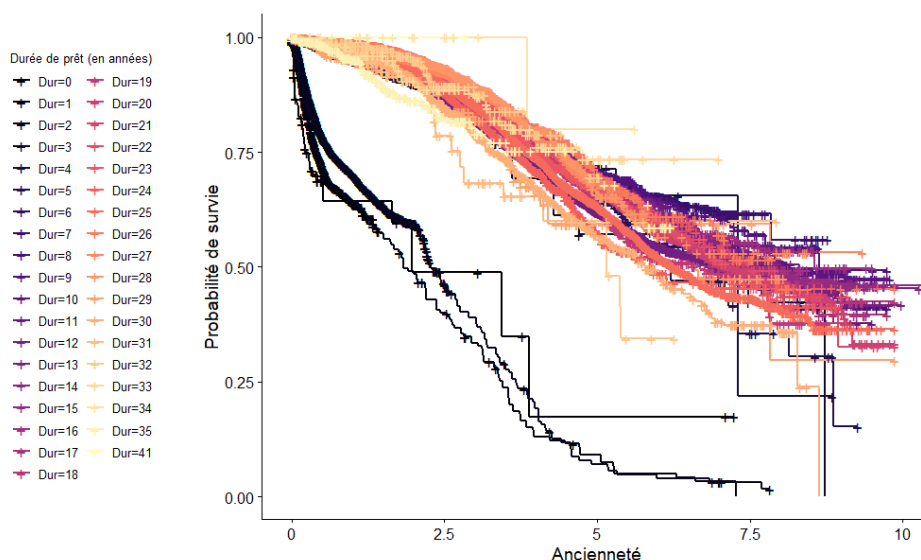
Nous concluons que le profil tabagique de l'assuré n'est pas discriminant dans la résiliation.

2.4.3.5 Durée de prêt initiale

La durée de prêt initiale est la durée sur laquelle l'assuré s'est engagé à rembourser son crédit.

Les durées de prêt que nous avons en portefeuille s'étendent jusqu'à 41 ans, nous traçons l'ensemble des courbes de survie correspondantes.

Figure 55 - Courbes de survie des contrats selon les durées initiales de prêt



Il peut arriver que l'assuré revoie ses mensualités à la baisse et prolonge donc son crédit de quelques années. Par conséquent le fait d'avoir des anciennetés supérieures à la durée de prêt initiale dans notre base de données n'est pas étonnant.

Une tendance ressort clairement du graphique en ce qui concerne les prêts de durées initiales 0 à 2 ans, dont la probabilité de survie est nettement inférieure au reste des durées. Rappelons quand même que les résiliations comptabilisées excluent les arrivées à échéance et les décès. Ce que nous observons est donc bien de la « pure » résiliation.

Une explication plausible serait que les prêts de moins de 2 ans sont plus faciles à rembourser par anticipation et surtout les capitaux empruntés sur ces durées sont généralement beaucoup plus faibles. Ainsi un clivage si prononcé apporte la justification sur le fait que la durée initiale de prêt est explicative de la résiliation. Mais à priori 2 tranches seraient largement suffisantes : durées inférieures ou égales à 2 ans et durées supérieures à 2 ans.

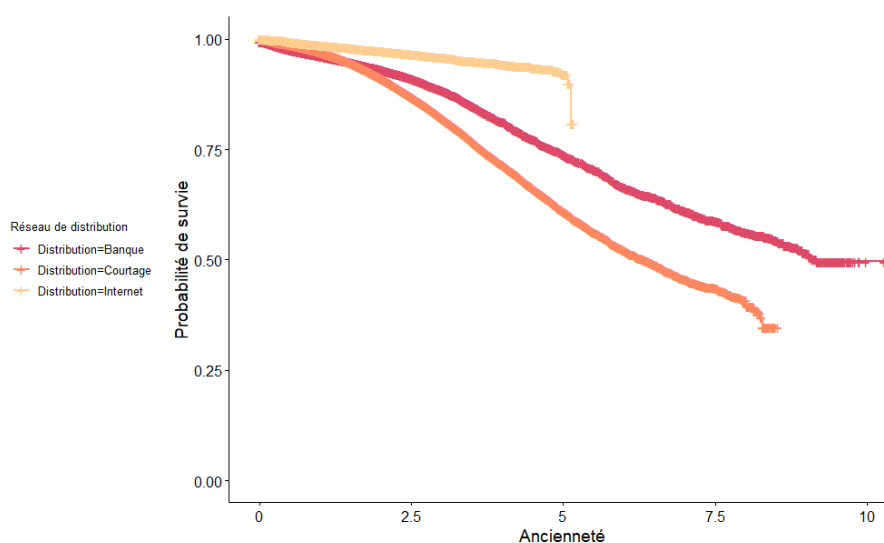
2.4.3.6 Mode de distribution

Le portefeuille observé dans le cadre de ce mémoire est historiquement le résultat de la commercialisation de 3 des plus gros produits vendus aux Partenariats Generali. Sa richesse et son intérêt réside dans le fait que ces produits sont vendus via 3 modes de distribution totalement différents avec chacun leurs spécificités : le réseau bancaire, le courtage et Internet.

Ainsi il est légitime de chercher à voir si ces spécificités se retrouvent également dans le comportement des assurés de chaque mode de distribution dans le cadre de la modélisation de notre loi de résiliation. De plus le produit bancaire étant essentiellement composé de tarifs en proportion du capital restant dû

et les produits courtage/Internet en proportion du capital initial, cette analyse aura également l'avantage de nous donner quelques indications sur l'influence de la typologie du tarif sur l'assuré.

Figure 56 - Courbes de survie des contrats par modes de distribution



A la vue de ces courbes de survie, il est indéniable que le réseau de distribution joue un rôle important dans l'explication de la résiliation. Les pratiques et les usages de chacun ont à priori un impact sur le comportement de la population assurée.

Par ailleurs rappelons que le produit Internet est le plus récent et le moins mature de notre portefeuille, ce qui explique que nous n'ayons plus de données au-delà de 5 ans.

Le fait que le réseau Internet ait la courbe de survie la plus haute est peut-être lié au profil de clients captés (plus jeunes, capitaux plus bas...).

S'agissant des réseaux bancaires, il s'agit probablement également d'une corrélation avec le profil de clients captés, avec en plus une probable aversion aux formalités administratives de renégociation de l'assurance, malgré l'entrée en vigueur de la loi Hamon en 2014 et de la loi Bourquin postérieurement.

Nous considérons que le réseau de distribution semble être une très bonne piste à étudier dans l'avancée de notre modélisation.

En conclusion, grâce à cette analyse graphique des courbes de survie des contrats, nous avons pu faire le tri dans les potentielles variables explicatives de la résiliation afin de pouvoir nous lancer à présent dans une modélisation statistique plus approfondie.

Les variables retenues pour l'instant sont les suivantes :

- Capital assuré,
- Durée de prêt,
- Réseau de distribution.

Dans les parties qui suivent nous mettons en parallèle deux méthodes statistiques qui vont nous permettre d'explicitier la significativité des variables retenues et de confirmer si oui ou non nos intuitions sont bien fondées.

Dans un premier temps à l'image de la loi de mortalité nous tentons une approche par le modèle de Cox. Puis dans un deuxième temps nous tentons une modélisation à l'aide des arbres de survie.

2.4.4 Approche par un modèle de Cox

Maintenant que nous avons observé graphiquement l'effet des différentes covariables sur les taux de résiliation, mesurons cet effet dans le modèle de hasard proportionnel de Cox.

2.4.4.1 Significativité des variables explicatives

Afin de mesurer la significativité des variables explicatives retenues précédemment nous utilisons la fonction *coxph* issue de package *survival* de R :

Figure 57 - Sortie R du test de significativité des covariables dans le modèle de Cox

```
Call:
coxph(formula = Surv(Ancienneté, Censure) ~ Tranche_capitaux +
      Dur + Distribution, data = data_complet)

n= 522860, number of events= 59948

              coef      exp(coef)  se(coef)      z      Pr(>|z|)
Tranche_capitaux]300K€-450K€] -0.01477  0.98534  0.02415  -0.611  0.5409
Tranche_capitaux+ de 450K€    -0.06500  0.93707  0.03040  -2.138  0.0325 *
Dur3 ans et +                 -2.49930  0.08214  0.01804 -138.531 <2e-16 ***
DistributionCourtage           0.40509  1.49944  0.01009  40.142  <2e-16 ***
DistributionInternet          -1.08991  0.33625  0.02537 -42.956  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

              exp(coef)  exp(-coef)  lower .95  upper .95
Tranche_capitaux]300K€-450K€]  0.98534  1.0149  0.93979  1.0331
Tranche_capitaux+ de 450K€    0.93707  1.0672  0.88286  0.9946
Dur3 ans et +                 0.08214  12.1740  0.07929  0.0851
DistributionCourtage           1.49944  0.6669  1.47008  1.5294
DistributionInternet           0.33625  2.9740  0.31993  0.3534

Concordance= 0.615 (se = 0.001 )
Likelihood ratio test= 17877 on 5 df,  p=<2e-16
Wald test              = 23680 on 5 df,  p=<2e-16
Score (logrank) test = 35969 on 5 df,  p=<2e-16
```

A la vue de ces résultats les covariables présélectionnées ont en effet un impact non négligeable sur les taux de résiliation (ou taux de hasard), mais la durée de prêt et le réseau de distribution apparaissent être les plus significatives (p-values nettement inférieures à 5% et statistiques de Wald élevées). De plus le niveau global de significativité du modèle est très bon avec des p-values pour les tests de Wald, du Score et du Ratio < à 5%. Cela montre que l'ajustement d'un modèle de Cox est très pertinent au seuil 5% et qu'il existe donc une influence significative d'au moins une covariable sur le taux de résiliation.

Cependant ces résultats n'ont de sens que si l'hypothèse de proportionnalité des hasards est vérifiée.

2.4.4.2 Validation de l'hypothèse de proportionnalité des hasards

Pour pouvoir utiliser le modèle de Cox il nous faut ainsi tester pour chaque covariable les hypothèses suivantes :

$$H_0 : \beta_j(t) = \beta_j \text{ contre } H_1 : \beta_j(t) \neq \beta_j .$$

2.4.4.2.1 Test du Khi-deux

Nous réalisons le test du Khi-deux à l'aide de la fonction *cox.zph* de R et voici ce que nous donnent les résultats :

Figure 58 - Sortie R du test d'hypothèse de proportionnalité des hasards

	chisq	df	p
Tranche_capitaux	59	2	1.5e-13
Dur	393	1	< 2e-16
Distribution	473	2	< 2e-16
GLOBAL	823	5	< 2e-16

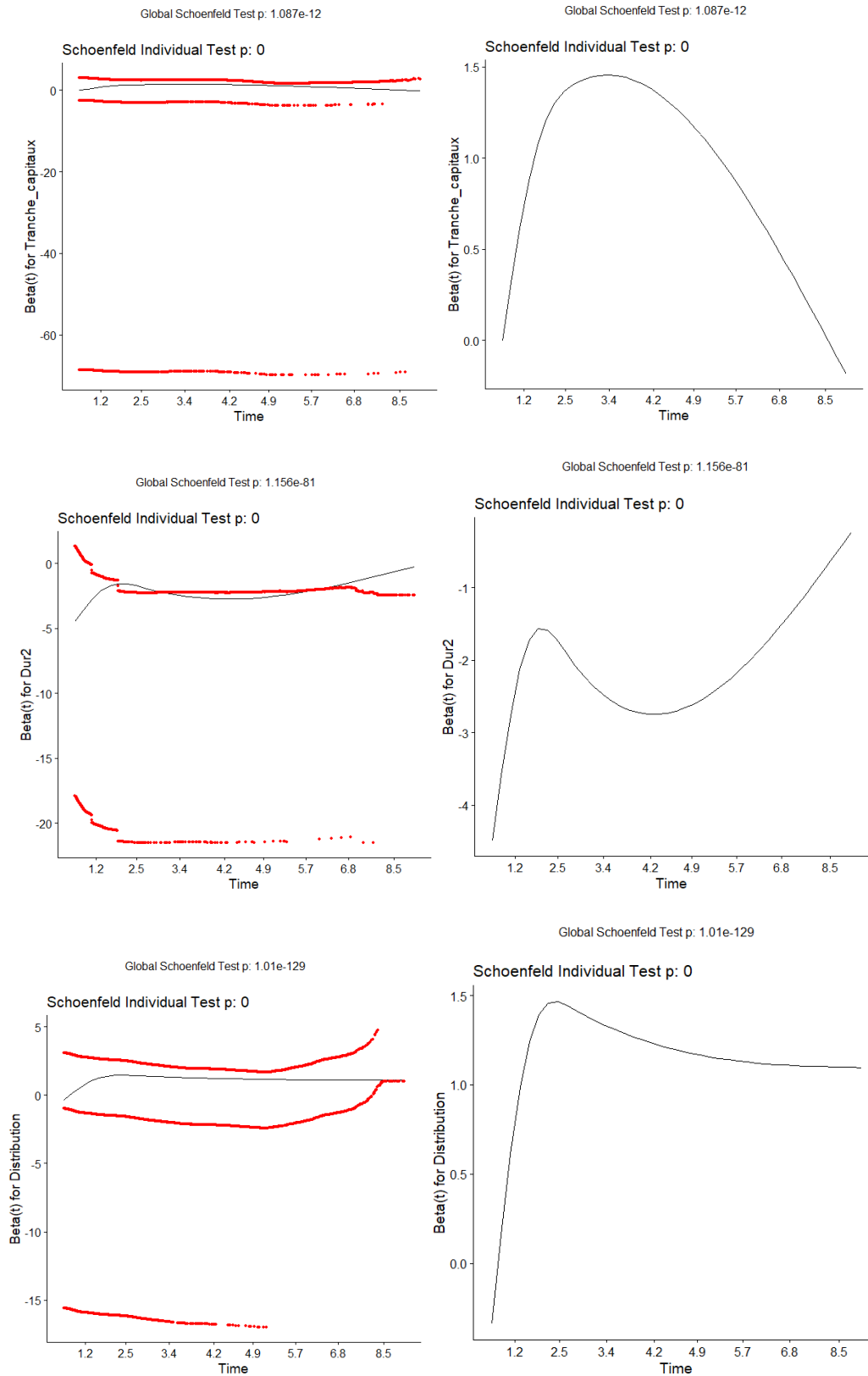
Quelle que soit la variable concernée, le test de validité de l'hypothèse de hasard proportionnel nous amène à la rejeter. Ceci impliquerait donc que l'effet des différentes covariables varie au cours du temps.

2.4.4.2.2 Analyse des résidus de Schoenfeld

Afin de confirmer le résultat précédent regardons également ce que nous donnent les résidus de Schoenfeld pour les 3 covariables (capitaux, durée, réseau de distribution). Une hypothèse de hasards proportionnels vérifiée signifie des résidus de Schoenfeld distribués de la même manière dans le temps et qui n'exhibent donc pas de tendance. En d'autres termes, les résidus doivent être répartis aléatoirement à l'intérieur d'une bande horizontale centrée en zéro.

La fonction *ggcoxzph* du package *survminer* de R trace les résidus de Schoenfeld ainsi qu'un lissage de ces résidus afin de pouvoir en analyser plus facilement la tendance. Mais pour des soucis de lisibilité à cause de l'échelle nous avons tracé un deuxième graphique pour chaque covariable présentant uniquement la courbe lissée des résidus.

Figure 59 - Résidus de Schoenfeld pour chaque covariable



Cette analyse graphique des résidus de Schoenfeld venant renforcer notre premier test d'hypothèse, comme nous nous y attendions il est clair que l'hypothèse de hasards proportionnels est invalidée pour les 3 covariables retenues.

Retravailler les covariables en fonction du temps serait même inutile car quelle que soit la fenêtre de temps observée les résidus ne sont jamais horizontaux centrés en zéro.

Même si l'hypothèse de hasards proportionnels n'est pas vérifiée, il reste avéré que les 3 covariables sélectionnées pourraient avoir un impact sur les taux de résiliations. Une modélisation par un modèle de Cox étant impossible, nous tentons une deuxième approche de type *Machine Learning* par arbres de survie afin de pouvoir confirmer pour une autre méthodologie l'hétérogénéité de la population selon le capital assuré, la durée de prêt et le réseau de distribution.

2.4.5 Approche par arbres de survie

Les méthodologies prédictives dites de *Machine Learning* ou d'apprentissage supervisé sont de plus en plus utilisées par les entreprises.

Parmi les utilisations les plus communes se trouvent :

- La détection des comportements frauduleux en matière de transactions en lignes,
- La prédiction des risques de non-solvabilité d'un client qui demande un crédit,
- La détermination des préférences d'un client afin de lui proposer de la publicité adaptée à ses besoins.

Ces préoccupations ne sont pas nouvelles et se propagent de plus en plus dans le monde et dans le domaine de l'assurance d'où un intérêt grandissant des actuaires pour la discipline.

La modélisation par arbres de survie comme décrite dans la partie 2.2.5, est une sous-discipline des méthodes d'apprentissage. Il s'agit d'une alternative non-paramétrique très populaire qui a su démontrer son efficacité et sa flexibilité. En effet contrairement aux arbres de survie, les modèles paramétriques ou semi-paramétriques, tel que celui de Cox par exemple, imposent une fonction de lien spécifique entre les variables explicatives et la variable à expliquer avec des hypothèses fortes pas toujours simple à valider. Ainsi nous faisons le choix d'utiliser l'arbre de survie dans le cadre de la construction de notre loi de résiliation pour palier à ces hypothèses assez contraignantes.

La finalité de cette partie est alors de confirmer les disparités de comportement des assurés en fonction des différentes caractéristiques de prêt présélectionnées dans les parties précédentes et d'en conserver un découpage optimal grâce à la segmentation donnée par l'arbre de survie.

2.4.5.1 Choix de la segmentation optimale

La segmentation par arbre repose sur la répartition des individus en groupes homogènes. La particularité d'un arbre dit de survie est que le critère d'homogénéité recherché est comme son nom l'indique, la survie.

En effet, dans le cas où nous observons la durée de vie des contrats en portefeuille, deux contrats appartiennent au même groupe seulement si leurs probabilités de survie sont « semblables ». Mais toute la difficulté de l'exercice réside justement dans le choix du niveau d'exigence souhaité pour cette similitude.

Le choix de l'arbre de survie optimal est le résultat d'un équilibre entre simplicité opérationnelle et précision. Cet équilibre peut être résumé par un paramètre nommé *complexity parameter* (CP). Ce terme que l'on retrouve dans le vocabulaire des méthodes par arbre traduit la complexité souhaitée du modèle. Plus celui-ci est petit plus l'arbre créé sera grand et donc complexe. Dans le modèle de régression, le

CP se traduit littéralement par : si une division a un nœud n'améliore pas l'erreur du modèle d'au moins CP alors il n'y a plus de division sur ce nœud car considérée non pertinente à ce niveau de complexité donné.

La problématique qui se pose avec des arbres trop complexes c'est que l'on doit faire face à des problèmes de sur-apprentissage. En d'autres termes, quand l'arbre s'adapte trop aux spécificités de la base de données il devient alors difficile à généraliser à un autre jeu de données. De plus pour des raisons opérationnelles évidentes, nous privilégierons un arbre simple à interpréter et facile d'utilisation.

Dans cette partie nous essayons donc à l'aide du package *DStree* (*Discret Time Survival Tree*) de R, développé spécialement pour la construction d'arbres de survie en temps discret, de choisir un arbre qui capte bien les hétérogénéités de nos données tout en gardant un niveau de complexité raisonnable.

Nous choisissons d'utiliser particulièrement cette méthode développée par Larocque et Bou-Hamad car elle répond parfaitement à nos besoins d'une part pour le type de données (donnée de survie) et d'autre part pour le format en temps discret du résultat produit. En effet la finalité est d'obtenir une loi de résiliation des contrats en fonction de leur ancienneté pour pouvoir l'utiliser dans nos calculs de rentabilité en des pas de temps finis $t = 1, 2, \dots, 20$.

Rappelons que les variables et modalités qui ont été retenues jusqu'à présent sont :

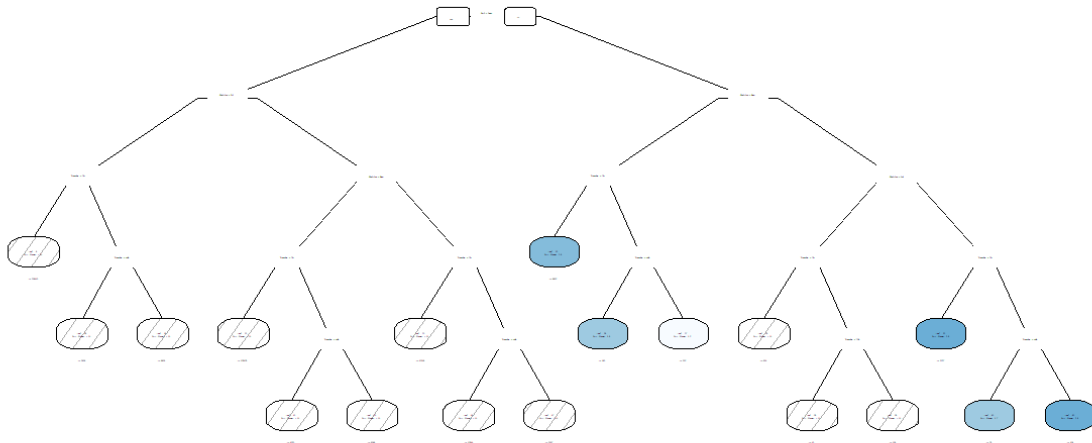
- Le capital assuré :
 - o Inférieur ou égal à 300 K€,
 - o Supérieur à 300 K€ et inférieur ou égal à 450 K€,
 - o Supérieur à 450 K€,
- La durée de prêt initiale :
 - o Inférieure ou égale à 2 ans,
 - o Supérieure à 2 ans,
- Le réseau de distribution :
 - o Banque,
 - o Courtage,
 - o Internet.

2.4.5.1.1 Construction de l'arbre maximal

La fonction *DStree* du package du même nom a l'avantage de permettre à son utilisateur de choisir le paramètre CP. D'autres paramètres sont modulables comme par exemple, le nombre minimum d'observations par nœuds avant ou après division, mais nous ne les utiliserons pas dans le cadre de notre étude.

Afin de pouvoir imaginer à quoi pourrait ressembler un arbre de complexité maximale nous avons construit un arbre avec comme paramètre $cp = 0$.

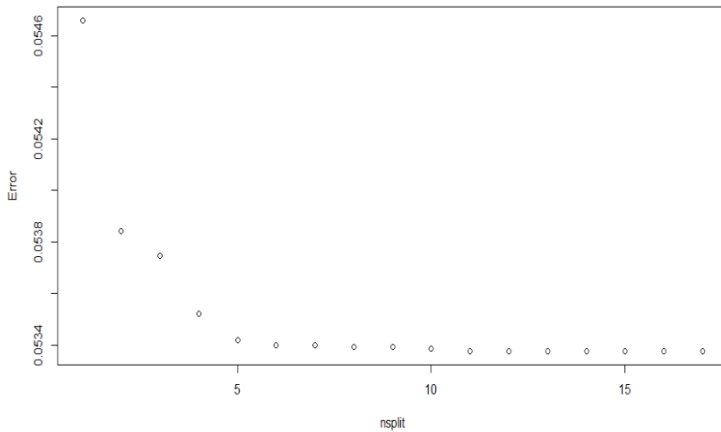
Figure 60 - Arbre de survie maximal des contrats



L'arbre maximal obtenu possède 6 niveaux de profondeur, 17 divisions et comme on l'imaginait est difficilement lisible.

D'un point de vue précision de l'estimation nous pouvons tracer l'erreur de Brier à chaque sous-arbre ou division supplémentaire :

Figure 61 - Erreur et complexité en fonction du nombre de divisions de l'arbre

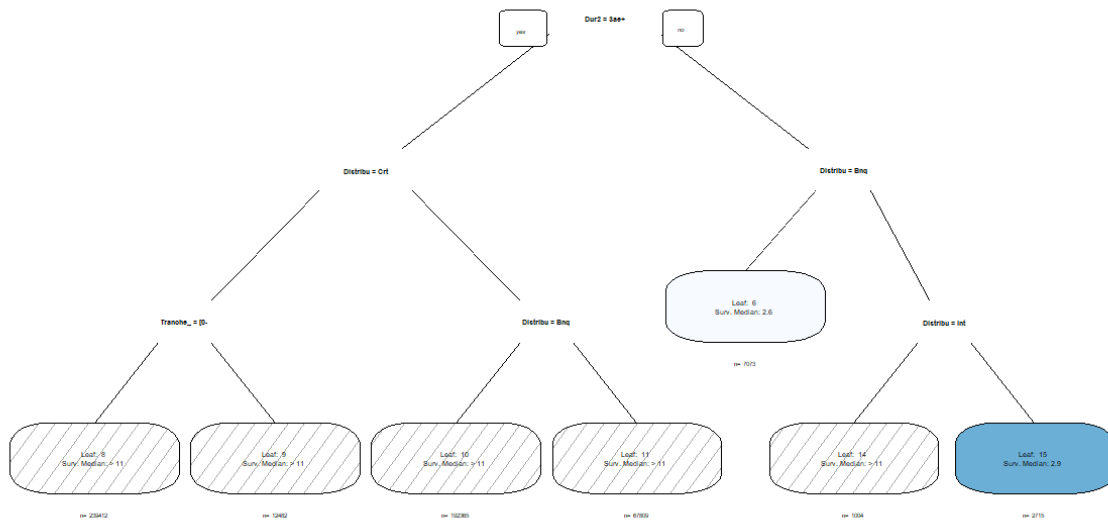


	CP	nsplit
1	2.314932e-02	0
2	2.021237e-02	1
3	3.547626e-03	2
4	1.320820e-03	3
5	6.763368e-04	4
6	1.308058e-04	5
7	8.715405e-05	6
8	4.744451e-05	7
9	4.007960e-05	8
10	3.531523e-05	9
11	3.373822e-05	10
12	2.784277e-05	11
13	1.168866e-05	12
14	1.138522e-05	13
15	1.100807e-05	14
16	7.530749e-06	15
17	2.355229e-06	16
18	0.000000e+00	17

2.4.5.1.2 Elagage de l'arbre

Nous constatons qu'au-delà de 5 divisions, la précision d'estimation apportée par une division supplémentaire est négligeable et que le niveau de complexité du sous-arbre de 5 divisions est de 1.31×10^{-4} . Nous reconstruisons donc un nouvel arbre en l'élaguant avec ce niveau de complexité :

Figure 62 - Arbre élagué (1ère version)



Certes au bout de 5 divisions l'erreur est presque à son minimum mais un arbre à 4 niveaux et d'une telle complexité reste toujours difficile à lire sachant que nous n'avons que 3 covariables avec peu de modalités à étudier. Par conséquent nous faisons le choix de fixer cp à $3,55 \times 10^{-3}$ de sorte à avoir au maximum 2 divisions et un niveau d'erreur raisonnable comme le montre la courbe d'erreur précédente :

Figure 63 - Arbre élagué (2ème version)

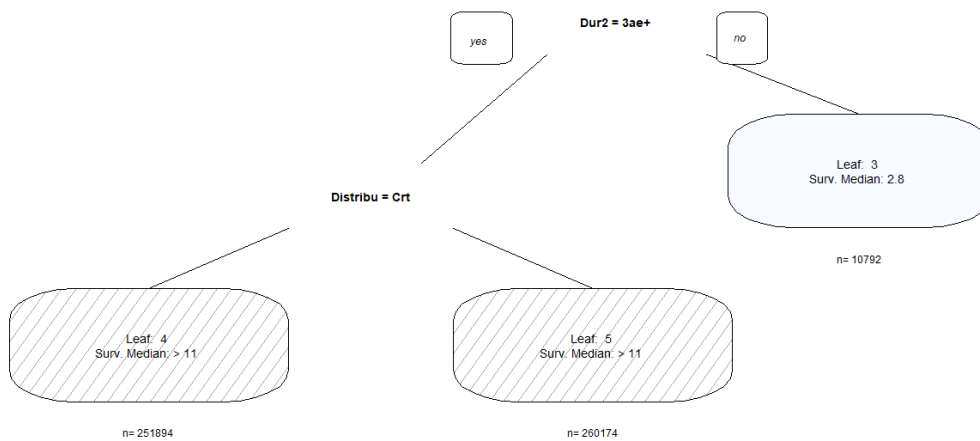


Figure 64 - Sortie R de la complexité à chaque nœud de l'arbre élagué

```
Call:
DStree(formula = Anc ~ Dur2 + Tranche_capitaux + Distribution,
        status = "Censure", data = data_complet, control = DStree.control(cp = 0.003547626))
n= 522860
```

	CP	nsplit	
1	0.023149322	0	
2	0.020212373	1	
3	0.003547626	2	

Ainsi nous obtenons une segmentation optimisée selon le niveau de complexité choisi qui est bien un entre deux entre facilité d'utilisation et précision. Comme nous le pressentions depuis le début les deux covariables retenues par l'arbre optimisé sont en premier lieu la durée de prêt (inférieure ou supérieure à 3 ans) puis le réseau de distribution (Courtage, Banque ou Internet). Cependant la variable réseau de distribution a été résumée par l'arbre en Courtage d'un côté et Banque/Internet de l'autre, tandis que celle des capitaux assurés n'a pas été retenue. Rappelons que dans le modèle de Cox le capital assuré était la moins significative des 3 variables.

2.4.5.1.3 Fonctions de survie et taux bruts de résiliation

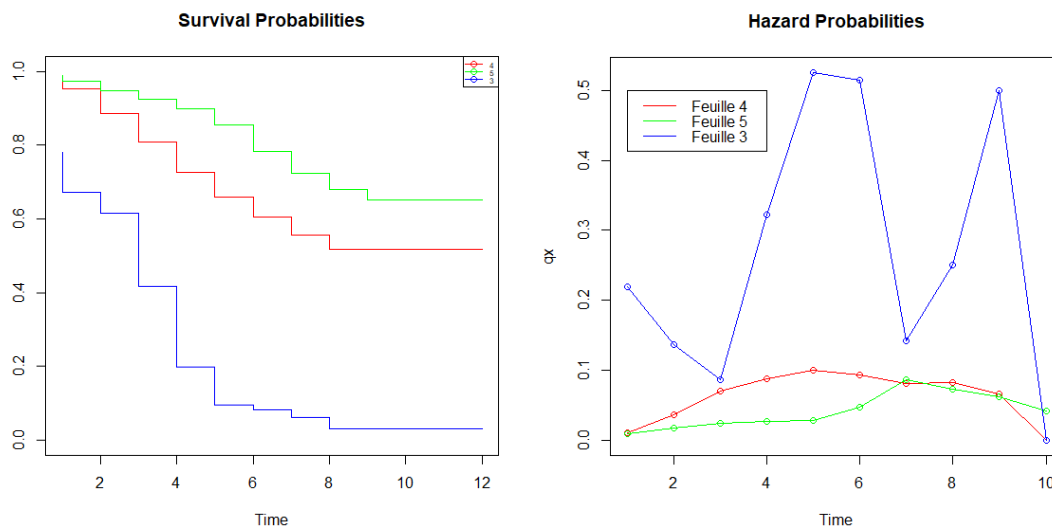
Chaque feuille est assimilée à une courbe de survie qui lui est propre, ici numérotées 3, 4 et 5 en lien avec l'arbre de survie précédemment tracé :

- **Feuille n°3** : Durée de prêt inférieure à 3 ans
- **Feuille n°4** : Durée de prêt supérieure ou égale à 3 ans et distribution via le Courtage
- **Feuille n°5** : Durée de prêt supérieure ou égale à 3 ans et distribution via la Banque ou Internet

Notons que les feuilles n°1 et n°2 correspondent aux nœuds de l'arbre, mais comme ce ne sont pas des nœuds terminaux nous ne pouvons pas les observer comme des feuilles sur l'arbre précédent.

Graphiquement ci-dessous les courbes de survie et taux bruts de résiliation associées à chaque feuille (nœuds terminaux) :

Figure 65 - Courbes de survie et taux bruts de résiliation des contrats par feuilles



Les courbes de survie apparaissent assez bien dissociées, d'autant plus pour les prêts de moins de 3 ans (nous comprenons bien pourquoi il s'agit de la première division choisie par l'algorithme).

De plus le test d'hétérogénéité du Khi-deux (log-rank) ci-dessous confirme bien que cette segmentation est légitime :

Figure 66 - Sortie R du test de Log-Rank

```
Call:
survdiff(formula = surv(Anc, Censure) ~ FEUILLE, data = data_complet)
```

	N	Observed	Expected	(O-E)^2/E	(O-E)^2/V
FEUILLE=3	251894	43480	35180	1958	5151
FEUILLE=4	260174	12938	24300	5313	9672
FEUILLE=5	10792	3530	468	20038	20980

Chisq= 28414 on 2 degrees of freedom, p= <2e-16

De manière univoque le test nous amène à refuser l'hypothèse H_0 : les trois distributions de survie sont identiques (p -value quasiment nulle).

Concernant les taux bruts, ceux de la feuille n°3 (durées de prêt < 3 ans) sont très irréguliers, il est donc très difficile d'en déduire une tendance. Rappelons que la qualité de l'estimation repose aussi sur le volume de données disponibles. Or on sait que la feuille 3 est celle qui contient le moins d'individus (cf. arbre Figure n°68). L'exposition à chaque pas de temps est donc un paramètre essentiel qui devra être pris en considération pour l'ajustement de la loi de résiliation finale, sachant que plus l'ancienneté est élevée moins nous avons d'individus.

En conclusion de cette analyse des covariables, nous retenons cette segmentation comme étant la segmentation optimale de notre portefeuille dans le cadre de la modélisation de la résiliation. Cependant au regard de la forme des q_x il est évident qu'une étape de lissage est nécessaire.

2.4.5.2 Lissage des taux bruts de résiliation par splines

Pour procéder au lissage de nos taux bruts nous choisissons d'utiliser un lissage par splines en particulier les *Smoothing Cubic Splines*. Cette méthode a plusieurs avantages dans le cadre de notre étude :

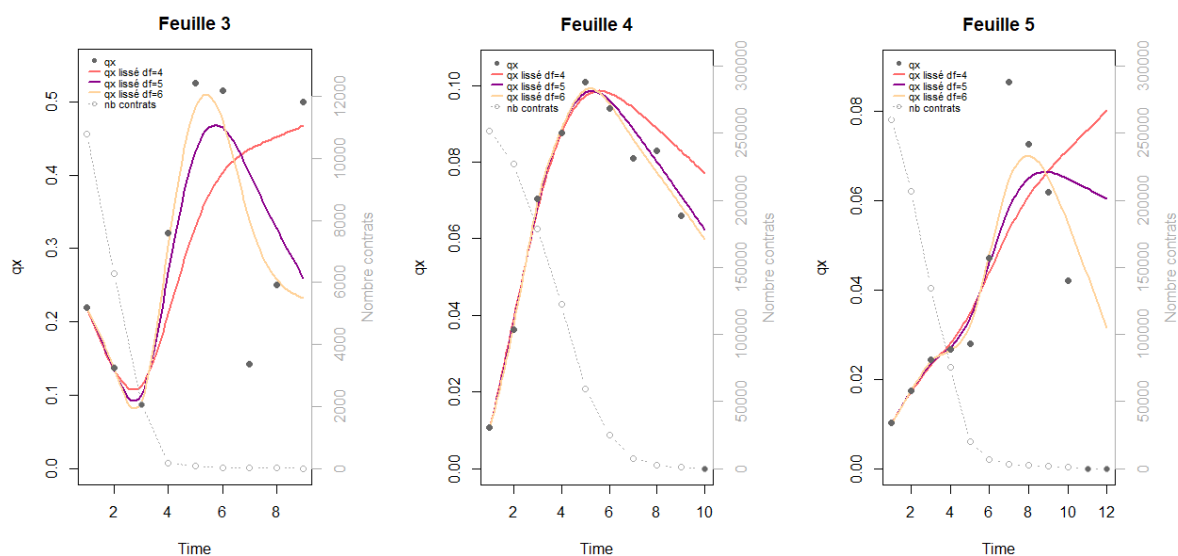
- Elle prend en considération un paramètre de pondération en chaque point,
- La fonction obtenue est aisément extrapolable,
- Un paramètre de régularité peut être défini au besoin.

Cette méthode est implémentée dans la fonction *smooth.spline* de R et c'est donc celle que nous utilisons. Deux composantes sont prises en considération par la fonction : la fidélité de l'estimation et sa régularité. Comme expliqué précédemment le niveau de régularité est paramétrable via le degré de liberté (df) compris entre 0 et le nombre de point. Dans le cadre de notre étude, c'est sur ce paramètre que nous effectuons des sensibilités afin de choisir la spline la plus adaptée à chaque feuille.

2.4.5.2.1 Comparaison graphique des lissages selon le degré de liberté

D'un point de vue pratique nous traçons donc les q_x bruts et les q_x lissés puis nous apposons au même graphique le nombre de contrats sous risque à chaque pas de temps. En année 1 nous avons donc l'ensemble des contrats du début de l'observation quel que soit l'année de souscription puis en année 2 les contrats toujours présents passée une année et ainsi de suite. Le portefeuille que nous étudions étant en forte croissance années après années il est normal d'avoir beaucoup plus de contrats à observer les premières années. Ce nombre de contrats à chaque année nous sert ainsi de pondération pour le lissage. En d'autres mots les années pour lesquelles le volume de contrats est très faible ne seront que très peu prises en considération.

Figure 67 - Comparaison des lissages des taux bruts de résiliation à chaque feuille



Ces premières approximations sont celles qui vont nous servir de base pour l'extrapolation. En effet pour notre étude nous choisissons d'observer une estimation des résultats sur un horizon de 20 ans, durée jugée suffisante sur une cohorte d'assurés souscrivant une nouvelle assurance. Le choix du paramètre df est donc important dans l'ajustement car c'est lui qui détermine la tendance de la fonction approximée sur la période où des q_x bruts ont pu être observés mais aussi au-delà une fois la courbe extrapolée.

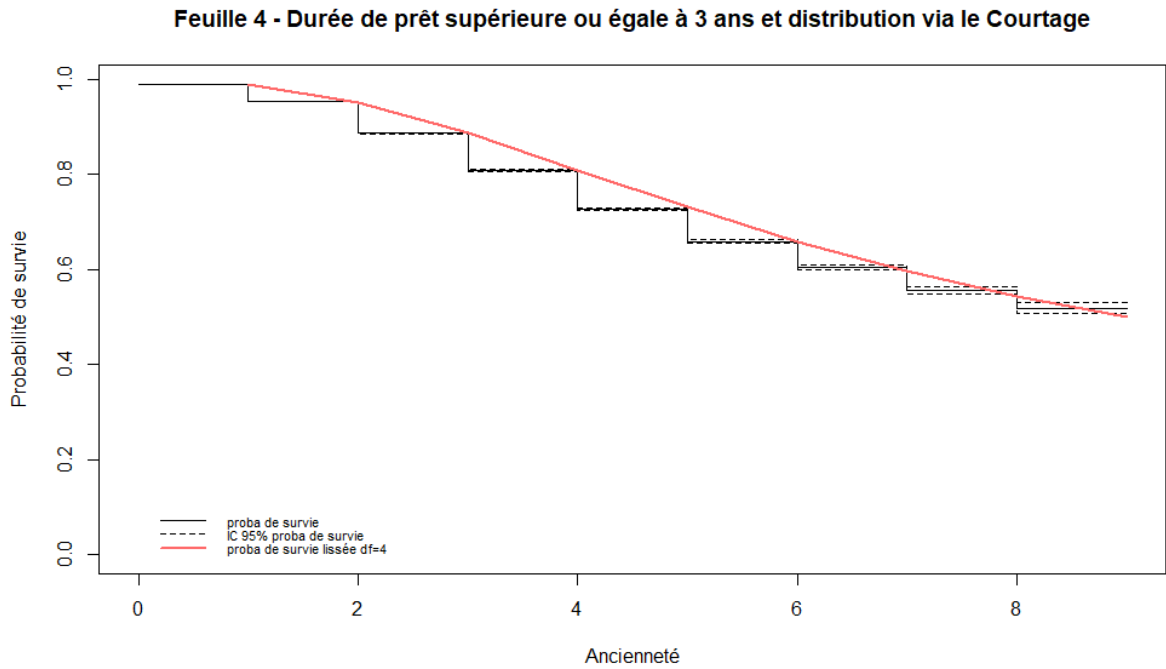
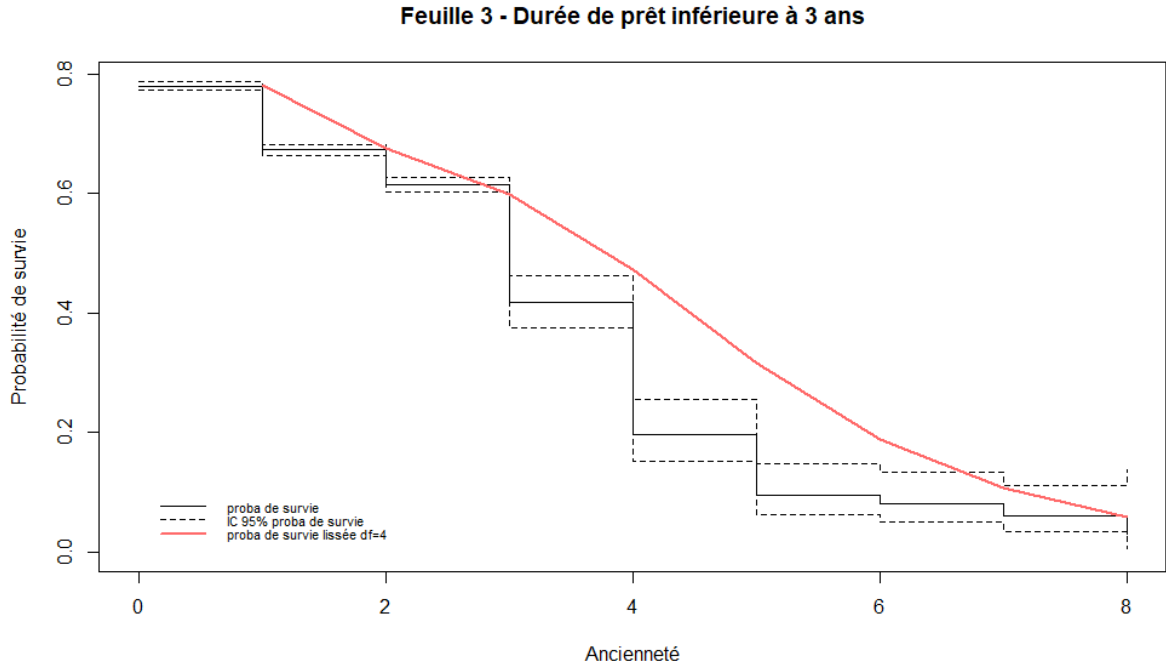
Concernant la Feuille n°3 (durées initiales de prêt inférieures à 3 ans), on constate que plus on se rapproche de la fin théorique du prêt plus les taux sont bas. Cependant au-delà de 3 ans les taux sont bien plus élevés et volatiles. Ce phénomène s'explique par le fait qu'à partir de ce temps d'observation nous n'ayons que très peu de données. De plus passé la durée théorique il n'est pas invraisemblable de penser que la probabilité de résiliation puisse augmenter. En effet les contrats de durée très courte sont pour beaucoup des prêts in-fine (remboursement unique à terme) qui sont généralement remboursés rapidement en raison des faibles montants empruntés. Le degré de liberté à 4 semble à première vue être une valeur adéquate pour ce cas particulier.

Concernant les Feuilles n°4 (durées de prêt supérieures à 3 ans / Courtage) et n°5 (durées de prêt supérieures à 3 ans / hors Courtage), l'idée d'avoir une courbe toujours croissante n'est pas envisageable car il est très peu probable qu'au bout de 20 ans le taux de résiliation soit à son maximum car de notre expérience de la Feuille n°3 plus on se rapproche de la fin théorique plus les taux de résiliations diminuent. Or avec une durée initiale moyenne de 18 ans nous sommes à l'aise avec le fait de supposer que plus on se rapproche de l'échéance finale, plus les taux ont tendance à décroître. Pour la Feuille n°4 un degré de liberté à 4 est un bon compromis car il nous permet d'avoir cette tendance baissière progressive et non rapide comme pour les degrés 5 et 6. Nous n'avons pas tracé le degré 3 mais ce dernier nous donnait une courbe toujours croissante ce qui va à l'encontre de nos hypothèses. Pour ce qui est de la Feuille n°5 le degré de liberté à 5 semble être le plus adapté pour les mêmes raisons que celles qui nous ont poussés à choisir le degré 4 pour la Feuille n°4.

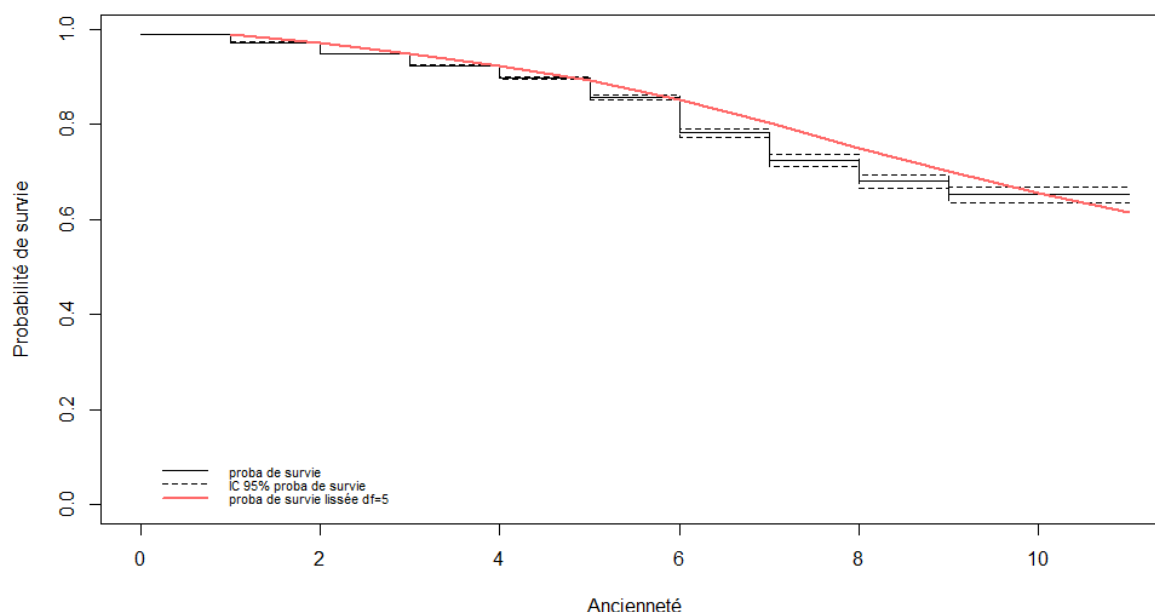
2.4.5.2.2 Tests et validations statistiques

A présent regardons dans un premier temps l'allure des nouvelles courbes de survies ainsi déduites et regardons ensuite à l'aide des tests statistique du Khi-deux et du changement des signes si nos choix de degrés de liberté nous permettent d'avoir une approximation satisfaisante.

Figure 68 - Fonctions de survie après lissage par feuilles



Feuille 5 - Durée de prêt supérieure ou égale à 3 ans et distribution via la Banque ou Internet



Dans la suite de nos travaux nous utiliserons la loi de résiliation approximée aux pas de temps discrets de 1 à 10 et c'est pour cela que l'évaluation des tests est réalisée en ces points-là. Ainsi graphiquement ce qui importe est la proximité de la courbe issue des q_x lissés avec la courbe de survie brute en ces pas de temps précis. Le fait que cette dernière soit en escalier étant donné sa précédente discrétisation n'est donc pas un problème. De plus rappelons que lors de l'estimation des q_x nous avons pondéré chaque point par l'effectif des assurées ce qui explique pourquoi sur les dernières années la courbe lissée n'est pas toujours dans l'intervalle de confiance donné par l'estimateur de Kaplan-Meier. Pour autant l'écart entre les deux courbes reste faible aux points d'évaluation quelle que soit la feuille.

Après cette visualisation graphique voici ci-dessous comment se présentent les tests statistiques :

a) Test du Khi-deux

Figure 69 - Résultats du test d'adéquation du Khi-deux selon le degré de liberté de la fonction spline

Test Khi-deux <i>p-value</i>	df = 4	df = 5	df = 6
Feuille n°3	0.8326	0.9981	1.0000
Feuille n°4	0.9859	0.9999	1.0000
Feuille n°5	0.9611	0.9946	0.9999

b) Test du changement des signes

Figure 70 - Résultats du test des signes selon le degré de liberté de la fonction spline

Test des signes	df = 4	df = 5	df = 6
Feuille n°3	p-value = 0.5078 IC 95% = [-0.188 , 0.112] médiane = 0.0020	p-value = 1.0000 IC 95% = [-0.073 , 0.091] médiane = 0.0021	p-value = 1.0000 IC 95% = [-0.008 , 0.039] médiane = 0.0008
Feuille n°4	p-value = 0.7539 IC 95% = [-0.016 , 0.002] médiane = -0.0033	p-value = 0.7539 IC 95% = [-0.007 , 0.003] médiane = -0.0013	p-value = 0.7539 IC 95% = [-0.004 , 0.002] médiane = -0.0009
Feuille n°5	p-value = 0.7744 IC 95% = [-0.027 , 0.003] médiane = -0.0008	p-value = 0.3877 IC 95% = [-0.021 , 0.001] médiane = -0.0003	p-value = 0.7744 IC 95% = [-0.012 , 0.0003] médiane = -0.0005

Au niveau global, la pertinence de notre ajustement est appréhendée sur deux niveaux :

- **L'écart entre l'ajustement et les observations** : il s'agit de jauger la distance entre ces deux quantités (test du Khi-deux),
- **La qualité du lissage** : il s'agit de déterminer si les données ont été sur-lissées ou sous-lissées (test des signes).

Le test du Khi-deux nous donne des résultats déjà très satisfaisants dès le degré 4 et excellents à partir du degré 5 (p -value proche de 1). La p -value la plus faible que nous ayons est celle de la Feuille n°3 au degré de liberté égal à 4 avec une valeur de 0.8326. Mais de toute manière lors de notre première approche graphique nous avons établi que pour la Feuille n°3 choisir $df = 4$ était préférable pour des raisons de cohérence avec la réalité du portefeuille. Pour les Feuilles n°4 et n°5 nous avons choisi respectivement $df = 4$ et $df = 5$ et le test du Khi-deux conforte notre idée.

Le test des changements de signes nous laisse par contre mitigés notamment pour les Feuilles n°3 et n°5. Rappelons que ce test a pour but essentiellement de nous dire si notre ajustement a tendance à surestimer ou sous-estimer les taux réels. Cependant précédemment nous avons précisé qu'il était plus judicieux de légèrement surestimer les q_x pour éviter lors de l'extrapolation d'avoir une chute brutale de la courbe avant la 20^{ème} année. De plus pour notre étude de rentabilité nous préférons privilégier la prudence. Par conséquent nous ne tenons pas compte du résultat du test pour la Feuille n°5 (p -value = 0.3877). Pour la Feuille n°3 toujours pour des raisons de cohérence nous restons sur le degré de liberté égal à 4 malgré le résultat du test (p -value = 0.5078). Pour la Feuille n°4 le test des signes ne permet pas d'arbitrer car les p -values sont identiques quelle que soit la valeur de df mais pour autant sa valeur est plutôt satisfaisante.

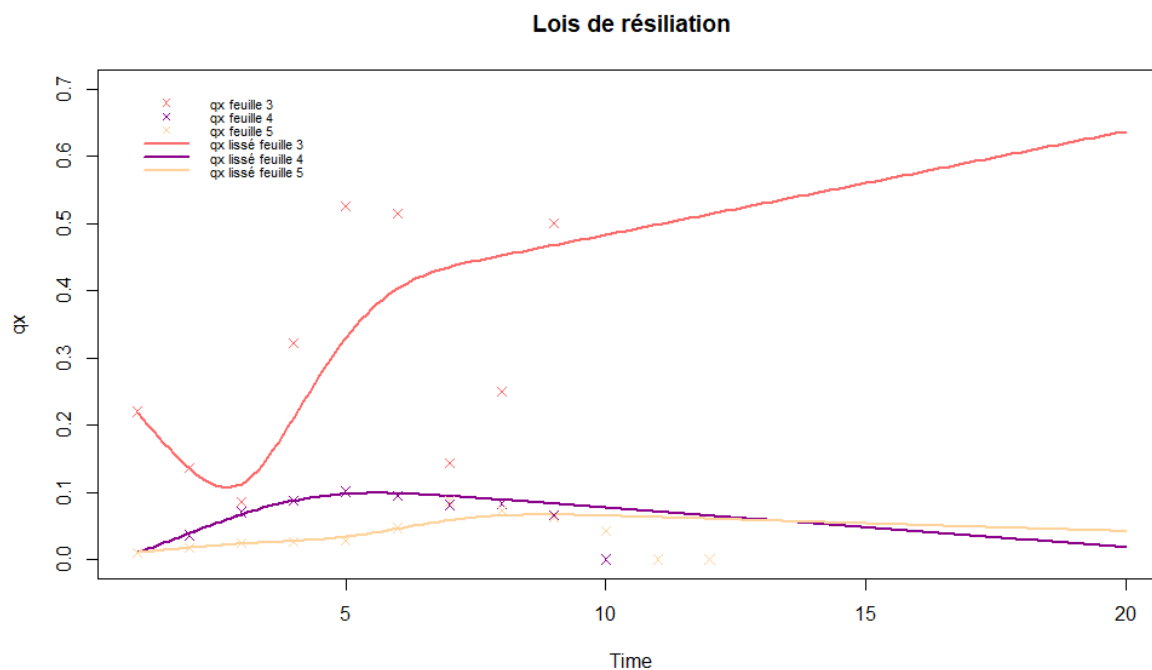
En conclusion après notre analyse des résultats des tests statistiques voici ci-dessous les degrés de liberté que nous avons choisis pour chacune de nos trois catégories :

Figure 71 - Récapitulatif des degrés de liberté de la fonction spline à chaque feuille

Feuille	df
Feuille n°3	4
Feuille n°4	4
Feuille n°5	5

Lors de l'extrapolation des fonctions splines nous nous sommes assurés que les qx lissés restent positifs sur la fin de la projection en capant les valeurs à 0. Voici alors la loi de résiliation que nous obtenons pour chaque feuille :

Figure 72 - Taux de résiliations lissés et extrapolés



Ainsi nous disposons à présent d'un groupe de lois de résiliation qui selon nous reflète au mieux notre portefeuille d'assurés et c'est ce dernier que nous utilisons par la suite dans le cadre de la modélisation d'un nouveau portefeuille d'assurance Emprunteur.

3 MODÉLISATION DE L'ACTIVITÉ D'ASSURANCE EMPRUNTEUR DANS LE NOUVEAU PARADIGME

Comme précisé dans l'introduction du mémoire, une évolution souhaitée par les acteurs en assurance du marché est maintenant la suppression des formalités médicales et la garantie du tarif dans le cadre d'un changement de banque prêteuse. Opérationnellement, cela permettrait donc la transférabilité du contrat d'assurance emprunteur.

Les impacts pour la rentabilité de l'assureur sont contradictoires :

- dans le sens positif, des clients conservés plus longtemps en portefeuille, ce qui a pour conséquence de réduire le coût d'acquisition ;
- dans le sens négatif une raréfaction des volumes d'affaires nouvelles souscrites et une possible hausse de la sinistralité moyenne (sous l'effet de l'absence de nouvelle sélection médicale au moment du transfert).

L'assureur qui souhaite maintenir sa rentabilité et sa compétitivité tarifaire doit étudier l'opportunité d'offrir cette faculté à la lumière d'une analyse de ces deux effets sur le profit à court-terme et à long-terme. L'objectif de cette partie est pour cela de modéliser la rentabilité de l'activité d'assurance Emprunteur dans le cadre du nouveau paradigme, dans lequel la transférabilité serait permise sur nos contrats.

Nous décomposons notre raisonnement en deux parties successives :

- Tout d'abord, nous résolvons un problème d'optimisation consistant à maximiser les profits futurs générés par une cohorte de production sous contrainte de respect d'un niveau de rentabilité et de règles de tarification acceptable et déployables dans des réseaux de distribution. L'output de cette partie est une approche tarifaire permettant d'optimiser le tarif d'une affaire nouvelle dans le nouveau paradigme.
- Ensuite, nous étudions la décision stratégique d'entrer ou non dans le nouveau paradigme en modélisant la rentabilité de l'activité d'assurance Emprunteur d'un acteur ayant adopté la transférabilité. L'output de cette partie est une projection du compte de résultat de l'activité d'assurance emprunteur dans le nouveau paradigme et sa sensibilité à différentes hypothèses sur la réponse du marché. Des critères pour la prise de décision sont proposés en conséquence.

Pour simplifier l'analyse, nous étudions dans ce mémoire la transférabilité du contrat d'assurance sur le capital restant dû : nous faisons l'hypothèse que le capital à assurer n'augmente pas ou que cette augmentation est couverte par un autre contrat, de telle sorte que le nouveau paradigme consiste simplement à prolonger le contrat existant dans ses termes initiaux.

3.1 Résolution du problème d'optimisation : maximisation de la rentabilité générée par une cohorte dans le nouveau paradigme

L'output recherché pour ce chapitre est une approche tarifaire permettant d'optimiser le tarif d'une affaire nouvelle dans le nouveau paradigme. Pour permettre une prise de décision (cf. partie 3.2), nous recherchons également une optimisation de l'approche tarifaire dans le paradigme actuel en tenant compte des conclusions sur la mortalité et la tenue des contrats de la partie précédente (cf. partie 2). Cette partie est décomposée de la manière suivante :

- Dans une première section, nous définissons les indicateurs de rentabilité à maximiser et choisissons les degrés de liberté du problème.

- Dans une seconde section, nous formulons des hypothèses sur le nouveau paradigme : impacts sur la loi de chute des contrats modélisée, changement de profil des clients captés...
- La troisième section présente le résultat des modélisations et une approche tarifaire optimisant la rentabilité dans le paradigme actuel et dans le nouveau paradigme.

3.1.1 Définition du problème d'optimisation

3.1.1.1 Expression formelle du problème d'optimisation

Nous commençons par introduire les notations utilisées dans l'expression du problème.

Pour un assureur $\mathcal{A}\mathcal{S}\mathcal{S}$ donné, le cadre de souscription en assurance emprunteur est définie par $\mathcal{S}_{\mathcal{A}\mathcal{S}\mathcal{S}} = \{\mathcal{X}_{\mathcal{A}\mathcal{S}\mathcal{S}}, \mathcal{K}_{\mathcal{A}\mathcal{S}\mathcal{S}}, \mathcal{J}_{\mathcal{A}\mathcal{S}\mathcal{S}}, \mathcal{N}_{\mathcal{A}\mathcal{S}\mathcal{S}}, \mathcal{E}_{\mathcal{A}\mathcal{S}\mathcal{S}}\}$ où :

- $\mathcal{X}_{\mathcal{A}\mathcal{S}\mathcal{S}} \subset \mathbb{N}$ désigne les âges acceptés pour un assuré,
- $\mathcal{K}_{\mathcal{A}\mathcal{S}\mathcal{S}} \subset \mathbb{R}$ désigne les niveaux de capitaux garantissables pour un assuré,
- $\mathcal{J}_{\mathcal{A}\mathcal{S}\mathcal{S}} \subset \mathbb{R}$ désigne les taux d'intérêts des crédits pouvant être couverts,
- $\mathcal{N}_{\mathcal{A}\mathcal{S}\mathcal{S}} \subset \mathbb{N}$ désigne les durées des crédits pouvant être couverts,
- $\mathcal{E}_{\mathcal{A}\mathcal{S}\mathcal{S}} \subset \mathbb{E}$ désigne l'ensemble des autres éléments de la politique de souscription (professions autorisées, pratiques sportives autorisées, caractéristiques de l'opération financière à couvrir...) (avec \mathbb{E} désignant l'ensemble des caractéristiques possibles des clients et des opérations financières à couvrir par une assurance).

Les caractéristiques de souscription d'un client entrant dans le cadre de souscription de l'assureur $\mathcal{A}\mathcal{S}\mathcal{S}$ pour la couverture d'une opération de crédit donnée sont notées $s = (x \in \mathcal{X}_{\mathcal{A}\mathcal{S}\mathcal{S}}, K \in \mathcal{K}_{\mathcal{A}\mathcal{S}\mathcal{S}}, i \in \mathcal{J}_{\mathcal{A}\mathcal{S}\mathcal{S}}, n \in \mathcal{N}_{\mathcal{A}\mathcal{S}\mathcal{S}}, e \in \mathcal{E}_{\mathcal{A}\mathcal{S}\mathcal{S}})$. On a évidemment $s \in \mathcal{S}_{\mathcal{A}\mathcal{S}\mathcal{S}}$.

L'approche tarifaire

- d'un assureur $\mathcal{A}\mathcal{S}\mathcal{S}$,
- pour un client et une opération de crédit ayant les caractéristiques de souscription $s \in \mathcal{S}_{\mathcal{A}\mathcal{S}\mathcal{S}}$ est définie par :

- $(a_t^{\mathcal{A}\mathcal{S}\mathcal{S},s})_{t \in \mathbb{N}}$ où $a_t^{\mathcal{A}\mathcal{S}\mathcal{S},s}$ est l'annuité demandée en année t par l'assureur à l'assuré en couverture de l'opération en question.
- On note $P_{\mathcal{A}\mathcal{S}\mathcal{S}}$ la politique tarifaire d'un assureur $\mathcal{A}\mathcal{S}\mathcal{S}$, que l'on définit comme l'ensemble des suites d'annuités demandées pour les clients et opérations relevant de son cadre de souscription : $P_{\mathcal{A}\mathcal{S}\mathcal{S}} = \left\{ (a_t^{\mathcal{A}\mathcal{S}\mathcal{S},s})_{t \in \mathbb{N}} \right\}_{s \in \mathcal{S}_{\mathcal{A}\mathcal{S}\mathcal{S}}}$.

La politique tarifaire d'un assureur $\mathcal{A}\mathcal{S}\mathcal{S}$ est un sous-ensemble de l'ensemble des politiques tarifaires possibles dans le cadre de souscription fixé, que l'on note $\mathcal{P} = (\mathbb{R}^{\mathbb{N}})_{(s \in \mathcal{S}_{\mathcal{A}\mathcal{S}\mathcal{S}})}$.

On prend par ailleurs en compte le jeu de contraintes (marché, juridique...) que doit respecter l'approche tarifaire d'un assureur via la notation suivante : on note \mathcal{C} l'ensemble des approches tarifaires admissibles du point de vue de ces contraintes. Autrement dit, une approche tarifaire admissible est un élément de l'ensemble $\mathcal{P} \cap \mathcal{C}$.

On note enfin $R(P, \text{paradigm})$ la rentabilité espérée par l'assureur ayant l'approche tarifaire P dans le paradigme paradigm (cette dernière variable valant *current* pour la situation actuelle dans laquelle la transférabilité n'est pas permise, *new* pour la situation étudiée dans laquelle la transférabilité est permise).

Mathématiquement, le problème que nous cherchons à résoudre consiste à calculer la différence Δ entre le résultat maximal espéré dans le nouveau paradigme et le résultat maximal espéré dans le paradigme actuel, soit :

$$\Delta = \sup_{P \in \mathcal{P} \cap \mathcal{C}} (R(P, \text{paradigm} = \text{new})) - \sup_{P \in \mathcal{P} \cap \mathcal{C}} (R(P, \text{paradigm} = \text{current})),$$

Pour calculer S , une étape importante consiste à rechercher l'optimum de la fonction R sur deux ensembles :

$$\mathcal{A} = (P \in \mathcal{P} \cap \mathcal{C} ; \text{paradigm} = \text{new}),$$

d'une part, et sur

$$\mathcal{A}' = (P \in \mathcal{P} \cap \mathcal{C} ; \text{paradigm} = \text{current}),$$

d'autre part.

Cette recherche d'optimum d'une fonction est appelée optimisation sous contrainte. Selon la forme de la fonction R et de l'ensemble de recherche \mathcal{A} , ce problème est plus ou moins complexe. Il est en général possible de trouver un maximum local via un algorithme de maximisation duale, par exemple le Lagrangien augmenté dans le cas où R est différentiable et \mathcal{A} est euclidien (Cerf, 2018). Trouver un maximum global est en général un problème beaucoup plus complexe qui fait encore l'objet de nombreuses études.

Pour simplifier le problème, nous allons restreindre \mathcal{A} et \mathcal{A}' à un espace euclidien uni-dimensionnel et R une fonction continue sur ces espaces, à valeur dans \mathbb{R} , de telle sorte que le problème soit intuitif, puisse être représenté graphiquement, et comporte nécessairement une solution.

3.1.1.2 Indicateur à optimiser et contraintes

La maximisation des profits futurs générés sera recherchée sous deux contraintes : l'une reflétant l'acceptabilité du business du point de vue de l'entreprise, l'autre reflétant l'acceptabilité du produit du point de vue du client.

Pour résumer en un indicateur à optimiser la maximisation des profits futurs générés, nous utiliserons la NBV (*New Business Value*). Celle-ci mesure les résultats techniques et financiers après impôts de la cohorte. Son avantage réside :

- dans la facilité à la calculer dans un contexte où le business envisagé ne déforme pas le profil de risque de l'entreprise,
- dans le fait qu'elle soit dénuée d'ambiguïté : une augmentation de la NBV reflète de manière univoque une augmentation des profits futurs.

En revanche, elle présente les biais suivants :

- en premier lieu, elle n'est pas totalement alignée avec les indicateurs présentés au marché : la communication financière des compagnies d'assurance axe son discours général sur le résultat opérationnel net de l'année en cours, le ratio NBV/PVNBV de l'année et les projections de résultat opérationnel à horizon 3 ans.
- en second lieu, elle est assez dépendante d'hypothèses financières qui la rendent volatile.

Pour traduire les contraintes sous lesquelles nous souhaitons travailler, nous utiliserons :

- **La contrainte Assureur**, reflétant l'acceptabilité du produit du point de vue de l'entreprise : un seuil minimal de ratio NBV/PVNBP, qui traduit les processus de validation des nouveaux produits tels que fixés dans les gouvernances de lancement de l'entreprise ;
- **La contrainte Client**, reflétant l'acceptabilité du produit du point de vue des assurés : une déformation maximale du tarif en comparaison du tarif actuellement pratiqué, cette déformation maximale étant imposée au global (mesurée sur la totalité de la durée de l'assurance) et en local (sur une période limitée autour de tout instant t donné). Ce choix traduit l'intuition que sur un marché traditionnel, une différence importante entre le tarif proposé et les tarifs habituellement disponibles pourrait avoir un effet délétère.

3.1.1.3 Fixation des degrés de liberté possibles pour l'optimisation

L'optimisation consiste à déformer le tarif payé par le client. Pour fixer cette déformation, nous nous basons sur une courbe de déformation qui consiste en une série de coefficients à appliquer, année après année, au tarif issu de nos calculs de prime annuelle reflétant l'estimation instantanée du risque.

Pour simplifier notre problème, nous modéliserons la courbe de déformation du tarif de la manière suivante :

$$P_{x,K,i,n}^{Déformé} = P_{x,K,i}^{Base} \times f(\rho, i, n) \times c(i, n),$$

Où

- $P_{x,K,i,n}^{Base}$ est le tarif technique initial (non déformé) en année i pour un client d'âge x , empruntant un capital K sur n années,
- $c(i, n)$ est une courbe de déformation usuelle considérée comme fixe dans notre problème (issue de l'expérience),
- f est une fonction de déformation de la courbe de déformation de base, que nous choisissons de la forme suivante :

$$f(\rho, i, n) = \begin{cases} \rho + (1 - \rho) \frac{i - 1}{\frac{n}{2} - 1} & \text{si } i < \frac{n}{2} \\ 1 & \text{si } i = \frac{n}{2} \\ \frac{1}{\rho^2} + \left(1 - \frac{1}{\rho^2}\right) \frac{i - n}{\frac{n}{2} - n} & \text{si } i > \frac{n}{2} \end{cases},$$

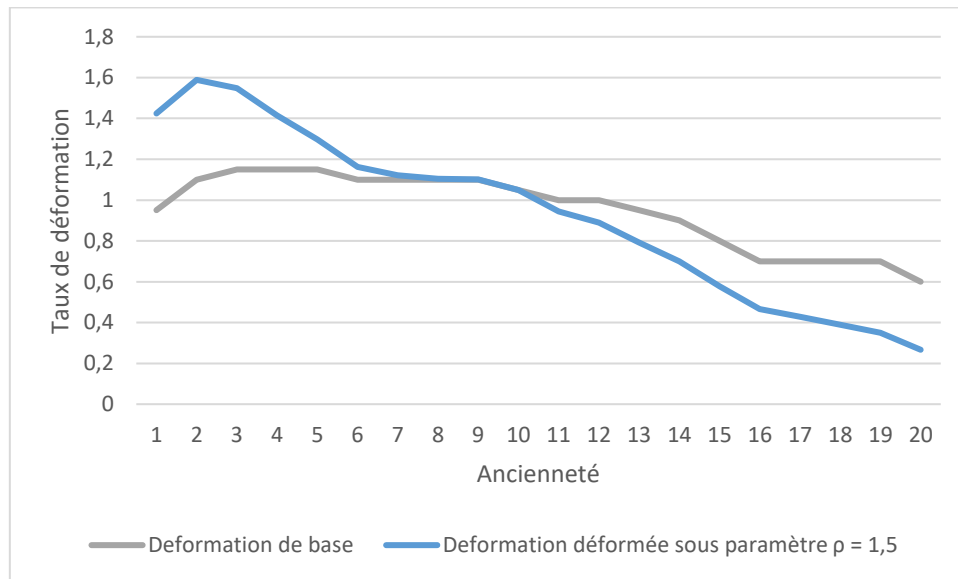
- ρ est le coefficient que nous allons rechercher.

En d'autres termes la fonction de déformation f est égale à :

- Si $i < n/2$: une interpolation linéaire entre 1 et $n/2$ qui ont pour valeur respectives ρ et 1,
- Si $i > n/2$: une interpolation linéaire entre $n/2$ et n qui ont pour valeur respectives 1 et $1/\rho^2$.

L'introduction de cette fonction f permet de faire dépendre la déformation de la courbe de déformation de base d'un paramètre unique ρ , que l'on va chercher à optimiser dans le problème suivant.

Figure 73 : représentation de la courbe de déformation de base et de la courbe de déformation déformée par un paramètre ρ



Nous obtiendrons donc un unique paramètre ρ identique pour tous les clients, indépendamment de leurs caractéristiques et des caractéristiques du contrat d'assurance.

3.1.1.4 Hypothèses à utiliser pour les calculs de rentabilité

Les calculs de rentabilité dans le paradigme actuel et dans le nouveau paradigme que nous souhaitons produire s'appuient sur des hypothèses concernant le marché de l'assurance, les frais liés à la souscription des contrats, les lois de décès et de résiliation ainsi que leur évolution dans le nouveau paradigme. Nous détaillons dans cette section ces différentes hypothèses, les raisons de nos choix et leurs limites.

3.1.1.5 Analyse des données du marché du crédit immobilier et notations

Pour estimer les hypothèses à prendre pour notre modèle, nous devons tout d'abord analyser les chiffres du marché de l'assurance emprunteur à partir des chiffres de l'immobilier, afin d'identifier et chiffrer les différentes causes de résiliations et sources d'affaires nouvelles. Nous exploiterons ensuite ces chiffres dans les hypothèses présentées dans les parties suivantes.

Les applications numériques sont faites à partir des données publiées par la Banque de France pour la période allant de 01/2018 à 01/2020.

L'encours de crédit immobilier (en €) en début de la période étudiée (A) est impacté par :

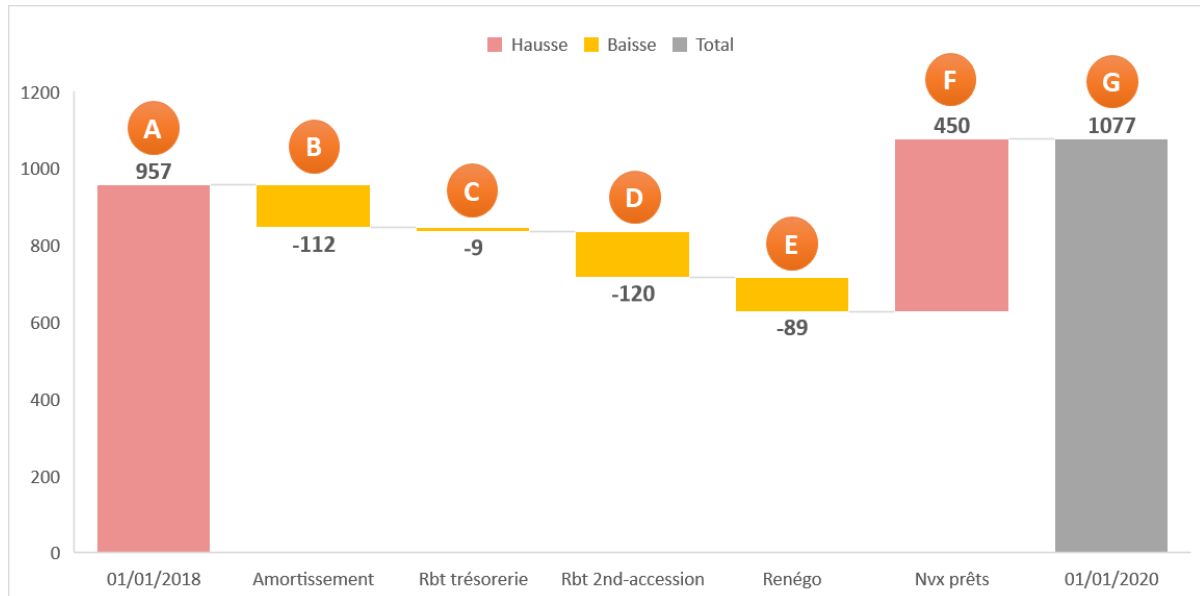
- l'amortissement des crédits (B),
- les remboursements anticipés par des clients disposant de liquidités (à l'exclusion des cas de vente du bien immobilier sous-jacent) (C),
- les remboursements anticipés par des clients disposant de liquidités issues du remboursement du bien immobilier sous-jacent (D),
- les renégociations (E),

- les nouveaux prêts (F).

L'encours de crédit immobilier en fin de la période étudiée (G) est :

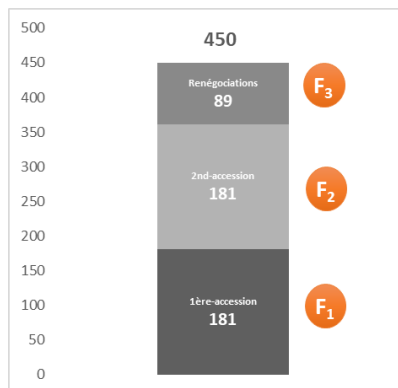
$$G = A + B + C + D + E + F .$$

Figure 74 - Evolution des encours de crédit entre 2018 et 2019



Les données A, E, F, G sont données par la Banque de France, il nous faut estimer les données B, C, D .

Pour cela, focalisons-nous sur la décomposition de F : $F = F_1 + F_2 + F_3$, avec :



- F_1 : prêts liés à la primo-accession
- F_2 : prêts liés à la secondo-accession (consécutifs, en général à un remboursement anticipé d'un précédent prêt, permis par la vente d'un bien immobilier).
- F_3 : nouveaux prêts liés aux renégociations. Cette donnée est donnée par la Banque de France ($F_3 = -E$).

Les spécialistes du marché immobilier indiquent en général une répartition à parts égales des primo et secondo-accédants, ce qui se traduit par $F_1 = F_2$

On en déduit donc : $F_1 = F_2 = \frac{F-F_3}{2}$.

D correspond au remboursement anticipé des secundo-accédants souscrivant un nouveau prêt. Les experts du marché estiment en général que le montant moyen des prêts remboursés par anticipation est égal aux 2/3 du montant moyen du nouveau prêt souscrit. Ceci est cohérent avec l'hypothèse canonique d'un prêt de 180K€ (montant moyen des prêts) souscrit pour 20 ans au taux de 2% et remboursé par anticipation au bout de 8 ans, date à laquelle le capital restant dû est de 116 k€ (soit les 2/3 du montant

initial emprunté). Il serait possible d'affiner ce rapport à partir d'un model point de prêts et en faisant des hypothèses sur la durée des prêts (en s'appuyant sur les lois de résiliation observées en assurance). En exploitant ce rapport, on en déduit que $D = -2/3 \times F_2$.

Par ailleurs, on peut estimer B , l'amortissement des crédits avec une approche simplifiée similaire : on considère une répartition des prêts en cours à parts égales entre les prêts de durée 10, 15, 20 et 25 ans ; on considère un âge moyen pour ces prêts de 6 ans et des taux de 2%. L'amortissement constaté dans la sixième année pour ces prêts est de 11,7% du capital restant dû, soit rapporté à notre assiette de crédits en cours : $B = 11,7\% \times A$.

De manière parallèle au calcul de D , il serait possible d'affiner cette approximation en faisant des hypothèses de durée et de taux de crédit appliquées à l'encours à l'aide d'un model point.

Enfin, C se déduit simplement par $C = G - A - B - D - E - F$.

Exploitions ces résultats pour trouver la correspondance dans les évolutions de stock et d'affaires nouvelles (AN) en assurance Emprunteur.

Par analogie avec l'évolution d'encours de crédit décrite ci-dessus, le stock d'assurance emprunteur (en nombre) en début de la période étudiée (A') est impacté par :

- les remboursements anticipés par des clients disposant de liquidités (à l'exclusion des cas de vente du bien immobilier sous-jacent) (C'),
- les remboursements anticipés par des clients disposant de liquidités issues du remboursement du bien immobilier sous-jacent (D'),
- les renégociations (E'),
- les nouvelles assurances (F'), avec $F' = F'_1 + F'_2 + F'_3$ correspondant aux trois composantes exposées ci-dessus (primo-acquisitions, secundo-acquisitions, renégociations).

On note M' le montant moyen assuré qui vaut 180 K€.

Pour notre étude, on introduit enfin un facteur x qui représente l'efficacité du nouveau paradigme pour conserver les assurances en cours. Dit autrement : dans le paradigme proposé, l'assurance actuellement détenue par un assuré pourra lui convenir pour un nouveau prêt avec une probabilité de x . Les $(1 - x)$ autres cas correspondront au cas où l'assurance, pour des raisons de contenu des garanties, lourdeurs administratives... ne saurait être conservée par l'assuré. Le paradigme actuel correspond au cas où $x = 0$.

Dans le nouveau paradigme, le stock d'affaires emprunteur total sur le marché et son évolution se déduit ainsi des éléments calculés sur le marché de l'immobilier et des notations introduites ci-dessus :

$$\begin{aligned} A' &= A/M', \\ C' &= C/M', \\ D' &= (1 - x) \times D/M', \\ E' &= (1 - x) \times E/M', \\ F'_1 &= F_1/M', \\ F'_2 &= (1 - x) \times F_2/M'. \end{aligned}$$

Enfin, on sait que les renégociations ne conduisent pas toutes à une chute du contrat d'assurance (en particulier dans les cas des renégociations menées dans un même établissement financier) : faute de mesure de cet aspect, nous faisons l'hypothèse que 50% des renégociations conduisent à une résiliation du contrat d'assurance. Cela donne :

$$F'_3 = x \times 50\% \times M'.$$

Compte tenu des difficultés administratives et des cas particuliers que ne manqueront pas de soulever les banques émettant des nouveaux prêts, l'atteinte d'un niveau $x = 1$ dans le nouveau paradigme semble illusoire et un objectif à très long terme. Pour poursuivre l'exercice de manière théorique et sans miser sur une efficacité immédiate et entière de l'option de transférabilité donnée au client, nous étudions un scénario d'efficacité moyenne dans laquelle $x = 0,5$ dans le nouveau paradigme.

Les valeurs obtenues dans ce cadre et utilisées dans la modélisation sont présentées ci-dessous.

Figure 75 - Table des paramètres descriptifs du marché du crédit immobilier et de l'assurance emprunteur dans le paradigme actuel et dans le nouveau paradigme

Paramètre	Valeur (en Mds € d'encours de crédits immobiliers)	Valeur' (en millions de contrats d'assurance) (x=0)	Valeur' (en millions de contrats d'assurance) (x=0,5)
A	957	5,3	5,3
B	-112	-	-
C	-9	-	-
D	-120	-0,67	-0,33
E	-89	-0,25	-0,13
F_1	180	1	1
F_2	180	1	0,5
F_3	89	0,25	0,13

3.1.1.6 Hypothèse de captation de nouveau business

3.1.1.6.1 Description de l'hypothèse retenue

L'objectif recherché dans ce nouveau paradigme est de s'adapter à un besoin de marché et générer ainsi un avantage concurrentiel pour nos produits :

- Cela conduira à court-terme à capter un business supplémentaire. On note μ_M le coefficient d'augmentation à court-terme des AN liée à cet avantage concurrentiel. μ_M n'est pas calculable à partir des données disponibles. Dans le cadre de la modélisation, des valeurs de 1 à 5 seront testées pour μ_M .
- A long-terme, on s'attend à une adaptation totale du marché, au nouveau paradigme dont nous aurons été précurseurs, conduisant d'une part à la disparition de l'avantage concurrentiel dont nous disposons, et d'autre part à la raréfaction des affaires nouvelles d'assurance emprunteur sur le marché. En effet, comme expliqué ci-dessus, l'assurance actuellement détenue par un assuré pourra lui convenir pour un nouveau prêt, entraînant un potentiel de production nouvelle moindre. La raréfaction à long-terme des affaires nouvelles dans ce nouveau paradigme est caractérisée par un coefficient de diminution de la production nouvelle μ_m calculable, avec :

$$\mu_m = F'(x = 0,5) / F'(x = 0) = 0,72 .$$

- La convergence du marché vers le nouveau paradigme se fera au rythme des décisions d'adaptation de leurs offres par les concurrents. On note T_c la durée de convergence du marché, auquel la raréfaction des affaires nouvelles aura atteint le niveau stable μ_m . $T_c = T_d + T_p$ où :
 - o T_d correspond au temps de décision par l'ensemble des concurrents d'adaptation de leur offre au nouveau paradigme (ce temps peut aller de 0 à l'infini). Passé T_d , les affaires nouvelles en emprunteur bénéficient toutes sur le marché de la faculté de transférabilité,

entraînant une raréfaction des affaires nouvelles captables, comme expliqué ci-dessus. On testera dans la suite des temps T_d allant de 5 à 20 ans.

- T_p est le temps d'effet à plein de cette adaptation sur les affaires du marché. La raréfaction expliquée ci-dessus est en effet graduelle, car les affaires nouvelles captées chaque année correspondront encore pendant un certain temps pour partie à des renégociations de prêts anciens qui ne bénéficient pas de cette option. Compte tenu de la durée maximale des prêts octroyés sur le marché (30 ans), l'effet à 100% sera atteint pour $T_p = 30$. En pratique, une fois T_d passé, c'est-à-dire une fois prise par l'ensemble des acteurs du marché la décision de s'aligner en ce qui concerne la transférabilité, l'effet de raréfaction des affaires nouvelles ne sera pas linéaire jusqu'à $T_p = 30$: la raréfaction se fera à un rythme dépendant de la distribution des durées initiale des prêts et des résiliations. Pour fixer T_p avec une plus grande significativité, il serait judicieux de fixer le temps d'effet à plein T_p comme le délai au-delà duquel un seuil (par exemple 90%) de l'effet maximal de la mesure sur la raréfaction des affaires nouvelles est atteint. Faire l'hypothèse d'un T_p très élevé (proche de 30 ans) revient à faire durer très longtemps l'avantage concurrentiel. Pour être prudent, on testera dans la suite des temps T_p allant de 5 à 30 ans.
- On fait enfin l'hypothèse que le niveau maximal μ_M de business supplémentaire est atteint à $T_c/2$, après la décision initiale d'un assureur d'offrir cette option, c'est-à-dire à mi-chemin de la convergence. La logique derrière cette hypothèse est celle d'un alignement progressif de la concurrence entraînant la disparition graduelle de l'avantage concurrentiel offert par cette option, auquel s'ajoute un temps d'effet progressif de l'alignement mesuré par avant l'effet total atteint à T_p .
- T_c n'est pas calculable dans la mesure où il dépend de prises de décision de concurrents. Dans le cadre de la simulation, les valeurs suivantes seront testées pour T_c : simulation de 10 ans à 50 ans.

3.1.1.6.2 Limites et biais de l'hypothèse retenue

L'estimation du business supplémentaire capté à court-terme est complexe et pourrait se faire par sondage (a priori ou a posteriori) et par expérimentation pour mesurer, tous autres paramètres inchangés, l'effet d'attraction de la nouvelle possibilité offerte.

Par ailleurs le calcul de gain court-terme lié à l'avantage concurrentiel et de la perte engendrée par la baisse durable des AN postérieurement à l'adoption par la concurrence des mêmes possibilités doit être pris avec prudence car d'autres paramètres, non pris en compte, pourraient jouer : vitesse d'adaptation de la concurrence, choix de certains concurrents de ne pas s'aligner, différences dans l'optimum entre croissance et rentabilité recherché par les acteurs du marchés...

3.1.1.7 Hypothèse du coût de la sélection médicale

3.1.1.7.1 Description de l'hypothèse retenue

Le coût de la sélection médicale est modélisé via l'application de frais généraux exprimés en % des primes captées, comme dans le paradigme actuel.

3.1.1.7.2 Limites et biais de l'hypothèse retenue

Cette hypothèse conduit à ne pas prendre en compte la déformation de la cohorte captée. Pour en tenir compte, il faudrait modéliser un coût des formalités médicales en considérant l'application des règles actuelles de sélection (seuils de déclenchement des formalités médicales, etc...) à une cohorte légèrement différente, en faisant l'hypothèse que la nouvelle faculté attire davantage les jeunes (qui auront l'occasion de l'utiliser), entraînant une déformation du mix de production.

Il pourrait aussi être intéressant de simuler une dégradation de l'état de santé général initial de la cohorte captée, causée par l'effet anti-sélectif de cette faculté de transfert sur les clients (des clients en plus mauvaise santé seraient attirés par ce contrat qui leur permettrait de bénéficier à long-terme d'une assurance même en changeant de bien et en cas de dégradation ultérieure supplémentaire de leur état de santé).

3.1.1.8 Loi de résiliation

3.1.1.8.1 Description de l'hypothèse retenue

La loi de résiliation actuellement observée sera amendée pour tenir compte de deux éléments :

- L'absence de résiliation pour cause de transfert.
- Des résiliations supplémentaires liées à la déformation du tarif actuel, qui peut générer des opportunités pour des clients d'utiliser les possibilités ouvertes par l'amendement Bourquin pour résilier en cours de prêt leur contrat d'assurance.

On définira la nouvelle loi de résiliation de la manière suivante :

$$TT = \max(0; \min(1; TS + TD)),$$

Où :

- TT est le taux de résiliation annuel total compris entre 0 et 1.
- TS est le taux de résiliation *statique*, c'est-à-dire le taux de chute structurellement lié au marché de l'immobilier et de l'assurance emprunteur, indépendamment de la concurrence offerte conjoncturellement par d'autres produits d'assurance emprunteur. En première approche, cette loi de résiliation statique correspond à la loi de résiliation modélisée dans la partie précédente. Dans le nouveau paradigme dans lequel la transférabilité du contrat d'assurance est permise, la loi de résiliation statique correspond à la loi de résiliation modélisée dans la partie précédente retraitée des résiliations causées par un transfert.
- TD est le taux de chute *dynamique*, c'est-à-dire un taux de chute conjoncturel correspondant à l'impact de la comparaison par les clients de leurs primes restantes à payer avec les produits disponibles à l'instant donné de comparaison sur le marché. Cette comparaison avec la concurrence disponible sur le marché s'exerce annuellement grâce à l'exercice du droit à résiliation offert par l'amendement Bourquin. Nous appelons ce taux *dynamique* car il est dépendant des conditions de marché (offres de la concurrence) et que nous le construirons par analogie avec la loi de rachat dynamique en assurance vie.

a) Modélisation des résiliations pour causes de transfert

Les résiliations évitées grâce à la faculté de transférabilité sont résumées par un coefficient η de diminution des chutes, où

$$\eta = \frac{C'(x = 0,5) + D'(x = 0,5) + E'(x = 0,5)}{C'(x = 0) + D'(x = 0) + E'(x = 0)} = 50\%$$

Avec

- C : les remboursements anticipés par des clients disposant de liquidités (à l'exclusion des cas de vente du bien immobilier sous-jacent),
- D : les remboursements anticipés par des clients disposant de liquidités issues du remboursement du bien immobilier sous-jacent,
- E : les renégociations,
- $x = 0,5$ correspondant au nouveau paradigme ; $x = 0$ correspondant au paradigme actuel.

Nous abattons donc la loi de résiliation de 50% pour tenir compte des chutes évitées par cette possibilité de transfert.

b) Modélisation des résiliations pour cause de concurrence

Par analogie avec la loi de rachat dynamique en assurance vie issue du QIS 5, nous construisons une loi de chute pour estimer 3 éléments :

- l'impact de la déformation de nos tarifs sur la perception de leur produit par nos clients,
- l'impact de la comparaison avec la concurrence,
- les résiliations supplémentaires qu'entraîneront l'exercice de leur droit annuel de résiliation Bourquin dans ce cadre.

Le taux de chute dynamique sera ainsi mesuré à chaque pas de temps en se basant sur la comparaison entre les primes restantes à payer par le client sur la durée du prêt et les prix disponibles sur le marché. Nous considérerons ces derniers comme égaux aux primes à payer par le client sur la durée restante du prêt dans le cadre du tarif actuel.

Par analogie avec la loi de rachat dynamique en assurance vie, nous souhaitons modéliser :

- Une sensibilité du client à un écart de prix positif (possibilité de faire des économies en s'adressant au marché).
- Une sensibilité du client à un écart de prix négatif (impossibilité de faire des économies en s'adressant au marché, entraînant une baisse du taux de résiliation).
- Un seuil de sensibilité du client à l'écart de prix faible, c'est-à-dire l'absence d'impact sur la loi de chute d'un écart (positif ou négatif) réduit.
- Un seuil d'insensibilité du client à un écart de prix supplémentaire lorsque l'écart de prix est déjà significatif, c'est-à-dire l'absence d'augmentation ou de diminution des chutes au-delà d'un certain écart de prix. On en déduit alors un taux dynamique maximal de résiliation (ce taux représente le taux de chute des contrats lorsque la comparaison des prix est particulièrement défavorable ; il est fixe, c'est-à-dire qu'au-delà d'un certain écart entre la prime restante et le prix de marché, ce taux s'applique) et un taux dynamique minimal de résiliation.

La formule de calcul du taux de résiliation dynamique est la suivante :

$$Taux\ dynamique = \begin{cases} T_{max} & \text{si } \sum (P - P_{marché}) > \alpha \\ T_{max} \times \frac{\sum (P - P_{marché})}{\alpha - \beta} & \text{si } \beta < \sum (P - P_{marché}) < \alpha \\ 0 & \text{si } \gamma < \sum (P - P_{marché}) < \beta \\ T_{min} \times \frac{\sum (P - P_{marché}) - \gamma}{\delta - \gamma} & \text{si } \delta < \sum (P - P_{marché}) < \gamma \\ T_{min} & \text{si } \sum (P - P_{marché}) < \delta \end{cases}$$

avec les valeurs suivantes issues de nos entretiens avec des experts du sujet :

Paramètre	Valeur
T_{max}	90%
T_{min}	0%
α	10 000 €
β	100 €
γ	-150 €
δ	-15 000 €

Pour fixer les seuils symbolisés par les lettres grecques, nous nous appuyons notamment sur les études publiées lors des débats sur l'amendement Bourquin, dans lesquelles avaient été estimés des marchés potentiels en faisant des hypothèses sur ces indicateurs (cf. Annexe 10).

3.1.1.8.2 Limites et biais de l'hypothèse retenue

La loi de résiliation est modélisée comme la somme entre les résiliations intrinsèques et les résiliations pour cause de concurrence.

Les résiliations intrinsèques sont calculées comme les résiliations actuelles diminuées des résiliations évitées par la faculté de transfert, c'est-à-dire, les chutes pour renégociation du prêt conduisant à un rachat du prêt et les chutes pour changement de bien avec rachat du prêt.

Les statistiques de la banque de France donnent la part des renégociations dans les nouveaux prêts. Nous savons que deux biais existent dans l'estimation des chutes évitées :

1. la part des renégociations qui ne conduisent pas à une chute du contrat d'assurance n'est pas connue (elle est prise forfaitairement à 50% dans notre modélisation) ,
2. nous faisons l'hypothèse que ces chutes sont constantes au cours des années,

3.1.1.9 Lois de décès

La loi de décès modélisée dans la partie précédente est utilisée pour les simulations et l'optimisation du tarif dans le nouveau paradigme. Nous considérerons donc que la probabilité de décès d'un emprunteur à un âge donné reste la fonction d'une probabilité de base, identique à celle mesurée dans le précédent paradigme, à laquelle des coefficients s'appliquent pour tenir compte du profil de l'assuré (Fumeur / Non Fumeur) et de l'effet de la sélection médicale. L'hypothèse implicite est que la mortalité ne dépend pas significativement de l'ancienneté dans un bien immobilier, qui est égale à l'ancienneté du prêt (pour les acquisitions) ou à l'ancienneté du prêt postérieurement à la renégociation pour les assurés en cours de prêt.

3.1.2 Présentation des résultats des modélisations

Le problème d'optimisation est réduit à l'identification du paramètre ρ de la fonction de déformation permettant d'optimiser la NBV générée par une cohorte.

Pour pouvoir prendre une décision quant à la mise en place du nouveau paradigme dans la partie 3.2, nous procédons dans cette section au calcul de la NBV d'une cohorte modèle d'affaires nouvelles dans le paradigme actuel, d'une part, et dans le paradigme futur, d'autre part. Pour cela, les étapes suivantes sont proposées :

- Présentation des paramètres de la modélisation
- Calcul d'un paramètre ρ_1 d'optimisation du tarif dans le paradigme actuel, en exploitant la connaissance nouvelle sur la loi de résiliation des contrats et la loi de mortalité issue des parties précédentes, et estimation de la NBV optimisée dans ce paradigme, notée NBV_1^{opt} , correspondant à la NBV d'une cohorte similaire au portefeuille actuel d'affaires nouvelles.
- Calcul d'un paramètre ρ_2 d'optimisation du tarif dans le nouveau paradigme, en exploitant la connaissance nouvelle sur la loi de durée des contrats et la loi de mortalité issues des parties précédentes, en mettant en œuvre les hypothèses présentées ci-dessous ; estimation de la NBV optimisée dans ce paradigme, notée NBV_2^{opt} , correspondant à la NBV d'une cohorte similaire au portefeuille actuel d'affaires nouvelles (sans prise en compte de la baisse prévisible à long-terme des affaires-nouvelles vers une cohorte de taille $\mu_m \times$ la cohorte initiale)
- Calcul de NBV_3^{opt} , correspondant à la NBV d'une cohorte similaire dans son profil au portefeuille actuel d'affaires nouvelles mais impactée par la baisse prévisible à long-terme des affaires-nouvelles vers une cohorte de taille $\mu_m \times$ la cohorte initiale.

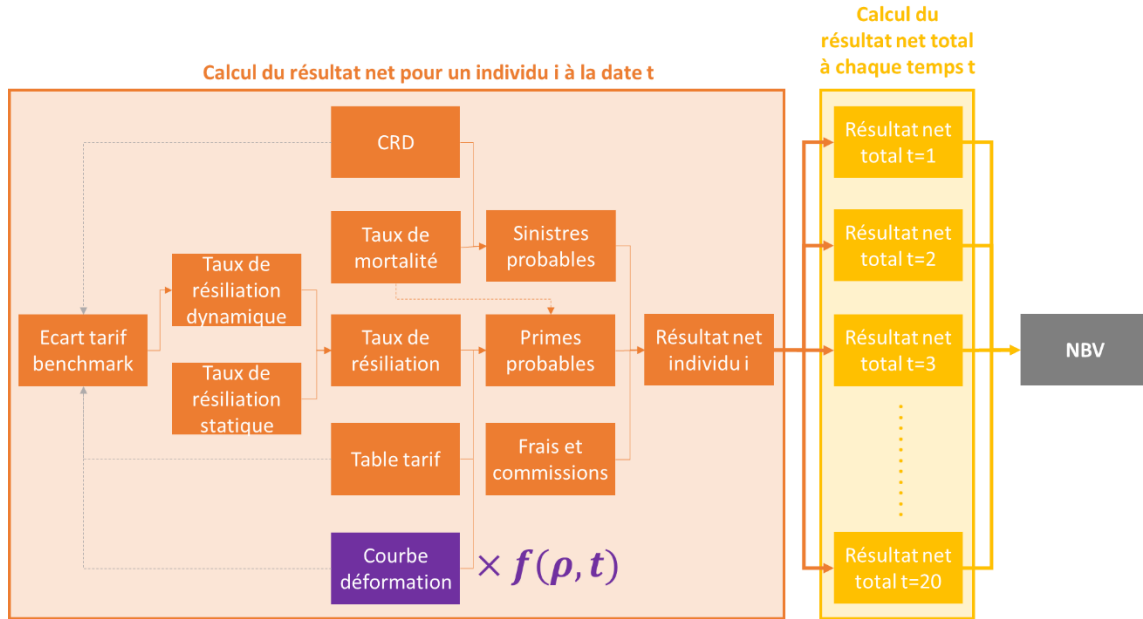
3.1.2.1 Paramètres initiaux de la modélisation

3.1.2.1.1 Description pratique du programme d'optimisation

Dans cette partie nous détaillons de manière synthétique la structure du programme d'optimisation qui va nous servir dans un premier temps à établir un compte de résultat puis à obtenir une NBV optimisée.

Ce graphique liste les briques nécessaires au calcul de la NBV pour un portefeuille d'assuré donné et met en exergue le paramètre ρ , unique levier de notre optimisation. La fonction de déformation f est celle que nous avons défini dans la partie 3.1.1.3.

Figure 76 - Architecture du programme d'optimisation



L'optimisation consiste donc à trouver ρ qui maximise la NBV évaluée sur 20 ans tout en tenant compte des liens multiples qui existent entre chaque brique de calcul. C'est donc cette multitude d'imbrications sur plusieurs dimensions (temporelle t et par individu i) qui en fait un problème complexe. Pour nous aider à le résoudre nous utilisons la fonction *constrOptim.nl* du package R *alabama*. Elle s'appuie sur un algorithme d'optimisation par Lagrangien augmenté pour une fonction non linéaire avec contraintes. Dans le cadre de notre étude, la contrainte que nous appliquons est d'avoir un ratio NBV/PVNBV supérieur au seuil que nous avons défini à 4%.

3.1.2.1.2 Cohorte initiale

Du fait de sa complexité, le programme d'optimisation décrit ci-dessus s'exécute en un temps de calcul relativement long qui augmente considérablement avec la taille du portefeuille initial d'assurés. Nous sommes ainsi limités par la puissance de nos processeurs et c'est pourquoi nous faisons le choix de créer une base d'assurés fictive regroupée en 20 Model Points qui reflètent notre portefeuille réel en termes de ventilation des capitaux. La définition des tranches d'âges est établie de sorte à avoir à partir du portefeuille réel un nombre d'assurés homogène sur chaque classe. Ces Model Points sont présentés en Annexe 11.

Nous faisons l'hypothèse que chaque assuré emprunte un capital sur 20 ans afin de pouvoir être alignés avec la durée de projection du calcul de la NBV. De plus le Courtage étant historiquement le mode de distribution privilégié nous supposons qu'il s'agit l'unique mode de distribution possible du portefeuille fictif. Cela implique donc que nous exploitons pour la suite une seule des lois de résiliations construites précédemment, à savoir celle des durées supérieures ou égales à 3 ans avec distribution via Courtage (Feuille n°4).

3.1.2.1.3 Frais généraux et commissions

Les frais et commissions sont définis de manière arbitraire comme suit :

	Valeur en % du CA
Frais généraux	15%
Honoraires médicaux	5%
Frais de distribution de 1ère année	80%
Frais de distribution de 2ème année et +	30%

3.1.2.2 Recherche du paramètre d'optimisation tarifaire dans le paradigme actuel

Nous exécutons le programme d'optimisation dans le paradigme actuel. Le programme exploite la connaissance nouvelle sur la loi de résiliation des contrats et la loi de mortalité (cf. partie 2) pour calculer un paramètre ρ_1 d'optimisation du tarif dans le paradigme actuel, conduisant à une NBV optimisée, notée NBV_1^{opt} , pour une cohorte similaire au portefeuille actuel d'affaires nouvelles.

La représentation de la NBV en fonction de ρ montre un maximum atteint pour $\rho_1=2,94$, c'est-à-dire pour une déformation du tarif conduisant à faire payer à l'assuré un tarif triplé au début de la période de couverture et un tarif réduit en fin de période de couverture. L'optimisation tarifaire via ce paramètre ρ_1 permet une augmentation de la NBV de 65% en comparaison de la NBV obtenue sans optimisation. Ce résultat est courant et cohérent avec les dires d'experts du marché.

	Sans optimisation	Avec optimisation
NBV	3 007 883	$NBV_1^{opt} = 4 934 635$
NBV/PVNBP	14,6%	20,1%
ρ	1	2,94

Figure 77 - Spectre de la NBV en fonction de ρ

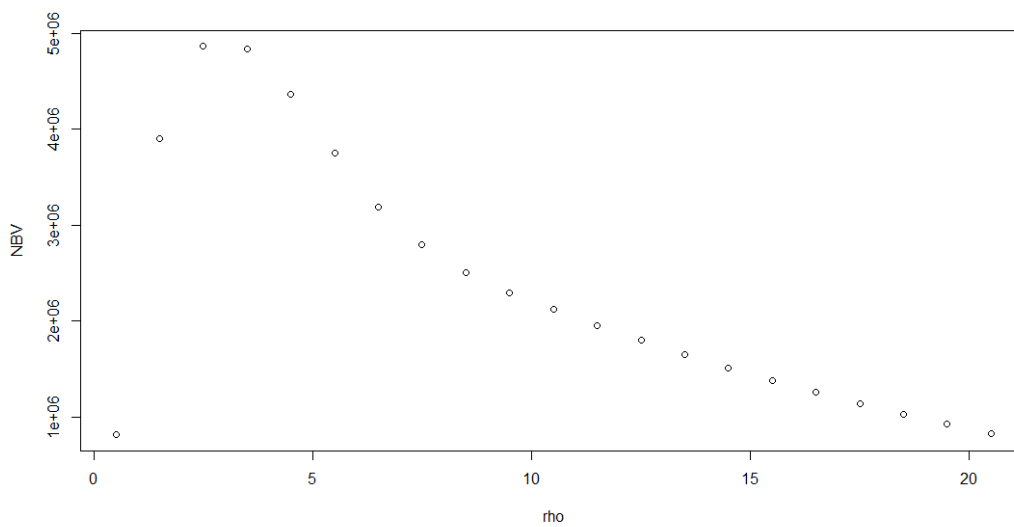


Figure 78 - Spectre de la NBV/PVNBP en fonction de ρ

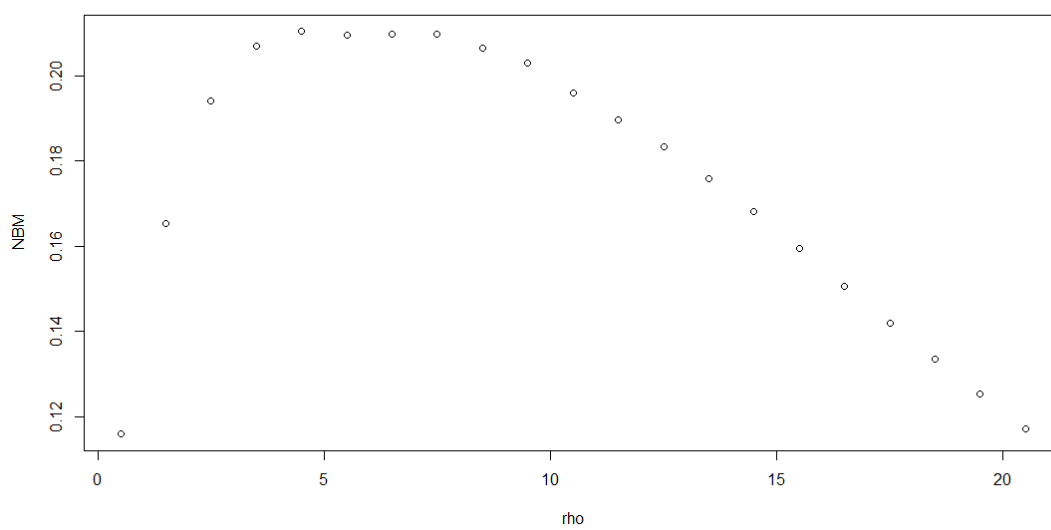
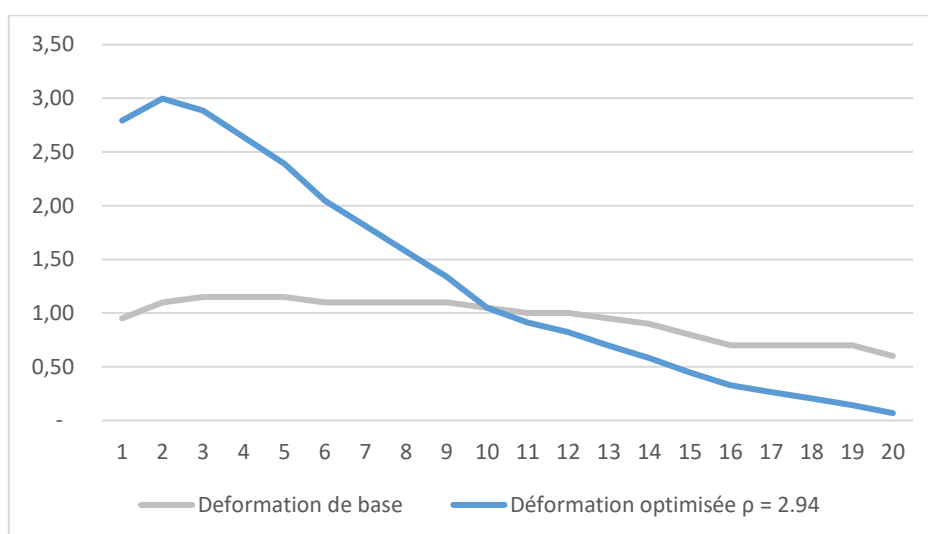


Figure 79 - Courbe de déformation optimisée dans l'ancien paradigme



3.1.2.3 Recherche du paramètre d'optimisation tarifaire dans le nouveau paradigme

Nous nous positionnons maintenant dans le nouveau paradigme. L'application des hypothèses présentées ci-dessus (cf. partie 3.1.1.4) conduit, avant optimisation tarifaire, à une augmentation de NBV de 30%, qui s'explique essentiellement par une plus grande tenue des contrats.

Nous exécutons le programme d'optimisation pour calculer un paramètre ρ_2 d'optimisation du tarif dans le nouveau paradigme, ce qui conduit à une NBV optimisée, notée NBV_2^{opt} , pour une cohorte similaire au portefeuille actuel d'affaires nouvelles. La représentation de la NBV en fonction de ρ montre un maximum atteint pour $\rho_2=2,89$, une valeur très proche de ρ_1 , ce qui est la conséquence du fait que, d'une part, les impacts estimés du nouveau paradigme sur les différentes hypothèses sont constants sur toute la durée d'un prêt, et d'autre part la courbe de déformation ne comporte qu'un seul degré de liberté. Là encore l'application du paramètre d'optimisation tarifaire conduit à faire payer à l'assuré un tarif triplé au début de la période de couverture et un tarif réduit en fin de période de couverture. L'optimisation tarifaire via ce paramètre ρ_2 permet une augmentation de la NBV de 46% en comparaison de la NBV obtenue sans optimisation. La NBV optimisée NBV_2^{opt} n'est supérieure que de 15% à la NBV optimisée du paradigme actuel NBV_1^{opt} , ce qui s'explique par le fait que l'essentiel du chiffre d'affaires est issu des premières années (pendant lesquelles le tarif est fortement déformé), au cours desquelles les cohortes sont assez similaires ; la différence, pourtant très importante, dans la taille des portefeuilles en fin de prêt dans les deux paradigmes a un impact limité sur le chiffre d'affaires total dans la mesure où le tarif est fortement réduit à ce moment.

	Sans optimisation	Avec optimisation
NBV	3 869 588	$NBV_2^{opt} = 5 695 851$
NBV/PVNBP	14,5%	19,6%
ρ	1	2,89

3.2 Modélisation de la rentabilité de l'activité d'assurance emprunteur dans le nouveau paradigme et prise de décision

Nous étudions dans ce chapitre la décision stratégique d'entrer ou non dans le nouveau paradigme du point de vue d'un assureur présent sur le marché de l'assurance emprunteur.

Ce chapitre fait l'hypothèse d'une réponse du marché, plus ou moins rapide, qui entraîne l'existence d'une phase de *transition* sur le marché (avant la totalité de la réponse du marché à la décision du premier acteur ayant décidé de mettre en place la transférabilité) avant une phase *stable*.

L'algorithme de décision proposé est le suivant :

- **Etude de la phase stable** : estimer la NBV NBV_3^{opt} générée par une cohorte d'affaires nouvelles dans la phase stable, en tenant en compte (i) de la NBV optimisée d'une affaire nouvelle telle qu'estimée dans la partie 3.1.2.3, (ii) de la raréfaction des affaires nouvelles telle qu'estimée dans la partie 3.1.1.6 :
 - Si $NBV_3^{opt} > NBV_1^{opt}$, c'est-à-dire si la NBV des cohortes d'affaires nouvelles dans le nouveau paradigme dépasse la NBV des cohortes actuelles d'affaires nouvelles, alors l'entrée dans le nouveau paradigme est profitable.
 - Si $NBV_3^{opt} < NBV_1^{opt}$, c'est-à-dire si la NBV des cohortes d'affaires nouvelles dans le nouveau paradigme est inférieure à la NBV des cohortes actuelles d'affaires nouvelles, alors il faut étudier la phase de transition pour pouvoir se prononcer.
- **Etude de la phase de transition** : l'entrée dans le nouveau paradigme n'est profitable que si la richesse créée dans la phase de transition compense la perte durable et pérenne de la phase stable. La phase de transition est en effet caractérisée par l'avantage concurrentiel donné par cette possibilité offerte aux clients, qui conduit à capter davantage de clients. On calcule la valeur actuelle de la richesse créée lors de cette phase de transition :
 - Si la valeur actuelle de la richesse créée lors de la phase de transition dépasse la valeur actuelle de la richesse détruite dans la phase stable, alors l'entrée dans le nouveau paradigme est profitable.
 - Si la valeur actuelle de la richesse créée lors de la phase de transition est inférieure à la valeur actuelle de la richesse détruite dans la phase stable, alors l'entrée dans le nouveau paradigme n'est pas souhaitable.

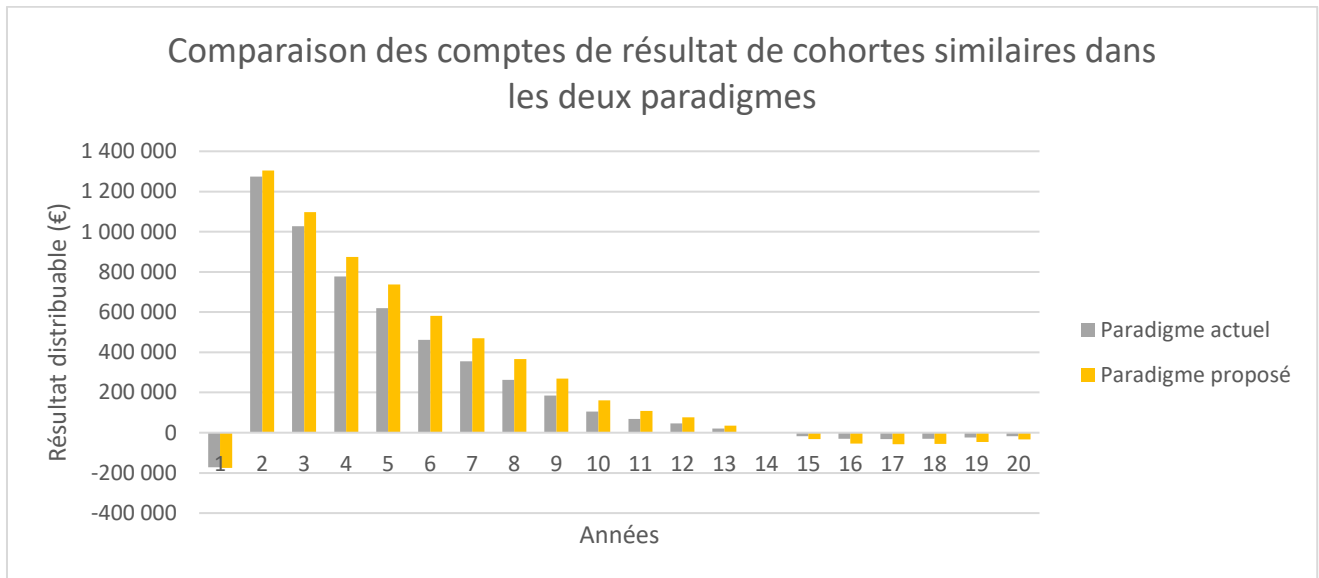
L'output de ce chapitre est

- une projection du compte de résultat de l'activité d'assurance emprunteur dans le nouveau paradigme et sa sensibilité à différentes hypothèses sur la réponse du marché
- des critères objectivables pour la prise de décision,
- une identification des limites de l'analyse et des questions qu'elle suscite.

3.2.1 Etude de la phase stable : comparaison des comptes de résultat des cohortes dans le paradigme actuel et dans le nouveau paradigme

Comparons d'abord les comptes de résultat d'une cohorte identique dans le paradigme actuel et dans le paradigme futur :

Figure 80 - Comparaison des comptes de résultat de cohortes similaires dans les deux paradigmes



Le compte de résultat traduit l'augmentation de NBV (qui est égale à la valeur actualisée des flux représentés), pour une cohorte donnée, permise par le nouveau paradigme et traduite dans les chiffres estimés ci-dessus. Cette augmentation de NBV vient d'une augmentation du résultat constaté dans les 13 premières années de la cohorte :

$$NBV_2^{opt} > NBV_1^{opt} .$$

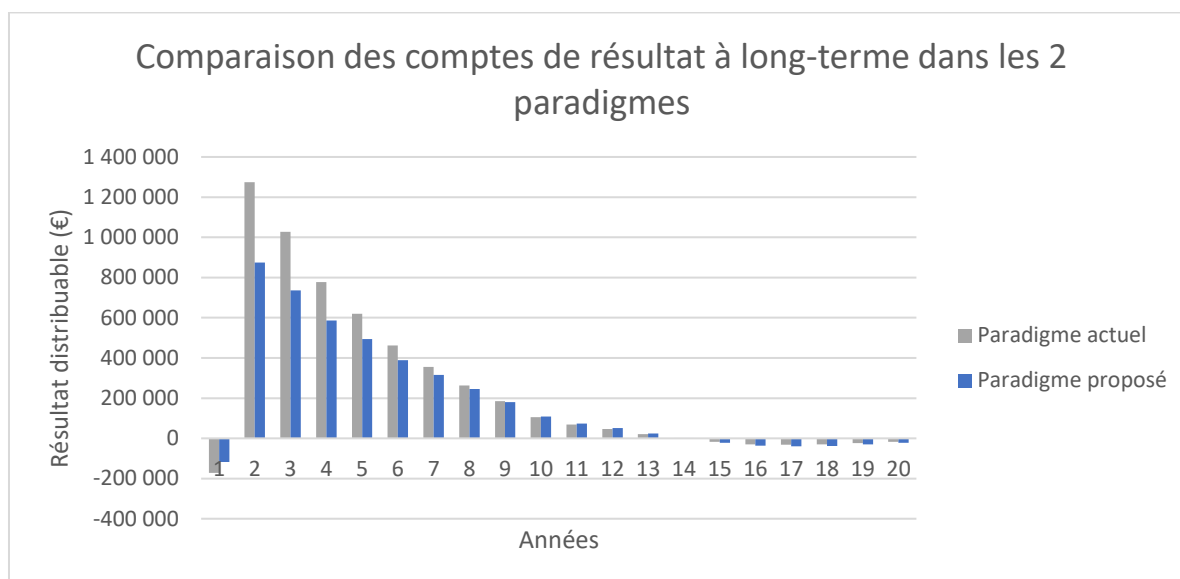
A court-terme (bien avant la durée de convergence du marché T_c), cette différence pourrait être encore plus grande, sous l'effet d'une captation supplémentaire de business, permise par l'avantage concurrentiel donné aux produits disposant de la transférabilité, et mesurée par l'application d'un coefficient multiplicateur de cohorte de μ_M . La NBV de la cohorte dans laquelle on bénéficierait au maximum de l'avantage concurrentiel serait alors :

$$NBV^{opt} = \mu_M \times NBV_2^{opt} .$$

A long-terme en revanche, il faut prendre en compte la baisse prévisible des affaires-nouvelles vers une cohorte réduite à $\mu_m \times$ la cohorte initiale, avec $\mu_m = 0,72$ comme calculé précédemment. Le compte de résultat d'une cohorte dans le nouveau paradigme est alors très inférieur au compte de résultat actuel et la NBV d'une cohorte sera :

$$NBV_3^{opt} = \mu_m \times NBV_2^{opt} .$$

Figure 81 - Comparaison des comptes de résultat à long-terme dans les deux paradigmes



Le gain de NBV par AN dans le nouveau paradigme ne compense pas la baisse en nombre des AN :

$$NBV_3^{opt} < NBV_1^{opt}.$$

On applique l'algorithme de décision exposé ci-dessus : dans la mesure où $NBV_3^{opt} < NBV_1^{opt}$, il faut étudier la phase de transition pour pouvoir se prononcer.

3.2.2 Projection du compte de résultat de l'activité Emprunteur

Nous avons conclu dans la partie précédente que dans le nouveau paradigme, la NBV supplémentaire générée par chaque AN, compte tenu de sa plus grande longévité, ne compenserait pas la raréfaction des AN. A long-terme, le nouveau paradigme est donc destructeur de valeur. Cette destruction de valeur est caractérisée pour chaque cohorte par l'inégalité $NBV_3^{opt} < NBV_1^{opt}$. La valeur totale détruite pour l'activité d'assurance Emprunteur peut se mesurer par l'actualisation des flux futurs des cohortes réduites d'AN captées.

En revanche, le nouveau paradigme génère à court-terme une opportunité, pour les premiers assureurs qui l'adopteront, de capter pendant quelques années (d'ici à l'adoption complète du nouveau paradigme par tout le marché), du business supplémentaire. La richesse générée par cette opportunité peut se mesurer par l'actualisation des flux futurs des cohortes d'AN captées durant cette phase de transition dans laquelle nous bénéficions de cet avantage concurrentiel.

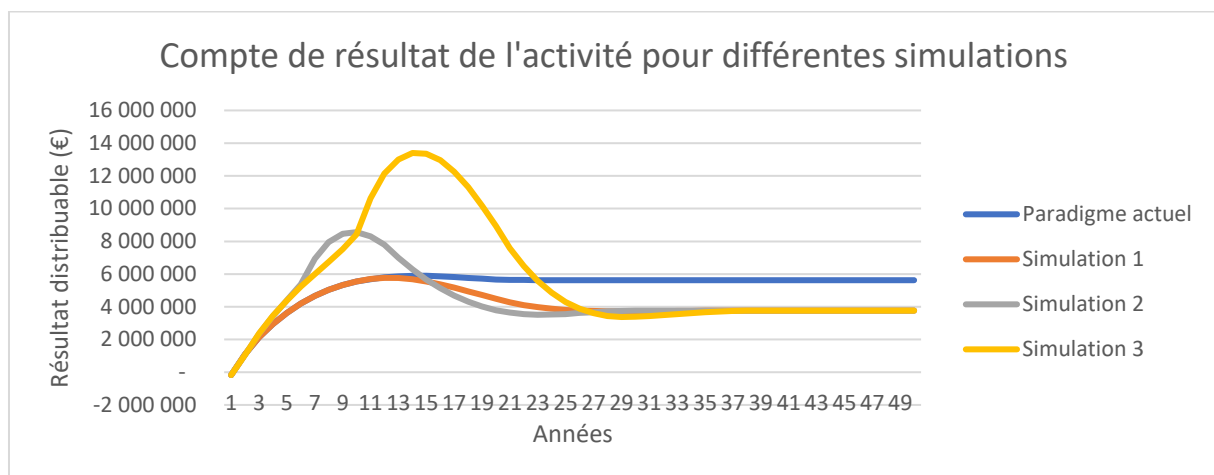
La question est donc de comparer la richesse détruite à la richesse captée pour pouvoir se prononcer sur l'opportunité de cette décision.

Pour calculer en un mouvement la différence entre la richesse détruite et la richesse générée, nous allons comparer le compte de résultat prévisionnel de l'activité d'assurance emprunteur dans le paradigme actuel et dans le paradigme proposé, en testant différentes hypothèses pour les paramètres T_c (temps de convergence vers le niveau d'AN stable du paradigme proposé) et μ_M (business supplémentaire capté à court-terme).

Nous représentons dans le graphique ci-dessous :

- **Paradigme actuel** : Le compte de résultat de l'activité dans le paradigme actuel.
- **Simulation 1** : Le compte de résultat de l'activité dans le paradigme futur, avec $T_c = 20$ et $\mu_M = 1$. Cette simulation correspond à un mouvement suiveur : la décision est déjà prise sur le marché par un concurrent et se diffuse, nous décidons d'adopter le nouveau paradigme, mais trop tardivement par rapport au marché pour pouvoir bénéficier d'un avantage concurrentiel.
- **Simulation 2** : Le compte de résultat de l'activité dans le paradigme futur, avec $T_c = 12$ et $\mu_M = 2$. Cette simulation correspond à une adoption rapide par les concurrents du paradigme, qui ne permet pas de bénéficier très longtemps d'un avantage concurrentiel, mais on fait l'hypothèse que pendant ce court laps de temps, l'avantage concurrentiel est significatif et permet de doubler le business capté.
- **Simulation 3** : Le compte de résultat de l'activité dans le paradigme futur, avec $T_c = 20$ et $\mu_M = 3$. Cette simulation correspond à une adoption progressive par les concurrents du paradigme, qui nous permet de bénéficier d'un avantage concurrentiel important et long.

Figure 82 - Compte de résultat de l'activité pour différentes simulations



On visualise bien sur ces courbes :

- Le positionnement abaissé du résultat généré à long-terme dans le paradigme futur (simulations 1, 2, 3) en comparaison du paradigme actuel, ce qui correspond à la destruction de valeur à long-terme présentée dans la partie 0 et caractérisée par $NBV_3^{opt} < NBV_1^{opt}$ pour chaque cohorte.
- La *bulle* de richesse générée à court-terme dans le paradigme futur pour certaines simulations (simulations 2 et 3), qui s'explique par l'augmentation temporaire, en raison de l'avantage concurrentiel, des affaires captées ($\mu_M > 1$).

Il est intéressant d'observer l'actualisation de résultats futurs du paradigme actuel et des trois simulations pour vérifier si la bulle initiale compense ou non la perte pérenne. Le tableau compare les résultats futurs actualisés en € et en base 1 pour le paradigme actuel.

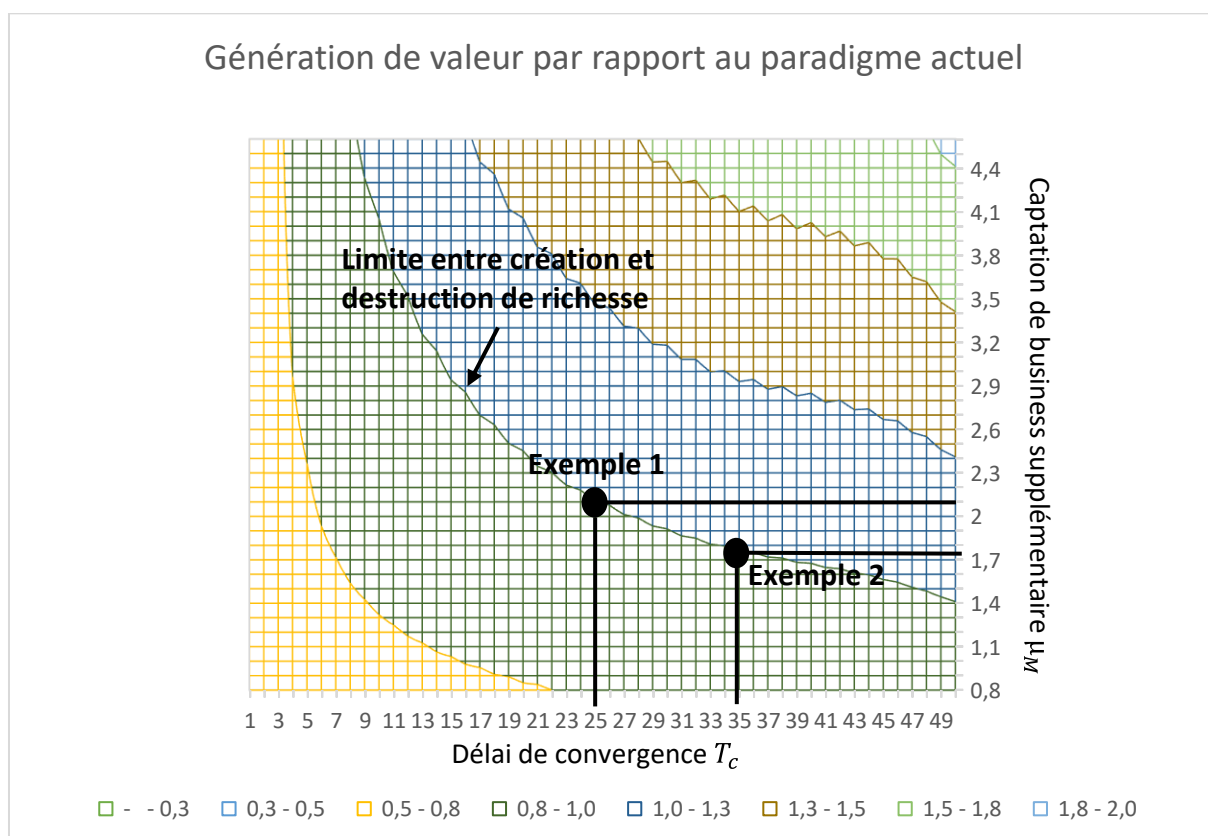
Figure 83 - Comparaisons des VAP selon différents scénarios

	Valeur actualisée des résultats futurs (€)	Valeur actualisée des résultats futurs (paradigme actuel = 1)
Paradigme actuel	VAP ₀ = 246 491 038	1
Simulation 1	VAP ₁ = 190 619 225	0,77
Simulation 2	VAP ₂ = 206 587 729	0,84
Simulation 3	VAP ₃ = 267 502 313	1,09

L'équilibre actuel est donc instable :

- à long-terme, tous les acteurs seraient perdant d'un mouvement vers le nouveau paradigme ($NBV_3^{opt} < NBV_1^{opt}$),
- mais aucun acteur n'a intérêt à faire le mouvement en retard, après les autres : $VAP_2 > VAP_1$,
- et par ailleurs, l'acteur qui ferait le mouvement en premier pourrait peut-être bénéficier d'une bulle qui compenserait sa perte à long-terme: $VAP_3 > VAP_1$.

Examinons de manière croisée les paramètres T_c et μ_M pour lesquels ce mouvement de l'acteur qui jouerait en premier, pourrait se révéler positif. Pour cela, nous calculons la valeur actuelle des résultats futurs VAP pour différentes valeurs de ces paramètres et la comparons à la valeur actuelle des résultats futurs dans le paradigme actuel VAP₀. Le graphique suivant présente le ratio VAP/VAP₀ selon un code couleur expliqué dans sa légende pour différentes valeurs de T_c (en abscisses) et μ_M (en ordonnées). Si ce ratio est supérieur à 1, alors le mouvement est générateur de richesse pour l'assureur qui l'enclenche. Si ce ratio est inférieur à 1, alors le mouvement est destructeur de valeur.



La limite entre destruction et création de valeur est symbolisée par la limite entre la zone bleue ($VAP/VAP_0 > 1$) et la zone verte ($VAP/VAP_0 < 1$). La lecture du graphique donne des indicateurs objectifs de décision pour un assureur d'entrer ou non dans le nouveau paradigme, selon les convictions qu'il a sur le business supplémentaire captable et le temps de convergence de la concurrence.

Par exemple :

- pour $T_c < 13$, le mouvement ne serait pas créateur de valeur pour l'assureur pour les valeurs de μ_M testées ($\mu_M < 4,4$),
- pour $T_c = 25$, le mouvement serait créateur de valeur pour l'assureur qui arriverait à multiplier son business par $\mu_M = 2$ (exemple 1 sur le graphique) ou plus au cours de la phase de transition,
- pour $T_c = 35$, le mouvement serait créateur de valeur pour l'assureur qui arriverait à multiplier son business par $\mu_M = 1,7$ (exemple 2 sur le graphique) ou plus au cours de la phase de transition.

3.2.3 Limites de l'analyse du nouveau paradigme

Nous avons étudié jusqu'à présent le sujet de la transférabilité des contrats d'assurance en faisant l'hypothèse du passage d'un équilibre existant, caractérisé par une réponse uniforme du marché à la question de la transférabilité (absence de possibilité de transfert), à un nouvel équilibre, caractérisé par une autre réponse uniforme du marché à cette question (possibilité de transfert incluse dans tous les contrats). La modélisation de la rentabilité de l'activité d'assurance emprunteur de notre entreprise d'une part, dans ce nouvel équilibre, et d'autre part, dans la phase transition, donne des éléments de prise de décision.

Au-delà des hypothèses, à parfaire, notre analyse nous semble faire face à deux limites :

- Premièrement, nous n'envisageons que la possibilité d'une réponse uniforme du marché. Dans la mesure où la transférabilité n'est qu'une faculté proposée au client mais que l'assureur ne dispose d'aucun moyen de l'imposer, une autre possibilité d'équilibre pourrait être un choix fait par certains acteurs de proposer la transférabilité comme une option facultative du contrat. Le client aurait alors le choix, lors de la souscription de la souscrire ou non. Dès lors, deux sujets complémentaires pourraient être traités :
 - o la question du caractère anti-sélectif de cette option se poserait : comme toute option laissée à la main du client, l'asymétrie de l'information fait courir ce risque à l'assureur (l'anti-sélection pourrait par exemple se manifester par une surreprésentation d'états de santé se dégradant davantage à long-terme).
 - o La question du pricing de l'option pourrait aussi être analysée, avec des paradoxes à résoudre : le prix de cette option se déduit de l'écart de prix entre le tarif optimisé du paradigme actuel (cf. paragraphe 3.1.2.2) et le tarif optimisé du nouveau paradigme (cf. paragraphe 3.1.2.3), ce qui conduit à un résultat négatif, c'est-à-dire que nous serions en apparence prêts à payer le client pour qu'il bénéficie de la faculté de transfert ? Quelle part de cet abandon de prime explique la perte de rentabilité estimée ?
- Deuxièmement, nous envisageons la transférabilité à iso-exigence de rentabilité des affaires nouvelles pour les assureurs. On voit que cette exigence de rentabilité inchangée conduit (i) une plus grande compétitivité des produits proposés, (ii) à une incertitude sur la rentabilité réelle, liée à l'approche du marché et à l'anti-sélection possible, (iii) à une moindre rentabilité de l'activité au global. Il semblerait donc rationnel pour les acteurs du marché de l'assurance emprunteur d'augmenter leur exigence de rentabilité en compensation de ces trois facteurs, ce qui aurait pour conséquence une augmentation généralisée des tarifs sur le marché... L'acteur qui ne proposerait pas cette option pourrait alors faire la différence en étant plus compétitif, ce qui nous ramènerait à l'équilibre initial.

3.3 Décision

Nous avons étudié le sujet de la transférabilité des contrats d'assurance en décomposant notre approche en deux étapes :

- Tout d'abord, nous avons résolu un problème d'optimisation consistant à maximiser les profits futurs générés par une cohorte de production sous contrainte de respect d'un niveau de rentabilité et de règles de tarification acceptable et déployables dans des réseaux de distribution. Pour résoudre ce problème, nous avons exploité les données les plus récentes dont nous disposons sur nos portefeuilles pour estimer les lois de résiliation et de sinistralité actuelles, avant de faire un certain nombre d'hypothèses sur l'impact de la transférabilité. La conclusion de ce problème est univoque : la hausse de NBV des AN captées dans le nouveau paradigme ne compense pas la raréfaction des AN sur le marché.
- Puis, nous avons cherché à projeter le compte de résultat de l'activité d'assurance emprunteur de l'assureur qui ferait le choix d'adopter la transférabilité. Cette projection est faite avec des hypothèses sur T_c (durée de convergence vers le niveau d'AN stable du paradigme proposé) et μ_M (business supplémentaire capté à court-terme).
 - Cette projection montre l'existence d'un gisement de valeur pour l'assureur qui arriverait à se trouver dans une configuration délimitée par une courbe de niveau $f(T_c, \mu_M) = 1$. Cela signifie que, pour un assureur qui serait capable de se forger une conviction sur :
 - le rythme d'adoption par la concurrence,
 - sa capacité de captation supplémentaire d'affaires nouvelles,une décision rationnelle est possible sur ce sujet.
 - Cette projection et la conclusion précédente montrent également que le paradigme actuel est en réalité un équilibre instable. Si un assureur A se forge la conviction qu'il existe potentiellement un assureur B capable de prendre la décision de proposer la transférabilité, alors l'assureur A a intérêt à le devancer, quand bien même cette décision ne serait pas créatrice de valeur pour lui, car il limiterait ainsi la perte de valeur causée par une décision trop tardive par rapport à un concurrent.

CONCLUSION

Dans ce mémoire, nous avons cherché à déterminer les critères objectifs permettant de juger l'opportunité de répondre positivement à une évolution souhaitée par les acteurs du marché de l'Emprunteur, qui réclament la suppression des formalités médicales dans le cadre du changement de bien assuré, c'est-à-dire la transférabilité du contrat d'assurance emprunteur.

La présentation du contexte et des éléments de marché a montré l'existence d'un marché capté par quelques gros acteurs et sur lequel les acteurs challengers ne peuvent se différencier que par la qualité de leur processus (cf. partie 1), ce qui justifie l'intérêt d'une telle étude.

Si le bénéfice pour le client d'une telle faculté est évident, le bénéfice pour l'assureur est plus ambigu, d'où la nécessité d'établir des outils permettant d'objectiver la décision d'offrir cette faculté.

Pour étayer notre étude, nous avons exploité les données les plus récentes d'un portefeuille d'assurance emprunteur pour établir une loi de décès et une loi de durée des contrats. Ces lois ont été établies à l'aide d'un modèle à risques proportionnels de Cox, d'un positionnement en fonction de lois de référence et en utilisant un arbre de décision (cf. partie 2).

Ces lois ont été exploitées pour déterminer une possible optimisation du tarif client dans la situation actuelle, puis une autre optimisation du tarif client dans la situation où la transférabilité serait offerte (cf. partie 3).

Enfin, la décision stratégique de proposer la transférabilité au client a été jugée en modélisant la rentabilité de l'activité d'assurance Emprunteur dans ce cadre et en la comparant à la rentabilité actuelle, en faisant différentes hypothèses quant à l'ampleur de l'avantage concurrentiel généré (mesurée par la variable μ_M , business supplémentaire capté à court-terme) et au rythme d'adaptation de la concurrence (mesuré par la variable T_c , durée de convergence vers le niveau d'AN stable du paradigme proposé) (cf. partie 3.2).

Cette projection montre l'existence d'un gisement de valeur pour l'assureur qui arriverait à se trouver dans une configuration délimitée par une courbe de niveau $f(T_c, \mu_M) = 1$. Concrètement, cela signifie que, pour un assureur qui serait capable de se forger une conviction sur les paramètres μ_M atteignables dans le cadre de son activité et T_c de diffusion de la mesure au sein de la concurrence, une décision rationnelle est possible sur ce sujet par observation de la valeur de f .

Nous avons également montré que le paradigme actuel est un équilibre instable. Si un assureur A se forge la conviction qu'il existe potentiellement un assureur B susceptible de prendre cette décision, alors l'assureur A a intérêt à le devancer, quand bien même cette décision ne serait pas créatrice de valeur pour lui, car il limiterait ainsi la perte de valeur causée par une décision trop tardive par rapport à un concurrent.

Notre mémoire ne permet donc pas de conclure de façon générale sur la question de la transférabilité mais fournit des outils d'aide à la décision pour un assureur qui s'interroge sur ce choix stratégique. L'analyse pourrait par ailleurs être complétée de sensibilités sur les hypothèses utilisées, de l'étude d'une augmentation du capital assuré dans le cadre de la transférabilité et d'une extension à la garantie incapacité-invalidité.

BIBLIOGRAPHIE

- ACPR. (2017). *Recommandation sur le libre choix de l'assurance emprunteur souscrite en couverture d'un crédit immobilier*.
- Actelior. (2017). *Etude comparée entre contrats collectifs et contrats individuels : tarifs et enjeux*.
- BAO. (2011). *La loi Lagarde et la liberté de choisir son assurance emprunteur*. Récupéré sur <http://baofrance.com/assurance-prest/publications-assurance-emprunteur/31-la-loi-lagarde-et-la-liberte-de-choisir-son-assurance-emprunteur>
- BAO. (2017). *Tarifs Bancaires : segmentés et tous margés*. Observatoire BAO de l'assurance emprunteur.
- Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. (1984). *Classification and Regression Trees*. CRC Press.
- Cantoni, E., & Hastie, T. (2000). Degrees of freedom tests for smoothing splines. *Biometrika*.
- Cerf, M. (2018). *Techniques d'optimisation*. Cours Master de Mathématique (M2) UPMC.
- Comité consultatif du secteur financier. (2015). *Avis du 13 janvier 2015*. Récupéré sur <https://www.moneyvox.fr/r/CCSF/CCSF-2015-assurance-emprunteur.pdf>
- Denoeux, T. (2016). *Splines and Generalized Additive Models*. Cours Master UTC.
- Groupe de travail Mortalité de l'Institut des Actuaire. (2016). Récupéré sur Méthodologies de construction de tables de mortalité pour Solvabilité II & Package ELT: <http://www.ressources-actuarielles.net/gtmortalite>
- Inspection générale des Finances. (2013). *Assurance Emprunteur*.
- Larocque, D., & Bou-Hamad, I. (2009). Discrete-Time Survival Trees. *Canadian Journal of Statistics*.
- Muet, P.-A. (2016). *Pour le droit au changement annuel de l'assurance emprunteur*. Récupéré sur <http://www.pierrealainmuet.fr/documents/pour-le-droit-au-changement-annuel-de-l-assurance-emprunteur.pdf>
- Pelissier, A. (2016). De la spécialité de l'assurance, la résistance des juges du fond s'impose. *Recueil Dalloz*, p. 1327.
- Planchet, F. (2019). *Modèles de durée*. Cours ISFA.
- Planchet, F., & Thérond, P. (2011). *Modélisation statistique des phénomènes de durée : Applications actuarielles*. Economica.
- Roussille, M. (Mai 2017). *Revue de droit bancaire et financier*, Alerte 27.

ANNEXES

Annexe 1 : Liste des critères d'équivalence en assurance Emprunteur

POUR LES GARANTIES DECES, PTIA, INVALIDITE ET INCAPACITE	
Couverture des sports amateurs pratiqués par l'emprunteur à la date de souscription	<input type="checkbox"/> Oui <input type="checkbox"/> Non
Maintien de la couverture en cas de déplacement dans le monde entier	<input type="checkbox"/> Oui <input type="checkbox"/> Non
A titre personnel	<input type="checkbox"/> Oui <input type="checkbox"/> Non
A titre professionnel ou humanitaire	<input type="checkbox"/> Oui <input type="checkbox"/> Non
GARANTIE DECES	
Couverture de la garantie décès pendant toute la durée du prêt ?	<input type="checkbox"/> Oui <input type="checkbox"/> Non
GARANTIE PTIA	
Couverture de la garantie PTIA pendant toute la durée du prêt ?	<input type="checkbox"/> Oui <input type="checkbox"/> Non
GARANTIE INCAPACITE	
Couverture de la garantie pendant toute la durée du prêt	<input type="checkbox"/> Oui <input type="checkbox"/> Non
Délai de franchise	<input type="checkbox"/> ≤ 30 jours <input type="checkbox"/> ≤ 60 jours <input type="checkbox"/> ≤ 90 jours <input type="checkbox"/> ≤ 120 jours <input type="checkbox"/> ≤ 180 jours
Pour une personne en activité, évaluation en fonction de la profession exercée au jour du sinistre	<input type="checkbox"/> Oui <input type="checkbox"/> Non
Pour une personne en activité, prestation égale à la mensualité assurée sans référence à la perte de revenu subie pendant le sinistre.	<input type="checkbox"/> Oui <input type="checkbox"/> Non
Maintien de la couverture en cas de temps partiel thérapeutique avec une prise en charge minimale de 50% sur une durée d'au moins 90 jours	<input type="checkbox"/> Oui <input type="checkbox"/> Non
Couverture des inactifs au moment du sinistre	<input type="checkbox"/> Oui <input type="checkbox"/> Non Si Oui Taux de prise en charge : <input type="checkbox"/> 1-49% <input type="checkbox"/> 50-99% <input type="checkbox"/> 100 %
Couverture des affections dorsales	<input type="checkbox"/> Sans condition d'hospitalisation, ni d'intervention chirurgicale <input type="checkbox"/> Avec conditions d'hospitalisation (<input type="checkbox"/> <10 jours <input type="checkbox"/> 10 jours et plus) ou d'intervention chirurgicale
Couverture des affections psychiatriques	<input type="checkbox"/> Sans condition d'hospitalisation <input type="checkbox"/> Avec conditions d'hospitalisation (<input type="checkbox"/> <10 jours <input type="checkbox"/> 10 jours et plus)

GARANTIE PERTE D'EMPLOI	
Couverture de la garantie pendant toute la durée du prêt sans limite d'âge	<input type="checkbox"/> Oui <input type="checkbox"/> Non
Délai de carence pour l'application de la couverture	<input type="checkbox"/> ≤ 3 mois ; <input type="checkbox"/> ≤ 6 mois ; <input type="checkbox"/> ≤ 12 mois ;
Délai de franchise	<input type="checkbox"/> ≤ 60 jours ; <input type="checkbox"/> ≤ 90 jours ; <input type="checkbox"/> ≤ 120 jours ;
Durée d'indemnisation par sinistre	<input type="checkbox"/> ≥ 12 mois ; <input type="checkbox"/> ≥ 24 mois ;
Durée d'indemnisation totale d'au moins 36 mois	<input type="checkbox"/> Oui <input type="checkbox"/> Non
Part de l'échéance prise en charge	<input type="checkbox"/> ≤ 50% <input type="checkbox"/> ≤ 75% <input type="checkbox"/> < 100% <input type="checkbox"/> 100%
Prestation égale à la prise en charge de la mensualité, sans référence à la perte de revenu subie au moment du sinistre	<input type="checkbox"/> Oui <input type="checkbox"/> Non
Prise en charge du sinistre sans condition d'ancienneté en CDI	<input type="checkbox"/> Oui <input type="checkbox"/> Non
GARANTIE INVALIDITE	
Couverture de la garantie pendant toute la durée du prêt	<input type="checkbox"/> Oui <input type="checkbox"/> Non
Evaluation en fonction de la profession exercée au jour du sinistre	<input type="checkbox"/> Oui <input type="checkbox"/> Non
Prise en charge de l'invalidité totale, sans référence à la perte de revenu subie au moment du sinistre	<input type="checkbox"/> Oui <input type="checkbox"/> Non
Prise en charge de l'invalidité partielle (IPP) à partir de 33%	<input type="checkbox"/> Oui <input type="checkbox"/> Non
Couverture des affections dorsales	<input type="checkbox"/> Sans condition d'hospitalisation, ni d'intervention chirurgicale <input type="checkbox"/> Avec conditions d'hospitalisation (<input type="checkbox"/> <10 jours <input type="checkbox"/> 10 jours et plus) ou d'intervention chirurgicale
Couverture des affections psychiatriques	<input type="checkbox"/> Sans condition d'hospitalisation <input type="checkbox"/> Avec conditions d'hospitalisation (<input type="checkbox"/> <10 jours <input type="checkbox"/> 10 jours et plus)

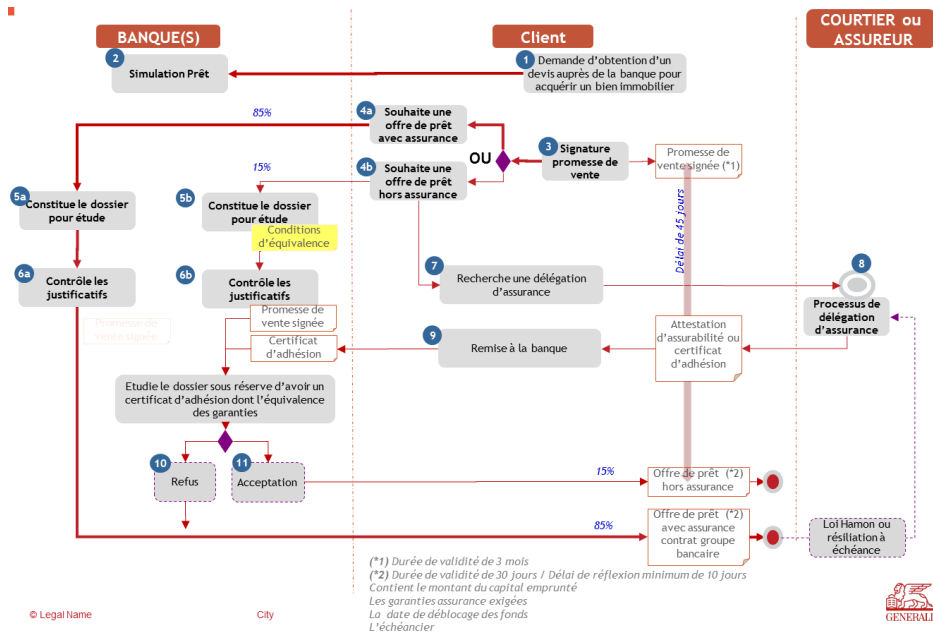
Annexe 4 : Comparatif des formalités médicales pour les emprunteurs âgés de moins de 46 ans (source : askapi.fr)

	Déclaration d'Etat de Santé	Questionnaire de Santé	Examens médicaux
APRIL Liberté	<=400 K€	400K€ à 500K€	>500K€
METLIFE	<=400 K€	400K€ à 1000K€	>1000K€
APRIL ADP	<=300 K€		>300K€
AFI ESCA	<=350 K€	350K€ à 500K€	>500K€
BPSIS Naoassur	<=300 K€	300K€ à 400K€	>500K€

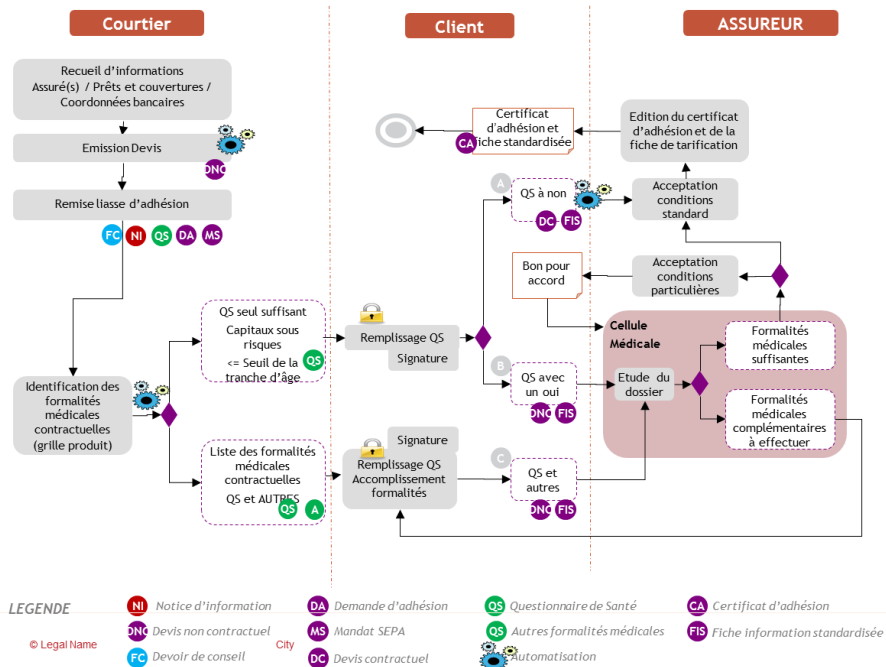
Annexe 5 : Grille des formalités médicales de l'assureur Axeria Prevoyance

	45 ans ou moins	De 46 ans à 54 ans	De 55 ans à 64 ans	De 65 ans à 80 ans
De 18 000 € ⁽¹⁾ à 100 000 €	Déclaration d'Etat de Santé	Questionnaire de Santé	Questionnaire de Santé	
De 100 001 € à 250 000 €			Questionnaire de Santé + Rapport Médical + Profil sanguin 2 + à partir de 300001 € : Dosage de la cotinine urinaire pour les non fumeurs	
De 250 001 € à 300 000 €		Questionnaire de Santé + Rapport Médical + Profil sanguin 2 + à partir de 300001 € : Dosage de la cotinine urinaire pour les non fumeurs	Questionnaire de Santé + Rapport Médical + Profil sanguin 2 + à partir de 300001 € : Dosage de la cotinine urinaire pour les non fumeurs	
De 300 001 € à 500 000 €	Questionnaire de Santé + Rapport Médical + Profil sanguin 2 + à partir de 300001 € : Dosage de la cotinine urinaire pour les non fumeurs			
De 500 001 € à 1 000 000 €	Questionnaire de Santé + Rapport Médical + Profil sanguin 1 + Dosage de la cotinine urinaire pour les non fumeurs			
De 1 000 001 € à 1 600 000 €	Questionnaire de Santé + Rapport Médical + Profil sanguin 4 + Chimie des urines + Dosage de la cotinine urinaire pour les non fumeurs	Questionnaire de Santé + Rapport Médical + Profil sanguin 3 + Chimie des urines + Rapport Cardiovasculaire avec tracé de l'ECG + Dosage de la cotinine urinaire pour les non fumeurs		
De 1 600 001 € à 2 200 000 €	Questionnaire de Santé + Rapport Médical + Profil sanguin 4 + Chimie des urines + Rapport Cardiovasculaire avec tracé de l'ECG + Dosage de la cotinine urinaire pour les non fumeurs + Questionnaire Financier	Questionnaire de Santé + Rapport Médical + Profil sanguin 3 + Chimie des urines + Rapport Cardiovasculaire avec tracé de l'ECG + Dosage de la cotinine urinaire pour les non fumeurs + Questionnaire Financier		Nous consulter
De 2 200 001 € à 4 600 000 € ⁽²⁾	Questionnaire de Santé + Rapport Médical + Profil sanguin 4 + Chimie des urines + Rapport Cardiovasculaire avec tracé de l'ECG + Dosage de la cotinine urinaire pour les non fumeurs + Formalités financières	Questionnaire de Santé + Rapport Médical + Profil sanguin 3 + Chimie des urines + Rapport Cardiovasculaire avec tracé de l'ECG + Dosage de la cotinine urinaire pour les non fumeurs + Formalités financières		Nous consulter

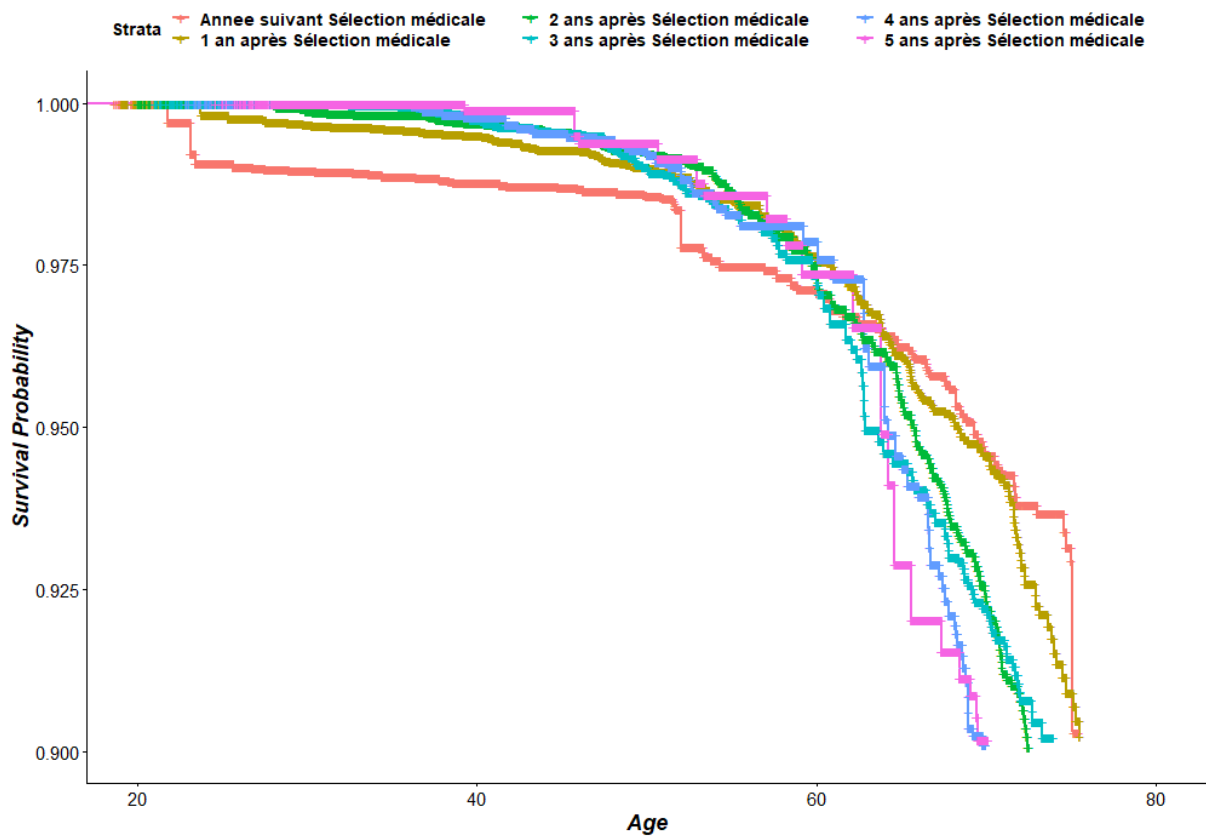
Annexe 6 : Parcours bancaire d'un client dans le cadre de l'octroi d'un prêt



Annexe 7 : Parcours assuranciel d'un client dans le cadre de l'octroi d'un prêt

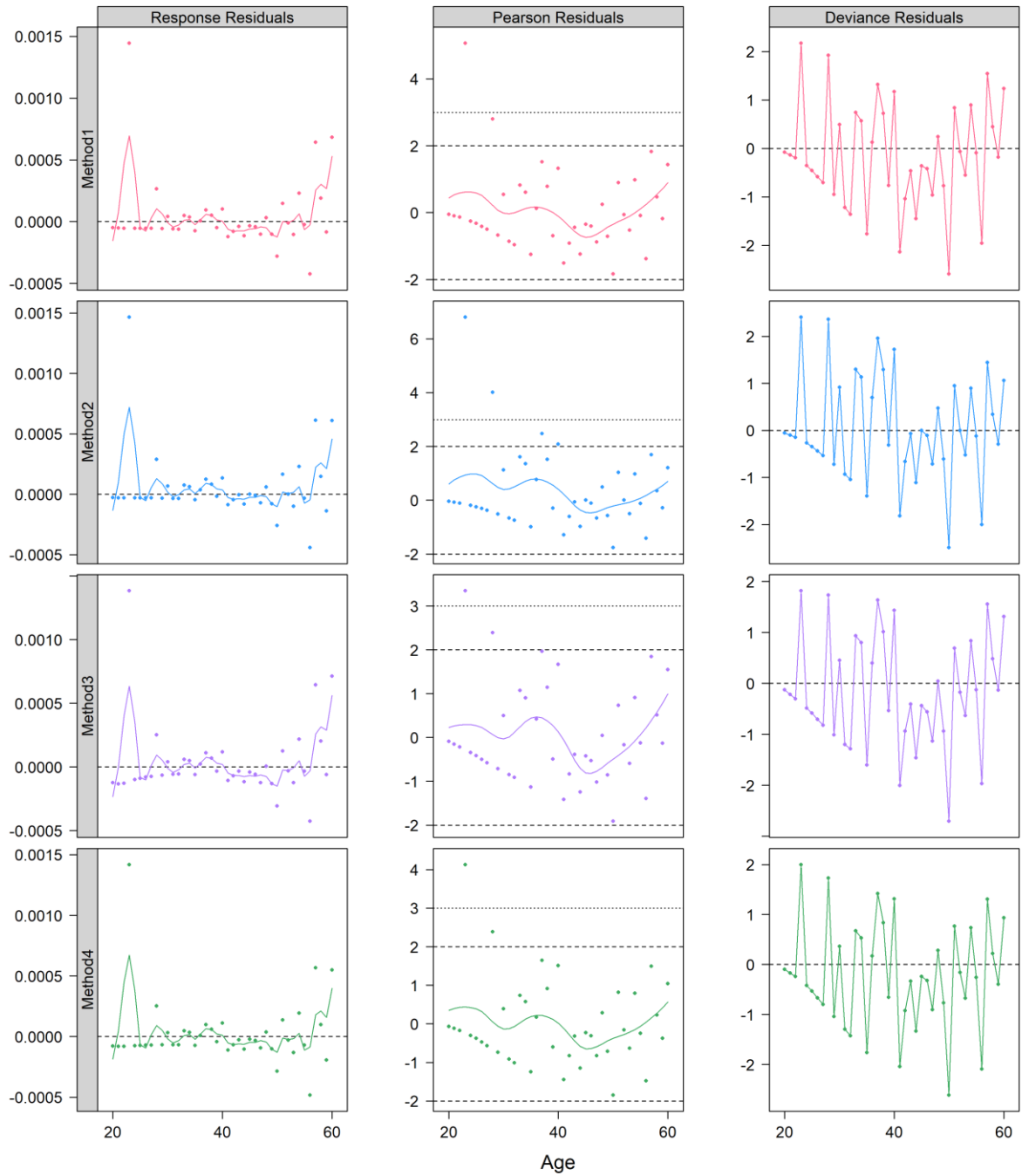


Annexe 8 : Observation de l'effet de la sélection médicale via la comparaison des lois de survie constatées sur le portefeuille par une estimation de Kaplan-Meier pour les clients selon l'antériorité de la sélection médicale



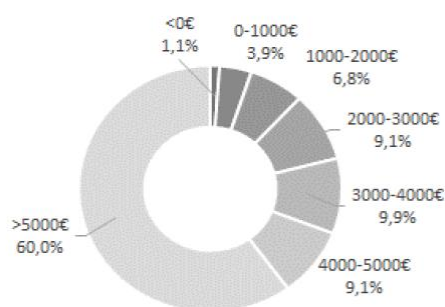
Annexe 9 : Représentation des résidus pour les quatre méthodes de modélisation pour les décès observés en 2018

Male - Plot of the residuals - year: 2018



Annexe 10 : étude du cabinet A. sur le potentiel du marché dans le cadre de la mise en place de l'amendement Bouquin

Répartition du stock d'assuré emprunteur selon l'écart entre leur tarif actuel et le tarif moyen de marché



Potentiel de marché par tranche d'écart moyen

Écart tarifaire	% adhésion (hypothèse)	Nombre d'assurés (en milliers)	tarif groupe	Tarif individuel*	potentiel individuel
<0M€	0%	2	65 M€	69 M€	0 M€
0-1000M€	0%	2 390	225 M€	128 M€	0 M€
1000-2000M€	0%	2 043	387 M€	193 M€	0 M€
2000-3000M€	10%	1 807	520 M€	265 M€	27 M€
3000-4000M€	15%	1 316	563 M€	282 M€	42 M€
4000-5000M€	20%	1 044	518 M€	251 M€	50 M€
>5000M€	30%	3 398	3 423 M€	1 516 M€	455 M€
Total	13,4%	12 000	5 700 M€	2 704 M€	574 M€

Sensibilité à l'hypothèse de captation de business selon écart tarifaire

Écart tarifaire	% adhésion hypothèse basse	potentiel individuel hyp. basse	% adhésion hypothèse basse	potentiel individuel hyp. basse
<0M€	0%	0 M€	0%	0 M€
0-1000M€	0%	0 M€	0%	0 M€
1000-2000M€	0%	0 M€	0%	0 M€
2000-3000M€	3%	9 M€	15%	40 M€
3000-4000M€	5%	14 M€	23%	63 M€
4000-5000M€	7%	17 M€	30%	75 M€
>5000M€	10%	152 M€	45%	682 M€
Total	4,5%	191 M€	20.1%	861 M€

Annexe 11 : Cohorte d'assurés utilisée pour le problème d'optimisation

	Tranche d'âge	Age moyen	Sexe	Fumeur	Capital moyen	Somme capitaux
1	18-32 ans	30	Female	Non	123 913 €	94 545 251 €
2	18-32 ans	30	Female	Oui	103 807 €	9 446 408 €
3	18-32 ans	30	Male	Non	121 456 €	61 699 495 €
4	18-32 ans	30	Male	Oui	111 564 €	13 833 933 €
5	33-36 ans	35	Female	Non	153 013 €	141 537 100 €
6	33-36 ans	34	Female	Oui	145 450 €	16 581 308 €
7	33-36 ans	35	Male	Non	164 644 €	139 453 419 €
8	33-36 ans	35	Male	Oui	170 973 €	33 852 632 €
9	37-41 ans	39	Female	Non	144 217 €	124 459 276 €
10	37-41 ans	39	Female	Oui	143 535 €	12 487 580 €
11	37-41 ans	39	Male	Non	161 804 €	126 854 260 €
12	37-41 ans	39	Male	Oui	160 301 €	23 083 318 €
13	42-46 ans	45	Female	Non	135 865 €	96 871 901 €
14	42-46 ans	45	Female	Oui	141 235 €	11 298 805 €
15	42-46 ans	45	Male	Non	163 007 €	114 430 745 €
16	42-46 ans	45	Male	Oui	141 470 €	12 873 749 €
17	47 ans et +	51	Female	Non	144 411 €	254 886 088 €
18	47 ans et +	50	Female	Oui	141 463 €	17 541 396 €
19	47 ans et +	51	Male	Non	199 426 €	517 309 780 €
20	47 ans et +	49	Male	Oui	188 433 €	36 179 157 €