

Mémoire présenté le : 13 janvier 2021

pour l'obtention du Master 2 Droit, Economie, Gestion, mention Actuariat
et l'admission à l'Institut des Actuares

Par : Sylvain ROZES

Titre : Construction d'une loi de rachat dynamique

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

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

Présidente du jury

Signature

Entreprise :

Sandrine LEMERY

Nom : La France Mutualiste

*Membres présents du jury de
l'Institut des Actuares*

Directrice de mémoire en entreprise :

Nom : Nathalie LE BRAZIDEC

Signature :

Edith BOCQUAIRE

Invitées :

Nom : Sara FERT

Arnaud COHEN

Signature :

Florence PICARD

Nom : Odile DELMAS

*Membres présents du jury du
CNAM*

Signature :

Olivier DESMETTRE

***Autorisation de publication et
de mise en ligne sur un site de
diffusion de documents actua-
riels (après expiration de l'éventuel
délai de confidentialité)***

David FAURE

Signature du responsable entreprise

François WEISS

Secrétariat :

Signature du candidat

Bibliothèque :

« Nous ne sommes pas des machines pensantes qui ressentent des émotions, nous sommes des machines émotionnelles qui pensent. »

- Docteur Antonio Damasio, neurologue à l'Université de Californie du Sud

En premier lieu, je tiens à remercier Nathalie Le Brazidec d'avoir accepté de diriger mon mémoire d'actuariat. Ses remarques et son expérience ont été essentielles dans son orientation et sa rédaction.

Je remercie également ma manager, Odile Delmas, de m'avoir encouragé et surtout poussé à certains moments de sa composition.

Mes remerciements vont par ailleurs à Christophe Maborough qui a donné de son temps pour les simulations Solvabilité II. Merci aussi de ton aide précieuse sur l'interprétation des impacts de la loi dynamique sur le taux de couverture.

J'adresse mes remerciements à Arnaud Cohen pour sa disponibilité en ce qui concerne les échanges à propos de la méthodologie et des points d'attention.

Je tiens aussi à remercier David Faure d'avoir orienter l'écriture de ce mémoire.

Pour finir, je tiens à dire un grand merci à mes amis, ma famille et particulièrement à ma femme qui m'ont soutenu et porté tout au long de cette longue aventure.

Résumé

La directive Solvabilité II préconise l'emploi d'une loi de rachat dynamique dans l'estimation des provisions techniques et des fonds propres. Les calculs prudentiels tentent de rapprocher les fluctuations de l'environnement économique avec la décision, prise par les détenteurs d'une assurance-vie, de récupérer leur épargne acquise. Cependant la mutuelle La France Mutualiste rencontre certaines difficultés dans son application car ses taux de rachats ne peuvent pas être tant atténués, du fait de l'application de la formule standard des rachats dynamiques, dans un contexte de taux d'intérêt obligataire très faible. La raison repose notamment sur le fait que son portefeuille se caractérise par la prédominance d'anciens combattants, dont une grande partie est à un âge avancé, ce qui semble l'éloigner du paramétrage des trois fonctions proposées par l'autorité de contrôle. C'est pourquoi l'objectif de l'étude est de construire, à l'aide de méthodes comportementales, une loi plus appropriée à destination du modèle de projection des flux de trésorerie. La persistance d'un environnement de taux bas ne permet pas de surcroît d'identifier les conséquences d'une remontée des taux obligataires et de ce fait d'anticiper la manifestation d'un rachat massif. Néanmoins, le même modèle de choix discrets pourra être appliqué à la fois sur les observations tirées du portefeuille et par la suite à partir de données simulées, dans le but de définir chacun des paramètres de la fonction de rachat dynamique. Pour finir, il est nécessaire d'examiner les incidences sur l'indicateur prudentiel du taux de couverture, qui requiert désormais une attention toute particulière.

Abstract

Lapse rate modelling is an important topic under Solvency II Directive for life insurers. For regulatory purposes, companies establish the amount of equity required to hold in order to meet its insurance obligations. The economic rationale to surrender a contract in anticipation is difficult to model. The reason of the termination can be influenced by a large number of parameters, in particular financial market and macroeconomic environment. La France Mutualiste encounters some problems in the application of the standard formula due to a specific contracts portfolio, composed largely of advanced age ex-servicemen. Since policyholders tend not to exercise their contractual option, the problem is to bring out the Interest Rate Hypothesis. This degree project aims to resolve this issue by calibrating the interest rate dependency of lapse rates with statistical techniques based on Conditional Logit Discrete-Choice Model. The objective is to explain the surrender choice from a set of two discrete alternatives where people choose the one of maximal benefit or utility. The lack of historical data in times of interest rates hikes leads to the solution of several scenario simulations using the same statistical method. Finally, it is important to perform the economical impact assessments on the Solvency Capital Requirement coverage ratio in putting the new dynamic lapse rates formula into practice.

Sommaire

Résumé	i
Table des figures	vii
Liste des tableaux	ix
Introduction	1
I La France Mutualiste et le risque de rachat	3
II Revue méthodologique et analyse exploratoire	17
III Hypothèse du taux d'intérêt	31
IV Loi conjoncturelle : partie droite	41
V Loi conjoncturelle : partie gauche	51
VI Loi structurelle et impacts Solvabilité II	61
Conclusion	71
Bibliographie	73
Annexes	77
A Généralités	77
B Modèle logistique	85
C Modèle à choix discret	91

D	Évaluation du modèle à choix discrets	109
E	Indice de Youden	123
F	Méthodes de lissage	127
G	Modèle mixte	131
H	Recherche d'un minimum	135

Table des figures

I.1	Décomposition des provisions mathématiques 2015 et 2016 par tranches d'âges	7
I.2	Décomposition des provisions mathématiques 2015 et 2016 par type de risque	8
I.3	Loi de rachats conjoncturels ACPR	15
II.1	Taux de rachat total du fonds Euro	22
II.2	Ancienneté moyenne de l'encours et des rachats	23
II.3	Taux de rachat en montant par tranches d'ancienneté	24
II.4	Age moyen de l'encours et des rachat	25
II.5	Taux de rachat en montant par tranches d'âges	25
II.6	Montant moyen de l'encours et des rachat	26
II.7	Montant des provisions mathématiques selon la présence d'un taux minimum garanti au contrat	27
II.8	Valeurs mensuelles du TME et Taux servi moyen $N - 1$	29
II.9	Valeurs de <i>delta taux</i> et taux de rachat total du fonds Euro	30
III.1	Taux de rachat mensuel en fonction de <i>delta taux</i>	32
IV.1	Taux de rachat par <i>delta taux</i>	43
IV.2	Lissage des taux de rachat	46
IV.3	Lissage des taux de rachat en fonction de <i>delta taux</i> par le modèle mixte	47
IV.4	Détermination du seuil γ à partir duquel les rachats conjoncturels sont stables	49
IV.5	Taux annualisés des rachats conjoncturels observés et formule ACPR en fonction de <i>delta taux</i>	50
V.1	Simulation d'un rachat du contrat initial en t	53
V.2	Simulation d'un rachat du contrat initial en t avec les notions mathématiques	55
V.3	Temps de retour sur investissement d'une valeur de 100€ en fonction de <i>delta taux</i>	57
VI.1	Taux de rachat total Euro observé et taux de rachat structurel théorique	62

VI.2	Taux de rachat total Euro observé contre taux de rachat conjoncturel . . .	63
VI.3	Taux de rachat global observé de 2010 à 2017	64
VI.4	Évolution de la différence $R - TA$ selon la valeur du quantile du <i>Best Estimate</i> au 31 décembre 2017	69
VI.6	Analyse des changements sur le taux de couverture entre 2016 et 2017 . . .	70
A.1	Nombre de mois accumulés par <i>delta taux</i>	80
A.2	Mois moyen par <i>delta taux</i>	80
C.1	La transformation logit	99
D.1	La courbe ROC	113
D.2	Exemple : Distribution des mesures de diagnostic de test pour une population en bonne santé et une population malade	115
D.3	Courbe ROC du modèle conjoncturel (variable <i>delta taux</i>)	121

Liste des tableaux

I.1	Fonctions de rachats conjoncturels proposées par l'ACPR	15
II.1	Délai en jours entre la date d'inventaire et le règlement du rachat par année de survenance	21
III.1	Tableau des coefficients de corrélation de Spearman entre <i>delta taux</i> et le taux de rachat	33
IV.1	Récapitulatif sur la sélection séquentielle dans le modèle à choix discret pondéré	44
V.1	Seuils optimums selon les divers scénarios en fonction de la prévalence de l'événement de rachat	59
VI.1	Coefficients de corrélation avec le taux de rachat observé	63
VI.2	Loi dynamique ACPR et loi dynamique LFM	67
VI.3	Impacts de la nouvelle loi dynamique sur le taux de couverture au 31 décembre 2017	67
VI.4	Loi dynamique ACPR et loi dynamique LFM	72
A.1	Exemple d'un contrat racheté	79
D.1	Matrice de confusion - Forme générique	110
D.2	Estimations des coefficients de régression par l'analyse du maximum de vraisemblance du modèle logistique pondéré	119
D.3	Valeurs de l'aire sous la courbe ROC du modèle logistique complet pondéré et marginaux	120
E.1	Estimations par l'analyse du maximum de vraisemblance des paramètres du modèle conjoncturel pondéré	124

Introduction

« Les scénarios de remontée de taux n'ont pas été traités de manière assez cohérente. La modélisation du risque de rachat conjoncturel est un exercice que nous allons devoir retravailler avec les assureurs », a déclaré Sandrine Lemery, première secrétaire générale adjointe de l'Autorité de Contrôle Prudentiel et de Résolution (ACPR), dans une interview à Reuters¹ en 2016.

Les nouvelles exigences européennes en matière de solvabilité ont permis au secteur de l'assurance de mieux maîtriser les aléas associés à leurs engagements. Néanmoins, le régulateur français estime que les compagnies sont tenues de mieux appréhender le risque d'une remontée brutale des taux d'intérêt.

Comment trouver un modèle assez fidèle dont le but est d'identifier dans quelle proportion les assurés, éventuellement peu avertis, pourraient réagir aux évolutions des marchés financiers ?

Le risque de rachat est ainsi une sortie anticipée que l'assureur doit savoir maîtriser afin de garantir sa solvabilité. Dans le cadre du régime prudentiel Solvabilité II, la Mutuelle utilise une loi de rachat dynamique proposée par l'ACPR. Cependant l'observation d'un taux de rachat faible entraîne de nombreux ajustements dans son emploi.

1. <http://marchésfinanciers.com/secteurs-natixis-france-les-assureurs-doivent-mieux-apprehender-le-risque-de-taux/>

Le but recherché est par conséquent d'estimer les paramètres de manière plus appropriée en conservant la forme par paliers de la fonction suggérée.

Ils existent par ailleurs de nombreux travaux dont le but est d'encadrer et d'établir une loi de rachat. Leurs motivations sont de connaître les raisons de récupérer l'épargne acquise d'un contrat d'assurance-vie. Ils présentent d'un côté, les rachats structurels, reposant sur les caractéristiques de l'adhérent ou du contrat. Les rachats conjoncturels, de l'autre côté, sont fondés sur la situation et les perspectives de l'environnement économique.

La difficulté majeure rencontrée dans ce mémoire sera la détermination, sur les données observées entre 2010 et 2016 caractérisées par un contexte de taux bas, d'un indicateur financier permettant de reconnaître une décision de rachat rationnelle.

La France Mutualiste et le risque de rachat

Ce chapitre présentera tout d'abord la Mutuelle et son portefeuille si particulier du fait de l'histoire de sa création.

L'assurance-vie, produit préféré des français, est contrôlée par la directive Solvabilité II. Cette dernière impose la modélisation du comportement de rachat dans un contexte de perturbations de l'environnement économique.

Ce risque est introduit comme un enjeu crucial dans les calculs prudentiels. Son évaluation reposant sur l'appréciation de l'horizon temporel de l'utilité de la détention d'une somme d'argent s'avère d'autant plus délicate...

I.1 La France Mutualiste

Mutuelle nationale d'épargne et de retraite, La France Mutualiste¹ (LFM) propose des solutions pour la préparation à la retraite, la valorisation de l'épargne et la transmission du patrimoine.

1. <https://www.la-france-mutualiste.fr/>

I.1.1 Présentation de La France Mutualiste

Son activité consiste à assurer des produits de rentes viagères immédiates et différées, ainsi que des produits d'épargne traditionnels, en euros ou en Unités de Compte² (UC). La Mutuelle s'est construite autour du monde des anciens combattants et de leur famille. La France Mutualiste est régie par le Livre II du Code de la Mutualité³.

La Mutuelle réalise des opérations :

En cas de vie : il s'agit d'opérations d'assurance prévoyant le versement de capitaux différés, de rentes viagères immédiates ou différées ou comportant des annuités certaines si l'adhérent est en vie au terme du contrat.

En cas de décès : il s'agit d'opérations d'assurance prévoyant le versement de capitaux aux bénéficiaires désignés, de rentes viagères immédiates ou différées au conjoint survivant si l'adhérent décède avant le terme du contrat (rentes de réversion).

I.1.2 Histoire

Le 17 mai 1891, le dentiste Léon Guillot réunit quelques-uns de ses patients dans un café de la rue Montorgueil et leur propose de verser une cotisation mensuelle dans le but de percevoir une retraite, le moment venu. Quelques mois plus tard, c'est la naissance de cette « Société philanthropique et humanitaire de retraite » baptisée « La Boule de Neige ».

La promulgation de la loi du 4 août 1923 (voir annexe A.1, page 77) institue un droit à réparation aux anciens combattants sous la forme d'une retraite spécifique : la Retraite Mutualiste du Combattant. Elle se fonde sur la notion fondamentale qui est le droit à réparation reconnu par la loi, en considération des dommages directs ou collatéraux

2. Les supports en unités de compte désignent les supports d'investissement autres que les fonds en euros et qui composent un contrat d'assurance-vie multisupports. Ces supports présentent des profils et des rendements variés : fonds actions, d'obligations, fonds diversifiés, parts ou actions de sociétés immobilières, *etc.*

3. Le code de la mutualité est un recueil des textes juridiques français qui déterminent le statut, la composition et le fonctionnement des mutuelles. Le livre II est relatif aux mutuelles pratiquant des opérations d'assurance, de réassurance et de capitalisation.

infligés par la guerre aux Combattants, à leurs familles, aux victimes civiles de guerre ou hors guerre.

La France Mutualiste est fondée en 1925 en tant que caisse autonome d'anciens combattants au sein de la Boule de Neige dans le but de diffuser ce produit. L'importance de la RMC est rapportée en annexe A.6, page 81.

Au 31 décembre 2017, la Mutuelle compte 227 222 adhérents pour un nombre de 259 484 contrats gérés. Le montant de ces actifs en valeurs de marché représente 8,2 milliards d'euros et celui des cotisations nettes de réassurance s'élève à 375,4 millions d'euros pour l'année 2017.

I.1.3 Produits commercialisés

Le contrat historique : la Retraite Mutualiste du Combattant

Le produit RMC est un contrat à adhésion individuelle dont le but est de se constituer une rente avec l'aide de l'État dans la limite d'un plafond. Ce dernier est fixé par référence à un indice des pensions militaires d'invalidité⁴. Une instruction fiscale paraît au Bulletin officiel des finances publiques chaque année dans le but de fixer ce plafond de déduction fiscale. La rente peut être viagère différée ou immédiate et associée en option à une contre-assurance viagère des versements au moment du décès. Les versements sont donc souscrits soit en régime de capitaux réservés soit en régime de capitaux aliénés. La constitution de la rente est réalisée sur une période :

- de minimum 10 ans et au moins jusqu'au 50^e anniversaire,
- selon la date de naissance, la durée peut être réduite d'une année à partir de 51 ans, sans pouvoir être inférieure à 4 ans,
- la rente peut être immédiate pour les adhérents qui bénéficient déjà d'une rente de même nature en cours de service dans un autre organisme.

4. Par exemple, l'arrêté du 1^{er} août 2017 (paru au Journal Officiel du 12 août 2017) porte la valeur du point d'indice de pension militaire d'invalidité à 14,40€ à compter du 1^{er} janvier 2017, ce qui a pour effet de porter le plafond de rente majorée par l'État de la Retraite Mutualiste du Combattant à 1 800€.

La provision mathématique⁵, est rachetable avant la date de liquidation initialement prévue au contrat d'après l'article L. 223-22 du Code de la Mutualité. Les souscripteurs sont les titulaires d'une Carte du Combattant, du Titre de Reconnaissance de la Nation ainsi que les veufs, orphelins et ascendants de militaires morts pour la France par leur participation à un conflit. La rente liquidée est versée à vie selon une périodicité trimestrielle.

Les contrats d'épargne

LFM lance en 1992 une gamme de contrats d'assurance-vie, complémentaires à la RMC, aussi proposés aux familles des anciens combattants. Des produits multi supports ont été lancés par la suite en 2002, 2008 et 2010.

À l'opposé du contrat mono support, qui ne propose qu'un seul support d'investissement (le fonds en euros), le contrat d'assurance-vie multi supports dispose d'au moins un fonds en euros et de plusieurs supports exprimés en Unité de Comptes (OPCVM, actions, obligations, parts de sociétés immobilières, *etc.*). Le souscripteur d'un contrat d'assurance-vie multi supports peut alors se constituer une épargne diversifiée en sélectionnant selon son aversion au risque⁶ les supports d'investissements qui lui sont proposés. Selon l'article L. 223-22 du Code de la Mutualité, ces contrats sont tous rachetables.

La Mutuelle intègre en 2005 le portefeuille de contrats R4 provenant de l'Union Mutualiste Retraite⁷ composé de 30 000 adhérents issus du monde de l'enseignement. Il est constitué de deux produits : un produit d'épargne mono support avec la possibilité d'une liquidation en rentes et un produit construit d'une façon similaire au produit RMC (rentes différées).

5. La valeur actualisée des engagements de chacun (assureur et assuré) pondérés par la probabilité d'occurrence.

6. L'aversion au risque est un concept dérivé de la psychologie, de l'économie et de la finance. La valeur que représente un investissement comportant des risques peut être amoindrie aux yeux d'un investisseur par la présence même de risques, qui pourraient être synonymes de pertes.

7. L'Union Mutualiste Retraite a été créée en 2002 afin d'offrir une solution retraite accessible à tous les mutualistes. C'est une union de mutuelles, unions et fédérations qui ont souscrit un contrat collectif à adhésion facultative afin de permettre à leurs adhérents de bénéficier du complément retraite Corem.

I.1.4 Caractéristiques du portefeuille LFM

Le portefeuille de la Mutuelle est essentiellement composé de membres des forces armées « régulières » ayant réalisé une opération extérieure⁸ (OPEX) et d'anciens combattants entrés par le biais de la RMC.

La décision de la France, en 1996, d'abandonner le système d'armée mixte en renonçant à la conscription s'est traduite par une réduction des effectifs au sein de la Défense. Entre 2002 et 2015, le volume des effectifs projetés en OPEX est donc passé de plus de 14 000 personnels à environ 7 000/8 000. Il semble alors plus difficile pour la Mutuelle de voir entrer suffisamment d'affaires nouvelles dans le but de rajeunir ses adhérents. Elle a donc ouvert à tous son portefeuille de contrats d'épargne depuis quelques années.

Le graphique I.1, de la présente page, montre que 75% des provisions mathématiques de la Mutuelle sur les deux dernières années observées se répartissent sur les âges compris entre 70 ans et 90 ans.

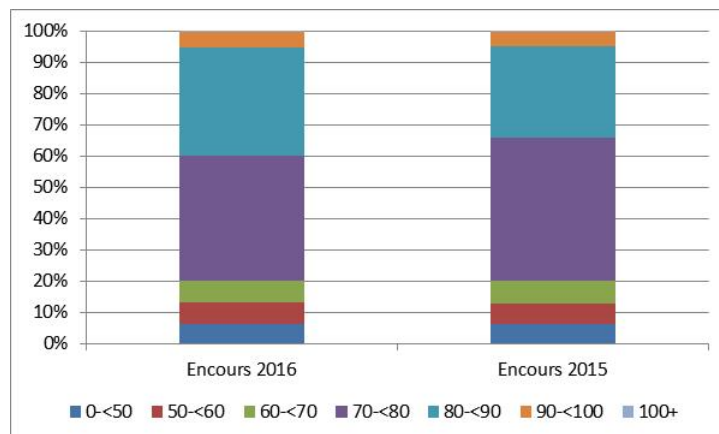


FIGURE I.1 – Décomposition des provisions mathématiques 2015 et 2016 par tranches d'âges

Le graphique I.2, page suivante, met en évidence le montant des provisions mathématiques rachetables de la Mutuelle. Le risque de rachat représente environ 40% en 2015 et

8. D'après la définition traditionnelle donnée par le Ministère des armées, les opérations extérieures sont les « interventions des forces militaires françaises en dehors du territoire national ». La qualification d'OPEX résulte d'un arrêté du Ministre des armées, qui porte ouverture du théâtre d'engagement en précisant la zone géographique et la période concernées. Les OPEX se distinguent des forces prépositionnées dans des bases en Afrique en vertu d'accords de défense ou en mer.

2016. Le montant des capitaux décès non rachetable est issu de la contre assurance des versements sur le produit RMC.

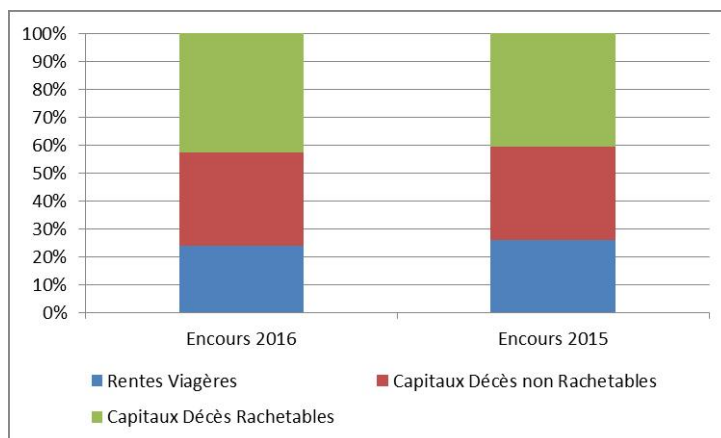


FIGURE I.2 – Décomposition des provisions mathématiques 2015 et 2016 par type de risque

I.2 Les produits d'épargne en France

D'après une étude de l'INSEE (2011), les détenteurs d'un contrat d'assurance-vie, âgés de plus de 70 ans, sont entre 3 et 4 sur 10 à le considérer comme un outil de gestion de patrimoine et de transmission successorale.

I.2.1 La sécurité du support garanti

L'étude sur le patrimoine des français réalisée par l'INSEE (2015) note que l'assurance-vie compte parmi les produits privilégiés : 36,5% des ménages métropolitains sont détenteurs d'au moins un produit. La raison de cet engouement réside principalement dans le cadre fiscal attractif et sur la possibilité de sécuriser tout ou partie des versements réalisés.

Les auteurs montrent également que les assurances-vie sont largement investies en support garanti ou à effet cliquet⁹, c'est-à-dire que les contrats sont pour deux tiers

9. L'effet de cliquet dans un contrat d'assurance vie en euros est le mécanisme qui garantit à l'épargnant les gains de l'année écoulée. Les intérêts générés par son contrat lui sont définitivement acquis : il

d'entre eux des mono supports et que 54% des contrats multi supports ne sont investis qu'à hauteur d'un tiers de fonds en unités de compte.

Les auteurs remarquent que ces chiffres sont stables depuis 2010, les français privilégiant la sécurité au rendement de leurs investissements depuis la crise de 2008. Ils mettent aussi en avant le fait que la détention de ces produits augmente avec l'âge car l'assurance-vie permet la transmission du patrimoine accumulé dans des conditions fiscales avantageuses.

I.2.2 Un placement pour préparer la retraite ?

Selon une étude de l'INSEE (2011) sur l'assurance-vie en 2010, les trois principales causes de détention sont pour les ménages résidant en France métropolitaine : avoir une réserve d'argent en cas d'imprévu (38% des ménages détenteurs), faire bénéficier un proche (29%) et préparer ou compléter la retraite (27%).

Cependant seulement 15% d'entre eux déclarent posséder une assurance-vie afin de bénéficier de ces rendements quelque soit la tranche d'âges interrogée. La détention d'un contrat se place principalement dans la cadre d'une préparation à la retraite, ce qui entraîne une ancienneté moyenne de 10 ans en 2010 : 32% des contrats ont au moins 12 ans et 13% au moins 20 ans.

L'objectif du placement de son patrimoine dans les contrats d'assurance-vie peut donc reposer davantage sur un souhait de transmission et de réserve d'argent que de rendement. Ceci peut conduire à des difficultés d'analyse d'un effet lié à la conjoncture économique sur la décision de rachat. Le détenteur du contrat serait moins impacté par les fluctuations des taux proposés par d'autres produits financiers disponibles sur le marché.

Le portefeuille de la Mutuelle est caractérisé par un âge avancé de ses adhérents. Il paraît par conséquent vraisemblable que la transmission d'un capital qualifierait l'objectif de placement des adhérents.

n'y a pas de risque de perte en capital. Le capital placé sur un fonds en euros, par ailleurs garanti, ne peut ainsi qu'augmenter.

I.3 Le risque de rachat

Le rachat sur les contrats d'assurance vie est le droit donné au souscripteur de retirer tout ou une partie de son placement capitalisé au prix éventuel de pénalités.

I.3.1 Les motivations incertaines

Le suivi des produits d'épargne est essentiel car la décision de racheter son placement peut être le fruit d'une décision rationnelle ou non.

Akerlof and Shiller (2009) présentent une vision keynésienne de l'économie et l'ont appliqué à une étude consacrée à l'épargne. Parienty (2009) précise qu'ils reprennent la notion d'« esprits animaux » de Keynes liée à l'incertitude.

L'économie néoclassique traditionnelle traite la décision d'épargne comme une décision rationnelle résultant d'un arbitrage intertemporel entre consommer immédiatement ou plus tard. Les auteurs démontrent au contraire que cette conception peut être remise en question car les individus présentent des difficultés à accorder de l'importance à l'avenir lointain ; ainsi le taux d'épargne peut varier considérablement d'une période à l'autre.

Les raisons d'un rachat peuvent donc être diverses : visions du monde (les histoires), facteurs institutionnels (les politiques publiques) ou culturels, *etc.*

I.3.2 Les conséquences

Le risque de rachat est un point d'attention dans le suivi des produits d'épargne. De par son caractère aléatoire il est très difficile à modéliser.

La sortie anticipée d'un assuré peut conduire l'assureur à ne pas recouvrir intégralement les dépenses engagées lors de l'acquisition et à ne pas prélever les frais de gestion sur encours.

Par ailleurs, une hausse brutale des taux d'intérêt pourrait déclencher des rachats massifs, obligeant alors la Mutuelle à réaliser une partie de son actif alors que la remon-

tée des taux entraînerait une baisse de la valorisation de son portefeuille de produits obligataires.

L'incertitude liée à l'estimation des taux de rachat peuvent donc impacter différents calculs comme le prix de l'assurance, l'évaluation des liquidités (Kuo et al., 2003) et l'évaluation de la rentabilité (Loisel and Milhaud, 2011). Le détail est présenté en annexe A.3, page 78.

Les conséquences d'une erreur d'estimation des taux de rachat peuvent mettre l'assureur en difficulté face à ces engagements. Il est donc nécessaire de s'efforcer à les estimer au mieux.

I.3.3 L'encadrement des risques

Dans le cadre de Solvabilité II, la projection des flux futurs probables de trésorerie en cas de vie doivent prendre en compte les options contractuelles et les garanties financières incluses dans les contrats. Le principe de la modélisation est de s'approcher au mieux de la réalité de la vie des contrats en décrivant le comportement des assurés et de l'assureur dans le temps et dans différents environnements économiques.

Les compagnies d'assurance sont en mesure d'observer deux types de rachat :

structurel : dépend uniquement des conditions du produit/adhérent (ancienneté, âge atteint, taux minimum garanti, *etc.*), ils sont observés *toutes choses égales par ailleurs* ;

conjoncturel : dépend de facteurs exogènes tels que le différentiel de taux servis entre l'assureur et celui de la concurrence, les conditions des marchés financiers. La modélisation repose sur une fonction dépendant uniquement de l'écart entre le taux servi et un taux dépendant de l'environnement économique, souvent appelé taux de rendement espéré par l'assuré (cf. Orientations Nationales Complémentaires¹⁰).

10. Les orientations de l'EIOPA sont des textes dits de niveau 3 dont l'objectif est de favoriser l'harmonisation des pratiques de contrôle en Europe. Elles sont adoptées par l'EIOPA puis, au sein de chaque autorité nationale, soumises à une procédure dite de *comply or explain*.

Le régime Solvabilité II impose aux assureurs une mesure prudente de leurs engagements et par conséquent un suivi rigoureux des risques afférents à leur portefeuille.

En conséquence, comment peuvent-ils estimer l'évolution du comportement de leurs assurés afin de garantir leur solvabilité ?

I.4 Mesure du risque rachat

Les recommandations de l'ACPR précisent que l'évaluation des rachats dans la projection, pour les produits d'épargne investis sur le fonds en euros avec participation aux bénéficiaires, doit être modélisée par une loi de rachat dynamique incluant une loi de rachats conjoncturels et une loi de rachats structurels.

I.4.1 Modélisation actuelle des rachats à La France Mutualiste

RMC

Les produits RMC en phase de cotisation sont considérés comme non rachetables à la vue des faibles montants de prestations¹¹. La part des produits d'épargne convertis en rentes ainsi que les produits de rentes différées en phase pensionnée ne sont pas rachetables au vu de l'article L. 223-22 du Code de la Mutualité. Par conséquent, des taux de rachats fixes nuls sont retenus en ce qui concerne ces produits.

Unités de compte

Pour la part des UC des produits multi supports, une loi de rachats fixe est appliquée. Pour la part Euro des produits mono et multi supports, une loi de rachats *dynamique* par produit est définie. Le taux de rachat global estimé, diffusé par le service Actuariat, est alors réparti entre la partie structurelle et la partie conjoncturelle.

11. Par exemple, les deux derniers taux de rachats annuels de 2015 et 2016 s'élèvent respectivement à : 0,09% et 0,08%.

Partie conjoncturelle

La partie conjoncturelle est évaluée à l'aide d'une fonction de rachats conjoncturels proposée lors de la cinquième étude d'impact (voir CEIOPS (2010)). Elle dépend de l'écart entre le taux servi à l'assuré l'année précédente et le taux attendu par ce dernier. Selon les préconisations de l'ACPR (2013), le taux attendu par l'assuré peut être évalué comme le Taux Moyen d'emprunt d'État (TME). Il représente le taux de rendement sur le marché secondaire des emprunts d'État à taux fixe supérieurs à 7 ans¹². La partie structurelle est ensuite déduite à partir du taux de rachat global.

Dans le modèle de la Mutuelle, pour estimer au mieux le TME, le taux attendu ou le taux cible de revalorisation est évalué à travers le taux d'intérêt 10 ans moyen des trois dernières années de projections car son horizon est semblable par hypothèse à ceux des contrats d'épargne.

À partir des résultats de la première année de projection, il est possible de distinguer les taux de rachat suivant ces étapes :

- le taux attendu par l'assuré est le taux d'intérêt 10 ans projeté,
- le taux d'intérêt 10 ans peut être calculé en déterministe et sa variance est faible en stochastique (cadre du pilier I¹³ où la projection est réalisée en risque neutre),
- un taux de rachat conjoncturel par produit est calculé à partir de ce taux d'intérêt 10 ans et de la fonction de rachat conjoncturelle choisie,
- un taux de rachat structurel par produit est obtenu par différence.

Cette technique permet de reconstituer un taux de rachat structurel qui sera ensuite impacté lors de simulations stochastiques de scénarios de tensions économiques.

12. A noter : le taux TME est par exemple, le taux d'intérêt applicable suite à l'acquisition d'obligations assimilables du Trésor (OAT) et reversé aux propriétaires des obligations chaque année.

13. Le pilier I de Solvabilité II regroupe les exigences quantitatives, c'est-à-dire les règles de valorisation des actifs et des passifs, ainsi que les exigences de capital et leur mode de calcul.

Partie structurelle

Les taux de la partie structurelle des lois de rachats dynamiques et les taux de la loi de rachat fixe des produits en UC sont répartis en rachats totaux et rachats partiels.

Les rachats totaux diminuent le nombre de contrats contrairement aux rachats partiels. Cet élément de modélisation a un impact sur la projection des frais puisque les frais de gestion définis en coût unitaire sont dépendants du nombre de contrats.

I.4.2 Choix de la fonction de rachats conjoncturels

La Mutuelle a choisi de se référer aux lois proposées par l'ACPR dans le cadre de la modélisation des rachats. Elle se décompose entre un taux de rachats conjoncturels RC et un taux de rachats structurels RS . Elle s'appuie sur le taux servi R et un taux attendu TA par les assurés. Si le taux servi est inférieur à TA alors le rachat est supérieur au taux de rachat indiqué par la courbe de RS . Le taux RC est donc une fonction de R et TA .

Elle se présente sous forme de paliers, figure I.3 page suivante, reposant sur six paramètres : α , β , γ , δ , RC_{max} et RC_{min} . L'interprétation des paramètres est la suivante :

- α est le seuil en-deçà duquel les rachats conjoncturels sont constants et fixés à RC_{max} . Ce n'est plus l'écart de taux qui explique le comportement des assurés ;
- β et γ sont respectivement les seuils d'indifférence à la baisse et à la hausse du taux servi. Entre ces deux seuils, le comportement de l'assuré n'est pas modifié ;
- δ est le seuil au-delà duquel la diminution du taux de rachat structurel est constante et fixée à RC_{min} . Ce n'est plus l'écart de taux qui explique le comportement des assurés.

Lors de la cinquième étude quantitative d'impact (voir CEIOPS (2010)), l'ACPR a proposé plusieurs fonctions de rachats conjoncturels : la fonction *Min*, *Moy* et *Max*. Les paramètres sont repris au tableau I.1, page ci-contre. Elles ont toutes les trois une forme par paliers comme présentée à la figure I.3, page suivante.

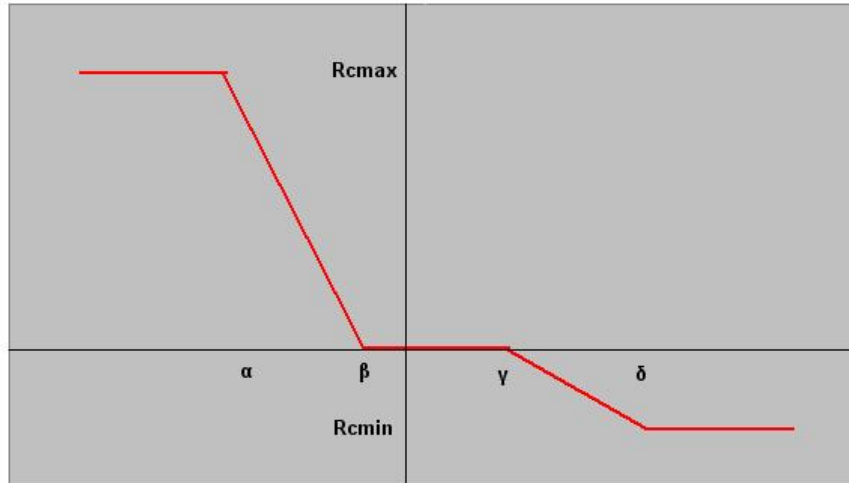


FIGURE I.3 – Loi de rachats conjoncturels ACPR

Fonction	α	β	γ	δ	RC_{min}	RC_{max}
Min	-6%	-2%	1%	2%	-6%	20%
Moy	-5%	-1%	1%	3%	-5%	30%
Max	-4%	0%	1%	4%	-4%	40%

TABLE I.1 – Fonctions de rachats conjoncturels proposées par l'ACPR

Elles se présentent analytiquement sous la forme générique suivante :

$$RC(R,TA) = \begin{cases} RC_{max} & \text{si } R-TA < \alpha \\ RC_{max} \times (R-TA-\beta)/(\alpha-\beta) & \text{si } \alpha < R-TA < \beta \\ 0 & \text{si } \beta < R-TA < \gamma \\ RC_{min} \times (R-TA-\gamma)/(\delta-\gamma) & \text{si } \gamma < R-TA < \delta \\ RC_{min} & \text{si } R-TA > \delta \end{cases} \quad (I.1)$$

Avec :

- RC : taux de rachat conjoncturel pour l'année N
- R : taux servi l'année $N - 1$
- TA : taux attendu par les assurés

Le taux de rachat total RT s'exprime alors comme :

$$RT(R, TA, \cdot) = \min\left(1, \max\left(0, RS(\cdot) + RC(R, TA)\right)\right) \quad (\text{I.2})$$

Le portefeuille des assurés de la Mutuelle ayant un âge moyen avancé, et étant par conséquent supposés peu enclin à arbitrer leur épargne en fonction de la situation économique, la fonction Min a été retenue dans le cadre des projections des flux de trésoreries.

I.5 Problématique

La difficulté rencontrée lors de l'utilisation de ces hypothèses est la suivante : les taux de rachats conjoncturels estimés par le modèle sont négatifs pour tous les produits. Cela signifie que l'écart entre le taux servi et le taux attendu se situe entre les paramètres γ et δ de la fonction de rachat conjoncturelle. La pente située entre ces paramètres est très importante dans la fonction Min choisie pour la projection. La valeur de RC_{min} est trop élevée par rapport à la valeur du taux de rachat total constaté.

Le département de Gestion Actif/Passif (ALM) a donc conduit plusieurs travaux sur la sensibilité des indicateurs prudentiels à la fonction de rachats conjoncturels et sur la ventilation des rachats totaux et partiels. Ces études ont démontré le besoin d'une construction appropriée d'une loi structurelle des rachats et une détermination plus adaptée des paramètres de la loi conjoncturelle des rachats aux comportements des adhérents de la Mutuelle.

Revue méthodologique et analyse exploratoire

La construction des lois de rachat est encadrée par la réglementation et plusieurs recommandations sont proposées par l'Institut des Actuaire (2016).

Le monde universitaire n'est pas en reste puisqu'il s'est aussi beaucoup intéressé au sujet des origines de la décision de rachat des contrats d'épargne. Ce dernier présente certaines hypothèses et outils de modélisation qui peuvent être appliqués aux données de la Mutuelle.

L'étude du portefeuille de contrats tentera dans une première approche de découvrir un lien entre un taux attendu et la décision du rachat total de l'épargne acquise.

II.1 Revue de la littérature

II.1.1 Les hypothèses

D'après Cerchiara et al. (2008), les variables explicatives d'une loi comportementale des rachats peuvent être classées suivant qu'elles modélisent une décision rationnelle ou irrationnelle.

La première provient de l'évolution des marchés financiers tandis que la dernière résulte des caractéristiques intrinsèques du souscripteur et du contrat.

L'hypothèse du taux d'intérêt

L'hypothèse du taux d'intérêt, dénommée *Interest Rate Hypothesis*, suppose que le taux de rachat augmente lorsque les taux d'intérêt des marchés augmentent.

En effet les investisseurs réagiraient à partir d'un coût d'opportunité afin d'obtenir de nouveaux contrats d'assurance. Lorsque les taux d'intérêt augmentent, la prime d'équilibre baisse, c'est pourquoi il y a plus de chances de voir apparaître des contrats proposant la même couverture mais à un prix plus faible. Les souscripteurs se tourneront alors vers des produits à meilleurs rendements disponibles sur le marché (Loisel and Milhaud, 2011).

Le taux de rachat est dans ce cas de figure négativement corrélé au taux de rendement interne (surplus de participation), et positivement lié au taux de rendement externe (taux d'intérêt des marchés), d'après les auteurs Dar and Dodds (1989) ou encore Kuo et al. (2003).

En résumé, cette hypothèse repose sur le fait que le souscripteur peut vouloir retirer ses fonds dans le but de tirer avantage d'un taux d'intérêt plus élevé.

L'hypothèse de fond d'urgence

Une autre hypothèse, celle de fond d'urgence, qualifiée comme *Emergency Fund Hypothesis*, est présentée par Outreville (1990). Elle met en avant le fait que le rachat d'un contrat peut être le résultat d'un besoin en capital lors d'une situation de difficultés personnelles (voir aussi Milhaud et al. (2011)).

Cette décision peut provenir d'un désir d'utiliser les fonds qui auraient dû aller aux paiements des primes pour d'autres besoins importants ou un désir de prendre les sommes accumulées pour couvrir diverses dépenses du ménage.

II.1.2 Les modèles proposés

Il existe en particulier deux modèles utilisés à partir de l'application de ces hypothèses dans le domaine de l'assurance. Le premier est considéré comme un modèle de statistiques d'apprentissages ou Machine Learning¹, l'arbre CART². Le second est un modèle statistique économétrique classique, le Modèle Linéaire Généralisé (Generalized Linear Model ou GLM). Le lecteur curieux pourra se référer à Charpentier et al. (2018) sur les différences fondamentales entre l'économétrie et l'apprentissage statistique.

Les GLM sont une extension aux modèles de régression linéaire largement utilisés et proposent de nombreux modèles s'affranchissant des relations linéaires et de l'hypothèse de normalité. La première application au domaine de l'assurance-vie a été apportée par Haberman and E. Renshaw (1996).

D'après Cerchiara et al. (2008), les modèles linéaires généralisés sont assez puissants pour la recherche de la relation entre des variables explicatives et une variable réponse observée, telle le comportement de rachat. Les GLM sont capables de mesurer la direction et l'intensité de l'impact sur le comportement de rachat d'une variation des caractéristiques retenues.

D'après Briere-Giroux et al. (2010), les GLM sont connus pour donner des résultats précis lorsqu'ils sont appliqués aux données de rachats, semblables aux modèles CART mais tout en demeurant interprétables.

Un modèle GLM peut être utilisé avec différents types de fonctions lien et peut prédire aussi bien des variables continues que binaires. Cette caractéristique implique sa grande utilisation dans le domaine de l'assurance puisque la variable représentant l'événement du rachat peut être représentée soit par 0 soit par 1 au niveau contrat ou continue entre 0 et 1 à un niveau agrégé. Pour cette raison, la théorie des modèles GLM sera employée par la suite.

1. Le Machine Learning est une technologie d'intelligence artificielle permettant aux ordinateurs d'apprendre sans avoir été programmés explicitement à cet effet.

2. L'apprentissage par arbre de décision désigne une méthode basée sur l'utilisation d'un arbre de décision comme modèle prédictif. L'algorithme CART dont l'acronyme signifie « Classification And Regression Trees », s'attelle à construire un arbre de décision en classifiant un ensemble d'enregistrements. Cet arbre fournit un modèle pour classer de nouveaux échantillons. Il a été publié par Leo Breiman en 1984.

II.2 Les données

II.2.1 Structure

Les données sont extraites du système de gestion. La base des contrats est construite à partir des observations de fin de mois entre le 31 janvier 2010 et le 30 novembre 2016. Aucune donnée contrat par contrat n'est disponible avant l'année 2010. L'objectif est de pouvoir modéliser un taux de rachat total par mois de survenance. La formule est la suivante :

$$\text{Taux de rachat du mois } M = \frac{\text{Montant racheté du mois } M}{\text{Provisions Mathématiques du mois } M - 1} \quad (\text{II.1})$$

Le montant du rachat est associé à la dernière situation du contrat en cours (exemple présenté au tableau A.1, page 79, en annexe). Les écarts observés avec les montants réglés enregistrés en comptabilité et les montants récupérés à partir du système de gestion sont négligeables. Cette méthode permet d'être cohérent entre la partie exposition des provisions mathématiques des contrats et le moment du règlement du rachat pouvant intervenir avec du retard.

Le manque de recul historique des produits multi supports et la faiblesse des taux de rachats sur les produits historiques en mono support de la Mutuelle contraignent à ne pas distinguer le type de contrat. Il est possible de cette manière d'obtenir assez d'observations présentant un rachat total. La loi dynamique ne s'appliquant qu'aux provisions mathématiques du fonds Euro, seuls les contrats possédant un encours non nul sont retenus.

II.2.2 Délai de règlements

Les statistiques sur les délais de règlements en nombre de jours, c'est-à-dire le délai entre la dernière date d'observation du contrat (date d'inventaire) et la date de règlement, sont présentées, dans le tableau II.1 page suivante, par année de survenance du rachat.

Elles permettent de mettre en évidence le décalage potentiel observé entre un taux de rachat par survenance et un taux de rachat comptable.

Année	N	Moyenne	Ecart-type	Mode
2010	594	1,72	4,94	1
2011	746	1,24	10,50	1
2012	741	8,81	11,02	1
2013	659	13,09	123,21	8
2014	598	18,99	11,65	7
2015	669	20,35	15,29	9
2016	800	17,48	60,35	7

N : nombre d'observations.

Mode : valeur la plus fréquente d'une série statistique.

TABLE II.1 – Délai en jours entre la date d'inventaire et le règlement du rachat par année de survenance

Le principe adopté dans l'étude est de pouvoir s'affranchir des retards de gestion qui peuvent se produire lors du règlement d'un rachat. Les années 2010 à 2012 ont de faibles moyennes et modes car beaucoup de dates de règlement se positionnent le 1^{er} jour du mois. L'ancienne programmation des dates de règlement fixait toutes les dates au premier jour du mois de règlement. Les statistiques sont plus vraisemblables à partir de 2013 où la Mutuelle procède aux règlements, la plupart du temps, sous une semaine (en observant le mode) et généralement dans le mois (en observant la moyenne). Les fortes valeurs de l'écart type en 2013 et 2016 sont dues à deux très fortes valeurs aberrantes. La Mutuelle s'efforce donc de régler le montant de rachat sous un mois.

La construction du taux de rachat à partir de la formule II.1, page ci-contre, semble approprier car le dénominateur reflète de manière adéquate le risque sous-jacent.

II.3 L'événement et ses origines

La variable à expliquer est la réalisation d'un rachat total, y compris les termes et les renonciations, enregistré uniquement sur le fonds Euro. Elle est codée d'une façon binaire c'est-à-dire 1 pour la présence d'un rachat, 0 sinon. Le montant du règlement des rachats réalisés est aussi incorporé à la base de données exploitées. Il est fréquent de

construire un taux de rachat en montant plutôt qu'en nombre, afin de mettre en évidence une éventuelle fuite des encours les plus importants.

Le graphique II.1, de la présente page, montre un taux mensuel toujours inférieur à 0,10%, ce qui représente un taux annualisé inférieur à 1,20%. Ce taux est très bas puisque selon Frey and Capitaine (2017), le taux de rachat moyen observé en 2016 sur le marché est de 4,8% pour les contrats n'ayant pas de pénalités de rachat comme ceux étudiés.



FIGURE II.1 – Taux de rachat total du fonds Euro

Les adhérents de la Mutuelle semble avoir davantage eu recours à l'option de rachat entre 2011 et 2012 d'après l'observation du graphique II.1, de la présente page. En cause peut-être une conjoncture morose basée sur la crise des dettes souveraines de la zone Euro. D'après Frey (2016), selon les données prudentielles collectées par l'ACPR, le taux de rachat sur l'ensemble des contrats d'assurance-vie individuels est passé de 4,7% en moyenne en 2010 à 6,1% en 2011, avec un pic à 7,7% au dernier trimestre de 2011. Un historique plus profond aurait permis de confronter les taux de rachats après la crise des *subprimes*³ en 2007.

Le creux observé à la suite de l'été 2010 et le pic observé en mars 2016 résultent des retards de gestion accumulés lors d'opérations commerciales (hausse du plafond de la RMC) ou de périodes de forte mortalité.

3. La crise des subprimes (en anglais : subprime mortgage crisis) est une crise financière qui touche le secteur des prêts hypothécaires à risque (en anglais : subprime mortgage) aux États-Unis à partir de juillet 2007. Avec la crise bancaire et financière de l'automne 2008, ces deux phénomènes inaugurent la crise financière mondiale de 2007-2008.

En rapprochant la forme de la courbe avec le contexte économique il semblerait que le portefeuille soit sensible à des situations de tension conjoncturelle. Le problème désormais est de découvrir de quelle manière il est possible de modéliser la différence de taux $R-TA$, comme il est préconisé dans la formule I.1, page 15.

II.3.1 Analyse exploratoire

Selon les préconisations fournies par l'Institut des Actuaires (2016), diverses variables ont été construites au préalable dans le but de déceler l'effet structurel. Les caractéristiques de l'adhérent ou du contrat sont retenues au sein de l'étude dans le but d'expliquer la décision de racheter son contrat.

Classe : ancienneté du contrat

Les tranches d'ancienneté sont basées sur la fiscalité de l'assurance-vie (hors loi *Flat Tax*⁴ puisqu'elle n'est entrée en vigueur qu'au 27 septembre 2017) : a) 0an - 3ans, b) 4ans - 7ans, c) 8ans - 11ans, d) 12ans - 15ans, e) 16ans - 19ans et f) sup 20ans.

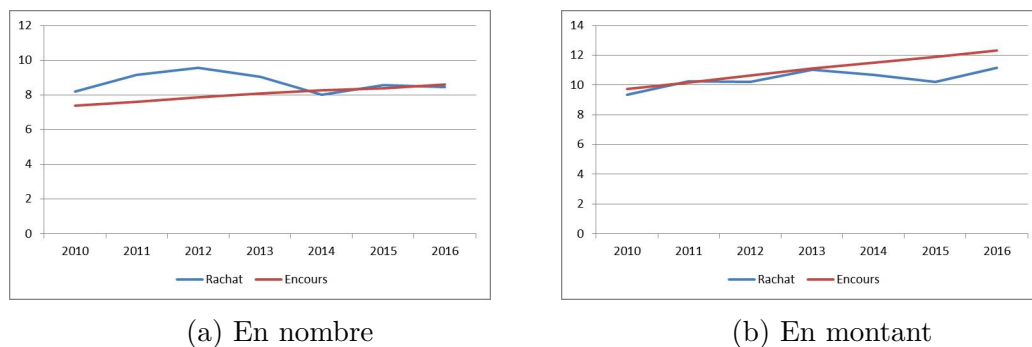


FIGURE II.2 – Ancienneté moyenne de l'encours et des rachats

L'ancienneté moyenne des contrats rachetés observée sur le graphique II.2, de la présente page, est supérieure à celle du portefeuille sur les premières années observées. Ce-

4. En 2017, le gouvernement a instauré une forme de flat tax avec la mise en œuvre d'un nouveau prélèvement forfaitaire unique (PFU) sur tous les revenus de l'épargne : dividendes, assurance-vie, intérêts, plus-values, revenus fonciers... Auparavant, ces revenus se voyaient appliquer des régimes fiscaux différents.

pendant l'analyse en montant révèle le contraire. Les contrats plus récents seraient moins fidèles que ceux du portefeuille.

Le graphique II.3, de la présente page, fait apparaître un mouvement à la baisse des rachats au cours des années observées quelque soit la tranche d'ancienneté. Tous les adhérents, exceptés les plus anciens, semblent par conséquent avoir été sensible à la conjoncture des années 2010 et 2011.

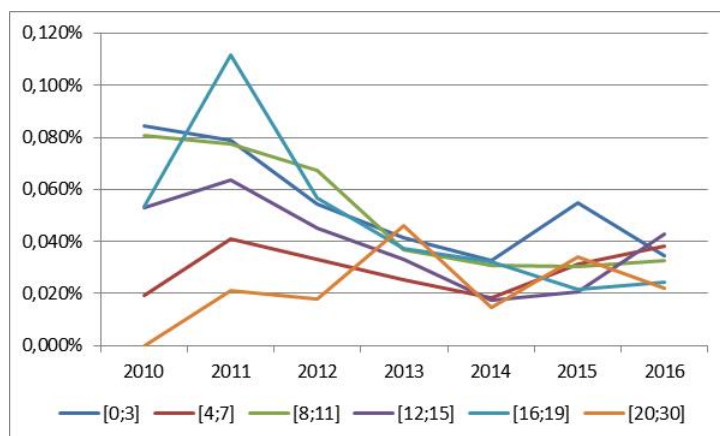


FIGURE II.3 – Taux de rachat en montant par tranches d'ancienneté

Classe : âges de l'adhérent

Les tranches d'âges sont basées sur l'étude INSEE (2011) : a) inf 30ans, b) 30ans - 49ans, c) 50ans - 59ans, d) 60ans - 69ans et e) sup 70ans.

La détention d'une assurance-vie après 70 ans présente une opportunité fiscale. Ainsi, contrairement à la fiscalité applicable pour les versements réalisés avant l'âge de 70 ans, seules les primes (c'est-à-dire la somme des versements initiaux ou périodiques ou exceptionnels) seront taxables aux droits de succession après un abattement de 30 500€.

Plus la vie de l'épargnant est longue, plus le montant des intérêts capitalisés non taxables au droits de succession seront élevés. Contrairement à la fiscalité applicable aux sommes versées avant les 70 ans de l'épargnant, le temps joue en faveur de ce dernier. La fiscalité après 70 ans pousserait les adhérents à choisir le dénouement du décès, ce qui peut avoir pour conséquence l'observation de taux de rachat total faibles.

La présentation du graphique II.4, de la présente page, en nombre plutôt qu'en montant, est révélatrice car elle permet de montrer qu'au contraire ce sont les adhérents plus âgés que la moyenne qui rachètent leur contrat.

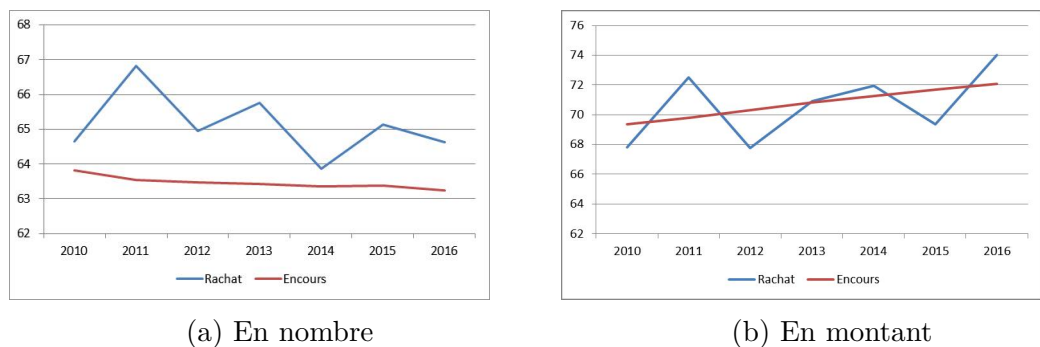


FIGURE II.4 – Age moyen de l'encours et des rachat

Le graphique II.5, de la présente page, fait apparaître le même mouvement qu'au graphique précédent par ancienneté, II.3, page ci-contre. Les taux de rachat entament, pour la plupart des tranches d'âges, une baisse à partir de l'année 2012. Ces deux sources de rachat, l'âge et l'ancienneté, paraissent s'atténuer à l'aide d'un phénomène qui peut se révéler être un effet conjoncturel.

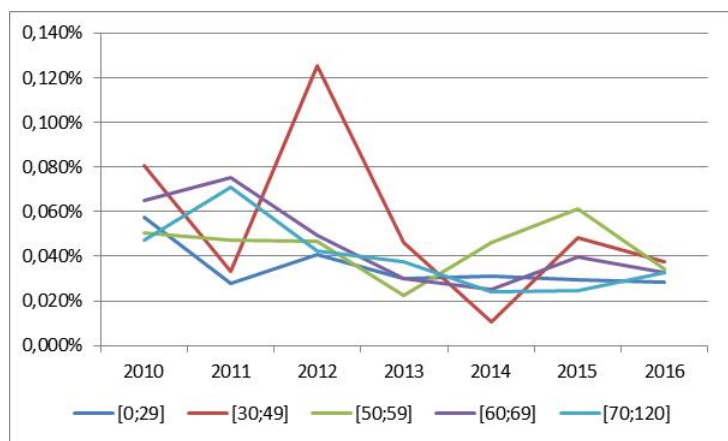


FIGURE II.5 – Taux de rachat en montant par tranches d'âges

Classe : provision mathématique du contrat

Les tranches d'encours sont basées à partir de statistiques réalisées sur le portefeuille durant toute la période étudiée : a) inf 8k, b) 8k - 20k, c) 20k - 44k et e) sup 44k.

Le graphique II.6, de la présente page, montre que les rachats sont issus de contrats avec une provision mathématique moins importante que les contrats en cours. De même que pour le critère de l'ancienneté, le montant des rachats serait inférieur à celui de l'encours moyen du portefeuille car les adhérents demeurant moins longtemps épargne moins.

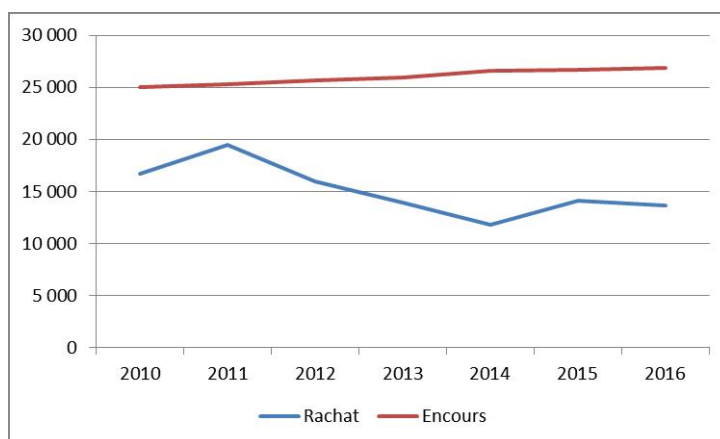


FIGURE II.6 – Montant moyen de l'encours et des rachat

Classe : taux minimum garanti du versement

Enfin une dernière variable statistique quantitative est créée en fonction de la présence d'un Taux Minimum Garanti moyen sur le contrat : a) TMG nul, b) TMG non nul et c) $TMG > Taux_servi$.

Le graphique II.7, page ci-contre, met en évidence la présence marginale des contrats pour lesquels le taux minimum garanti sur les versements est supérieur au taux de participation aux bénéfices moyen servi par la Mutuelle. Elle ne représente qu'en moyenne 7% des provisions mathématiques sur la période pour 6% des montants rachetés. Il est évident que cette décision de racheter son contrat concerne le cas de l'hypothèse de fond d'urgence et non celle de rendement.

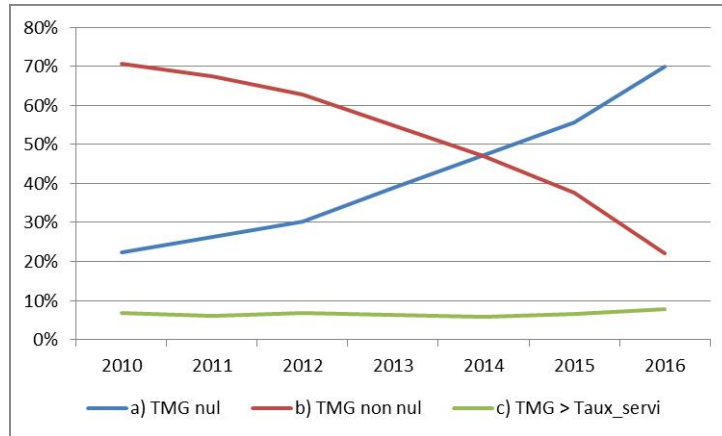


FIGURE II.7 – Montant des provisions mathématiques selon la présence d’un taux minimum garanti au contrat

II.3.2 Conclusion

Les variables présentées ci-dessus sont toutes associées à la raison structurelle des rachats c’est-à-dire à l’hypothèse de fonds d’urgence (EFH). En effet les caractéristiques des contrats et le profil de l’adhérent favorisent dans certaines circonstances une anticipation du retrait de toute son épargne.

L’analyse de ces graphiques mène à la conclusion qu’un adhérent plus âgé aurait tendance à plus racheter tout comme un adhérent possédant un contrat récent.

L’objectif de la préparation à la retraite peut alors être mis en évidence dans la première observation. L’adhérent semble en effet avoir besoin de disposer d’une réserve d’argent dans laquelle il pourrait puiser en cas de nécessité. Quant à la seconde, il pourrait s’agir de cas de nouveaux entrants plus sensibles aux besoins de leur ménage ou aux produits de la concurrence.

II.4 Le taux attendu et le taux servi

Le premier problème auquel est confronté la construction de la loi conjoncturelle est la détermination d’un taux adverse.

II.4.1 Choix du taux attendu

Une hypothèse intéressante à tester serait de se reposer sur un taux réglementaire auquel l'ensemble des acteurs du marché de l'assurance-vie devrait se référer dans le but de proposer un taux concurrentiel.

L'Arrêté⁵ du 7 juillet 2010 (annexe A.7 page 82) portant modification des modalités de garanties d'un taux minimum par les entreprises d'assurance vise à mieux encadrer les taux garantis dans le but de :

1. assurer plus d'équité entre assurés,
2. garantir une information plus loyale des assurés,
3. empêcher que des promesses déraisonnables ne fragilisent la situation prudentielle des assureurs.

Les taux garantis seront notamment plafonnés à un niveau qui dépendra de l'environnement des marchés obligataires et donc de la rémunération pouvant être espérée à moyen terme sur le contrat. En cas de baisse significative des taux, l'assureur doit en revanche limiter ses offres commerciales à un taux ayant pour référence 110% du Taux Moyen des emprunts d'État (TME). En cas de hausse significative des taux, l'assureur peut dynamiser son offre commerciale en proposant un taux ayant pour référence 90% du TME.

Le taux attendu TA pourrait, à partir de ces conditions stipulées, être estimé à partir du TME. Il peut aussi être retenu comme le taux attendu parce qu'il est déjà employé dans le modèle ALM, du fait de sa publication mensuelle et qu'il détermine un couloir en ce qui concerne les taux garantis.

II.4.2 Conséquences

Un taux servi unique a été considéré compte tenu qu'aucune distinction par produit n'a été retenue. Face au problème d'un rachat structurel potentiellement prédominant,

5. Journal officiel de la république française numéro 0174 du 30 juillet 2010 page 14111 texte numéro 13 modifiant le Code des assurances (articles A. 132-2, A. 132-3 et A. 331-5).

il a été décidé de modéliser la « santé économique » de la Mutuelle. En d'autres termes, ceci permet de tenir compte d'un seul comportement reflétant l'hypothèse de taux de rendement.

L'observation sur le graphique II.8, de la présente page, des taux moyens servis $N - 1$ montre qu'il décroît (exceptée l'année 2013) sur la période observée en même temps que le TME, cependant pas à la même vitesse. Compte tenu de la baisse des taux, l'écart taux servi et taux attendu est peu souvent négatif et s'amplifie à partir de l'année 2012.

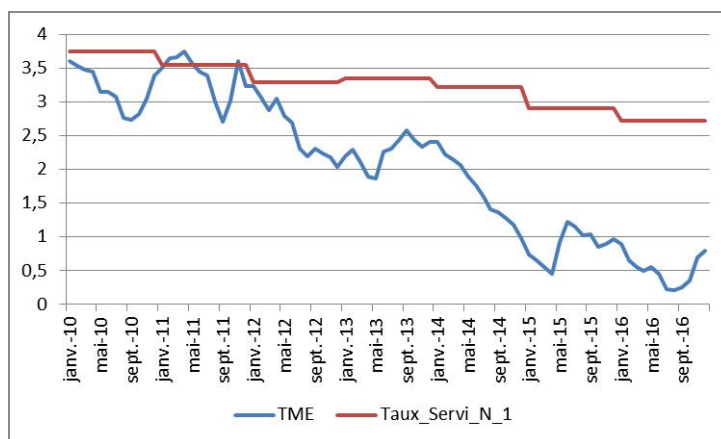


FIGURE II.8 – Valeurs mensuelles du TME et Taux servi moyen $N - 1$

Les données observées semblent appropriées à la construction de la loi dynamique sur le segment comprenant les valeurs γ et δ , autrement dit le moment où la loi conjoncturelle vient diminuer les rachats structurels.

II.5 Premier résultat

Les valeurs de la variable *delta taux* sont pratiquement toujours positives. Les valeurs négatives seront par conséquent retirées de l'étude parce qu'elles ne concernent que la partie gauche de la loi conjoncturelle, où le taux attendu est supérieur au taux servi.

Dans l'ensemble, la tendance à la baisse est telle qu'elle est décrite à la figure II.9, page suivante.

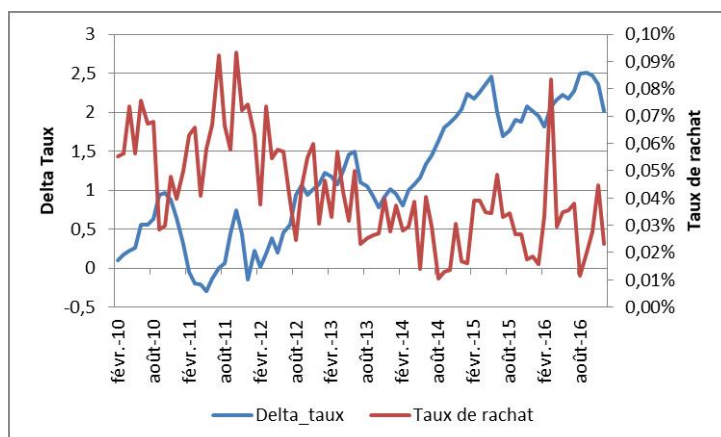


FIGURE II.9 – Valeurs de *delta taux* et taux de rachat total du fonds Euro

Toutefois l'erreur serait d'interpréter ces deux phénomènes de structure et de conjoncture indépendamment l'un de l'autre. L'analyse est d'autant plus difficile qu'il faut extraire assez d'informations pour construire une loi de rachat dynamique.

La méthode sera alors d'élaborer dans un premier temps une loi conjoncturelle pour ensuite en déduire la loi structurelle.

En effet, le choix s'est porté sur l'importance de dégager tout d'abord l'effet attribué à l'environnement économique s'imposant à tous les adhérents à un période donnée.

Hypothèse du taux d'intérêt

La raison de racheter son contrat d'épargne peut reposer sur une décision rationnelle liée à l'environnement économique.

Ce chapitre présente dans un premier temps deux outils classiques dont l'objectif est de déterminer son effet sur le taux de rachat.

Dans un second temps, il est démontré de manière théorique pourquoi l'utilisation d'un modèle de choix discrets semble être appropriée à la détermination d'une grandeur financière. Cette dernière pouvant influencer la volonté de récupérer son épargne lorsqu'un contrat alternatif semble plus rentable.

III.1 Mesure de l'intensité et du sens de la relation

Une première approche consiste à tester une relation linéaire entre le taux de rachat et la variable *delta taux*. Le nuage de points III.1, page suivante, est l'outil privilégié afin d'analyser la liaison entre deux grandeurs statistiques.

Le coefficient de corrélation de Pearson¹ obtenu présente une valeur de $-59,47\%$. Il mesure la tendance qu'ont ces deux variables à être simultanément au dessus ou en dessous de leurs espérances respectives.

Par conséquent il existerait une liaison monotone entre *delta taux* et le taux de rachat en montant, sur la période observée. Le sens de la relation se lit de la manière suivante : le taux de rachat diminue lorsque le l'écart de taux augmente.

Cependant elle pourrait être éventuellement non monotone car les points ne révèlent pas une forme curviligne et semblent ne pas s'ajuster autour d'une droite.

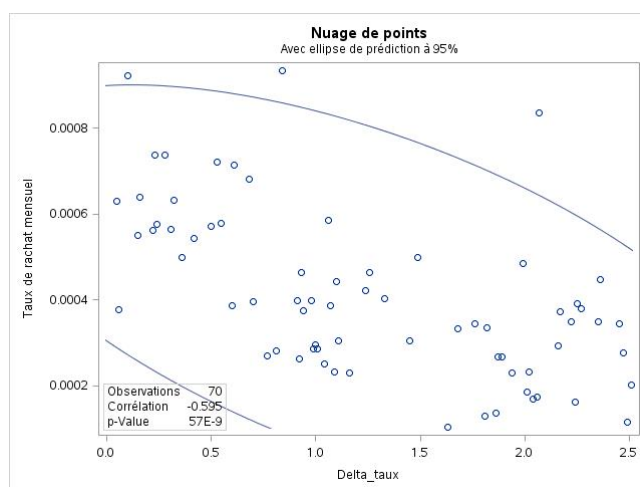


FIGURE III.1 – Taux de rachat mensuel en fonction de *delta taux*

Un autre moyen de mettre en évidence une liaison est le coefficient de corrélation de rang, aussi appelé coefficient de Spearman. Il examine l'existence d'une relation entre les rangs des observations pour deux variables dans le but de détecter une relation monotone quelque soit sa forme (linéaire, exponentiel, puissance, *etc.*).

Au demeurant, l'adhérent ne suit, en général, pas l'actualité financière au jour le jour. Par conséquent il existerait peut être un lien plus fort entre le taux de rachat observé et des décalages de *delta taux* dans le temps.

1. La corrélation est une quantification de la relation linéaire entre des variables continues. Le calcul du coefficient de corrélation de Pearson repose sur le calcul de la covariance entre deux variables continues. Le coefficient de corrélation est en fait la standardisation de la covariance. Cette standardisation permet d'obtenir une valeur qui variera toujours entre -1 et $+1$, peu importe l'échelle de mesure des variables mises en relation.

Il est intéressant, par exemple, de comparer le taux servi $N - 1$ avec le TME du mois $M - 2$. Les corrélations du taux de rachat du mois M avec la valeur de *delta taux* ont été obtenues jusqu'à $M - 4$ mois, en supposant qu'une décision est prise au pire dans les 4 mois à venir. Les résultats sont présentés dans le tableau III.1, de la présente page.

	M-1	M-2	M-3	M-4
Coefficient de corrélation de Spearman	-60,85%	-58,31%	-58,54%	-58,54%
Proba $> r $ sous $H_0 : \text{Rho}=0$	<0.001	<0.001	<0.001	<0.001

TABLE III.1 – Tableau des coefficients de corrélation de Spearman entre *delta taux* et le taux de rachat

Aucun *delta taux* ne se détache significativement et prendre le taux $M - 1$ semble optimiste car la culture financière n'est peut-être pas commune à tous les individus. Le taux auquel les adhérents semblent se référer n'est alors peut-être pas le taux brut du TME observable mais une combinaison de ces derniers. L'adhérent peut en effet ne pas prendre sa décision sur une moyenne mensuelle de taux qu'est le TME mais peut-être sur l'observation d'une tendance de marché.

L'idée sera alors de tenter de construire un taux permettant d'observer nettement un mouvement de baisse des rachats tout en discriminant au mieux les individus entre eux.

III.2 Du modèle Logistique au modèle de choix discrets

L'outil le plus communément employé parmi les articles académiques, en ce qui concerne la modélisation des rachats, est le modèle logistique ou autrement dit le modèle *logit*.

III.2.1 Raisons du modèle logistique

La première démarche lors du traitement d'une variable binaire est simplement le modèle Binomial (annexe C.1, page 92). Ce dernier ne considère cependant pas le cas

des censures et des troncatures dans l'observation des données, ainsi qu'il doit être fait d'après l'Institut des Actuaire (2016).

Dans ce cas, il est alors possible de s'appuyer sur le modèle de durée de survie de Cox (annexe C.2, page 95). Nonobstant, à cause de la structure longitudinale des données, il est démontré (annexe C.3, page 97)) que ces deux modèles conduisent tous deux au modèle logistique appartenant à la famille des Modèles Linéaires Généralisés.

À ces éléments, la fonction *logit* semble être la méthode la plus adaptée à la modélisation des taux de rachat. Guszcz (2005) a comparé trois modèles avec la régression logistique et a conclu que la régression logistique est meilleure que le modèle **CART** mais moins bonne que le modèle sophistiqué **MARS**².

Un des avantages de la fonction *logit* est qu'elle permet de calculer un rapport de chance ou *odds-ratio* (Anderson et al., 2010). Il indique la chance d'opter pour un choix plutôt qu'un autre. Le facteur d'échelle augmente l'interprétation et l'application du modèle.

III.2.2 Puis du modèle de choix discrets

Les modèles de choix discrets décrivent le comportement d'un individu face à un ensemble d'actions mutuellement exclusives. Les économistes supposent généralement que les individus choisissent toujours l'alternative qu'ils préfèrent de façon déterministe.

Il semble néanmoins que les choix des individus peuvent fluctuer selon le contexte, c'est pourquoi les modèles de choix discrets adoptent une démarche probabiliste plutôt que déterministe. Les choix des individus sont par conséquent soumis à un certain aléa qui conduit à calculer des probabilités de choix pour chaque alternative et à confronter ces probabilités aux fréquences observées.

2. La Régression multivariée par spline adaptative (en anglais MARS pour « Multivariate adaptive regression splines ») est une méthode statistique; plus précisément, c'est une forme de modèle de régression présentée pour la première fois par Jerome H. Friedman et Bernard Silverman en 1991. C'est une technique de régression non paramétrique pouvant être vue comme une extension des régressions linéaires qui modélisent automatiquement des interactions et des non-linéarités.

III.2.3 Processus d'un choix

Ben-Akiva and Lerman (1985) indique que le choix peut être vu comme le résultat d'un processus séquentiel de prises de décisions suivant les étapes :

1. définition du problème de choix,
2. mise en place d'alternatives,
3. évaluation des caractéristiques des alternatives,
4. choix,
5. réalisation.

Ceci montre qu'un choix n'est pas réduit au seul fait d'une décision à un moment donné mais est plutôt le résultat d'un processus. Les choix sont alors faits sur les caractéristiques ou attributs des alternatives.

Dans le but de prendre une décision, l'individu a besoin d'une règle : l'attractivité de l'alternative peut en être une. Elle renvoie souvent à l'utilité. Cette mesure, que le décideur essaie de maximiser, est adoptée dans les modèles récents.

III.2.4 Comportement rationnel

Le terme de comportement rationnel est basé sur les croyances qu'un individu a du gain qu'il en retirerait s'il prenait cette décision. Cela semble évident que différents individus ont différentes croyances. Par conséquent il n'y a pas un seul type de comportement rationnel. C'est pourquoi la rationalité n'est pas très facile à mettre en œuvre pour décrire le comportement de choix individuels.

Un autre concept décrit dans la littérature est l'opposé de l'impulsivité : un individu ne prendra pas une décision basée sur son état psychologique. Il suivra un processus cohérent et calculé, ce qui ne veut pas dire que cela ne répondra pas à ses propres objectifs.

La rationalité est un concept ambigu et nécessite donc d'introduire un jeu de règles capables de la représenter.

Pour simplifier, il suffit de faire l'hypothèse que pour un preneur de décision si l'alternative A est plus rentable que l'alternative B , il choisira tout le temps A .

III.3 Le modèle de choix discrets

III.3.1 Théorie

Le modèle *logit* peut être appliqué à la théorie économique des comportements individuels en posant les hypothèses qui suivent.

Soit un individu i pouvant choisir entre deux options possibles, notées 0 et 1. Sa décision s'appuie sur le modèle sous-jacent suivant, appelé modèle d'utilité stochastique additive (en anglais : *additive random utility model* – **ARUM**) : l'utilité U_{ji} qu'il retire (ou retirerait) de l'option j , où j peut prendre la valeur 0 ou 1, est la somme d'une composante *déterministe* V_{ji} et d'une composante aléatoire u_{ji} :

$$U_{ji} = V_{ji} + u_{ji} \quad (\text{III.1})$$

La première est nommée ainsi car elle est entièrement déterminée ou expliquée par un ensemble de caractéristiques individuelles observées et notées x_i : $V_{ji} = V_j(x_i)$. La forme généralement retenue de la fonction $V_j(x)$ est linéaire en x : $V_j(x_i) = x_i\beta_j$.

La seconde composante rassemble les variables inobservées et inobservables qui peuvent jouer sur la décision de l'agent i . Elle est supposée varier de manière aléatoire.

En résumé, l'utilité que i retire(raît) de l'option j s'écrit :

$$U_{ji} = x_i\beta_j + u_{ji} \quad (\text{III.2})$$

La règle de décision est alors la suivante : l'individu choisit une des deux options si l'utilité qu'il en retire est supérieure à l'utilité attendue de l'autre option. Si y est la variable binaire repérant l'option choisie (i.e. $y_i = 0$ si l'individu i a choisi l'option 0, et

$y_i = 1$ si i a choisi 1), alors :

$$\begin{cases} y_i = 0 & \iff U_{0i} > U_{1i} \\ y_i = 1 & \iff U_{1i} > U_{0i} \end{cases} \quad (\text{III.3})$$

Ce cadre théorique (annexe C.5, page 103) est adapté aux cas où les deux options entre lesquelles l'individu doit trancher ont elles-mêmes des caractéristiques qui font partie des critères de décision.

Le cas où des caractéristiques des options entrent dans les critères de décision, se formalise de la manière ci-après. Soit z_{ji} , pour $j = 0$ ou 1 , les caractéristiques pour l'individu i . En supposant qu'elles agissent de manière additive sur l'utilité, l'expression (III.2) devient :

$$U_{ji} = x_i\beta_j + z_{ji}\gamma + u_{ji} \quad (\text{III.4})$$

En appliquant toujours la même règle de décision (III.3), la probabilité de prendre l'option 1 s'écrit maintenant :

$$P(y_i = 1|x_i, z_{0i}, z_{1i}) = G[x_i(\beta_1 - \beta_0) + (z_{1i} - z_{0i})\gamma] = \frac{1}{1 + e^{-x_i\beta - (z_{1i} - z_{0i})\gamma}} \quad (\text{III.5})$$

Formalisé ainsi, le modèle est parfois appelé modèle logit conditionnel selon la dénomination que lui a donné McFadden (1973) (conditional logit model). Il contient en effet, en plus des caractéristiques individuelles \mathbf{X} , des variables - les z_j - qui varient avec l'option proposée.

De plus, et surtout, les z_j sont les variables du modèle à privilégier dans l'analyse. Les caractéristiques individuelles sont introduites d'abord pour prendre en compte l'hétérogénéité observée des individus. L'utilisation de ce type de modèle est d'estimer l'impact sur les comportements d'une modification des z_j .

III.3.2 Évaluation

Le cadre méthodologique de la construction de la loi de rachat conjoncturelle repose sur l'aptitude d'un modèle à discriminer de manière satisfaisante les observations.

Cette étape nécessite l'utilisation d'un critère de classification permettant de définir une valeur en-deçà ou au-delà de laquelle un évènement se produit. Le lecteur pourra se reporter à l'annexe D.1, page 109, pour trouver les différentes manières d'évaluer l'efficacité d'un modèle de régression binomiale.

Le modèle à choix discrets présente le même outil que le modèle logistique permettant de détecter efficacement la survenance d'un évènement précis. La courbe **ROC** (annexe D.2, page 112) est ainsi un graphique représentant les performances d'un modèle de classification pour tous les seuils de classification.

Une méthode plus avancée repose sur la maximisation d'un indice synthétique représentant l'aire sous la courbe **ROC** qui pourrait déterminer la meilleure composition de la variable *delta taux*.

III.4 Application

Dans le cas de l'étude sur les comportements de rachats, les caractéristiques individuelles \mathbf{X} sont les variables de l'âge, de l'ancienneté, *etc.* énoncées plus haut et les caractéristiques économiques les z_j . La variable z_0 représente la valeur du TME et z_1 celle du taux servi $N - 1$.

En définitive, la régression logistique comprend les covariables structurelles âge, encours, ancienneté, *etc.* et une unique covariable conjoncturelle *delta taux*. Or cette dernière demeure la plus sensible car ses variations influencent le taux de rachat total.

Par conséquent, il semble raisonnable de privilégier dans un premier temps la construction de cette variable dans l'objectif de déterminer quelle différence de taux pourrait mettre en évidence la prise de décision du rachat.

Afin de construire la variable *delta taux* qui reflétera au mieux le comportement des individus, la méthode utilisée est définie à partir de la qualité du modèle. La rationalité est différente d'un individu à l'autre, ce qui entraîne des décalages entre la prise de décision et celle qui aurait dû être observée au moment opportun.

Il faut par conséquent déterminer une variable *delta taux* qui tentera de rassembler tous ces comportements plus ou moins rationnels en fonction du contexte économique.

Loi conjoncturelle : partie droite

Ce chapitre présente la méthode qui a permis de déterminer les valeurs de *delta taux* en classant de la meilleure manière les observations de la décision de racheter.

À partir de ce résultat, des modèles complémentaires ont permis d'obtenir les paramètres restant de la partie droite de la loi conjoncturelle.

IV.1 Estimation de γ

La classification des rachats en fonction de *delta taux* est le point d'attention de la modélisation. Étant donné qu'aucune valeur observée ne s'est vraiment démarquée, le principe est alors de tester toutes les combinaisons de *delta taux* à partir des quatre valeurs du TME (hypothèse de décision dans les quatre mois) dont les poids varient par pas de 5.

Cela représente 1 771 combinaisons. Le but est de maximiser la valeur de l'**AUC** en trouvant la combinaison *delta taux* qui conduira le modèle à la meilleure classification.

Une première approche aurait pu être la construction de la variable *delta taux* en maximisant simplement le coefficient de corrélation de Spearman. Le résultat est le suivant :

$$R - (\mathbf{0}, \mathbf{10} \times TME_{M-1} + \mathbf{0}, \mathbf{15} \times TME_{M-2} + \mathbf{0} \times TME_{M-3} + \mathbf{0}, \mathbf{75} \times TME_{M-4})$$

pour une valeur de -68% . Elle est supérieure à celles trouvées avec les TME bruts au tableau III.1, page 33, où la valeur maximale était de $-60,85\%$ pour le TME $M - 1$.

Mais cette méthode ne repose pas sur une discrimination des comportements. Ce résultat montre néanmoins que la piste d'une combinaison est opportune.

IV.1.1 Résultats du modèle de choix discrets

Variable *delta taux*

La maximisation de l'aire sous la courbe **ROC** a nécessité de trouver un moyen d'estimer sa valeur lors de chaque combinaison obtenue de *delta taux*.

La première approche est une approximation à partir de la courbe empirique (annexe D.2.1, page 114) et repose sur la méthode binomiale. Or la contrainte de la pondération des observations par les montants d'encours a conduit vers une autre solution. Chevalier (2013) présente une mesure équivalente à la valeur de l'**AUC** à partir de la statistique **c** (annexe D.2.2, page 117) en faisant intervenir les paires d'observations concordantes et discordantes.

La valeur construite s'élève à 0,5813. La combinaison est la suivante :

$$R - (0,50 \times TME_{M-1} + 0,15 \times TME_{M-2} + 0,10 \times TME_{M-3} + 0,25 \times TME_{M-4})$$

Cette tendance de l'environnement économique sur 4 mois ainsi déterminée mène à la meilleure classification des adhérents.

Point d'attention

Avant d'interpréter les résultats de la régression logistique, il serait judicieux de s'assurer de quelques points sur la construction de la variable *delta taux*.

Les graphiques en annexe A.1 et A.2, page 80, démontrent qu'il n'y a pas d'accumulation des observations pour des valeurs faibles de *delta taux*, ni de mois particulier, qui expliqueraient la tendance de la courbe des taux de rachat.

En d'autres termes, la forme de la courbe n'est pas expliquée par une sur-exposition de certains mois.

Représentation

La figure IV.1, de la présente page, met en évidence le taux de rachat en fonction de la nouvelle variable *delta taux*, sous une forme agrégée, à partir des données observées. Elle présente un palier de taux de rachat élevé puis une décroissance au fur et à mesure des valeurs de *delta taux*. Ceci est la meilleure discrimination possible sur la période retenue 2010-2016.

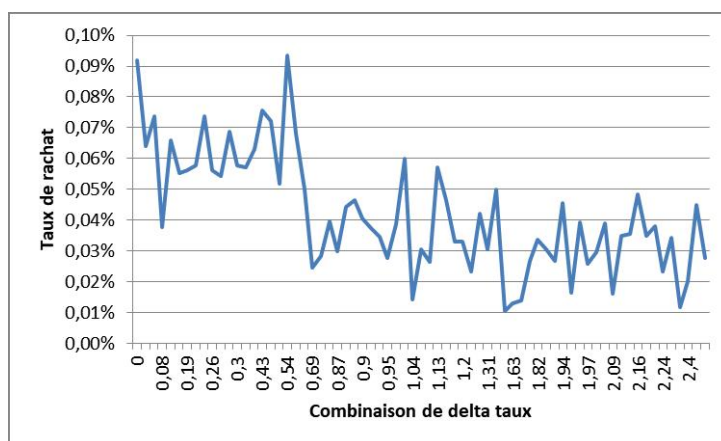


FIGURE IV.1 – Taux de rachat par *delta taux*

IV.1.2 Validation de l'effet additif

L'objectif est à présent de savoir si le modèle complet, contenant les variables structurales et la variable conjoncturelle ainsi modélisée, est statistiquement significatif. Le taux de rachat est, d'après la sélection séquentielle présentée au tableau IV.1, page suivante, expliqué par ses deux natures.

Les résultats détaillés sont présentés en annexe D.3, page 119. Ils démontrent un résultat important : plus l'écart entre le taux servi et le taux attendu est important moins les individus sont enclin à exercer leur option. Cette interprétation conforte le fait

Etape	Saisi	DDL	Nombre dans	Khi-2 du score	Pr > Khi-2
1	<i>delta taux</i>	1	1	224,2290	<,0001
2	lib_anc	5	2	109,1882	<,0001
3	lib_age	4	3	94,7110	<,0001
4	top_tmg	2	4	22,5874	<,0001

DDL : degrés de liberté

Nombre dans : ordre dans lequel la variable est entrée

TABLE IV.1 – Récapitulatif sur la sélection séquentielle dans le modèle à choix discret pondéré

d’avoir fait l’hypothèse forte que les adhérents seraient sensibles à la conjoncture à partir d’une combinaison du TME.

IV.1.3 Seuil optimal

Le fait d’avoir choisi de maximiser l’**AUC**, conduit à identifier, de la meilleure manière possible, les adhérents qui ont le plus de chances de prendre la décision de racheter, connaissant le contexte conjoncturel représenté par la valeur de la variable *delta taux*.

Sachant que l’aire sous la courbe **ROC** peut, d’une autre manière, être approximée par la méthode des trapèzes c’est-à-dire :

$$\begin{aligned}
 \text{AUC} &= \sum_{i=0}^n \frac{(y_{i+1} + y_i) \times (x_{i+1} - x_i)}{2} \\
 &= \frac{TV\text{P} - TFP - 1}{2} \\
 &= \frac{S_e - (1 - S_p) - 1}{2} \\
 &= \frac{S_e + S_p}{2}
 \end{aligned} \tag{IV.1}$$

Maximiser l’**AUC** est alors équivalent à maximiser la somme de la sensibilité et de la spécificité c’est-à-dire $S_e + S_p$. Cette maximisation correspond à l’étape de détermination des coefficients de la variable *delta taux* lors de son estimation par la valeur \mathbf{c} .

Il est dans ce cas possible d’estimer la valeur du paramètre γ en maximisant l’indice de Youden (1950) (annexe E, page 123) et ainsi de trouver la valeur du seuil qui discriminera le mieux les individus. La valeur de *delta taux* est de **0,63**. Cela signifie qu’à partir de ce

seuil, le modèle considère que les adhérents ont moins de chances de racheter leur contrat d'épargne.

L'hypothèse est alors de prétendre qu'à partir de la valeur 0 jusqu'à la valeur dans ce cas 0,63, la courbe conjoncturelle n'a pas d'effet sur les taux de rachats structurels. Sur cet intervalle, les adhérents sont dès lors indifférents à l'écart entre le taux servi et le taux proposé par le marché.

Il convient désormais de calculer les valeurs de δ et RC_{Min} . Cependant toutes les informations issues de la discrimination par la courbe **ROC** ont été extraites.

IV.2 Estimation de δ

La détermination de ce paramètre requiert de poser le problème d'une autre manière en considérant les observations de manière agrégée.

IV.2.1 Lissage

L'observation agrégée des données permet de construire des taux de rachats en montant en fonction uniquement de la variable conjoncturelle *delta taux*.

Hypothèse

Le principe qui va être utilisé repose sur l'idée de l'existence d'une tendance qui pourrait être extraite à partir des méthodes de lissage. Le paramètre du déclenchement de la baisse des rachats étant estimé avec le modèle de choix discrets, il reste à déterminer le seuil à partir duquel les adhérents ne sont plus sensibles aux variations de *delta taux*. La question repose sur l'existence d'une valeur de cet indicateur pour laquelle la courbe des taux de rachats conjoncturels se stabiliserait.

Il est présenté à la figure IV.2, page suivante, les résultats des lissages par les méthodes B-spline et Loess. L'observation montre que les lissages ne tiennent pas compte d'un effet sous-jacent qui pourrait définir la perturbation d'une origine structurelle.

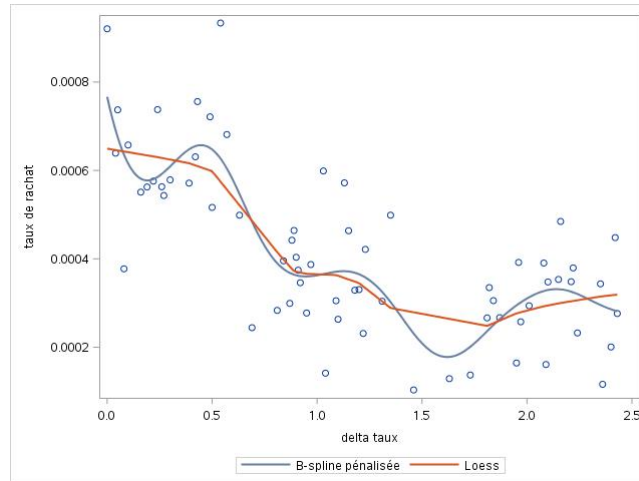


FIGURE IV.2 – Lissage des taux de rachat

Problème

L'utilisation des méthodes de lissage (annexe F, page 127) semblait être une bonne approche dans la modélisation de la tendance. Cependant aucune solution n'a été trouvée au problème d'estimation de la valeur de δ .

Une réponse a été obtenue grâce au fait que ces données longitudinales ne respectent pas l'hypothèse d'indépendance des observations. Il est donc indispensable d'introduire une structure de covariance pour les variables aléatoires associées à chaque adhérent.

IV.2.2 Modèle mixte

Le modèle linéaire à effets mixtes est une extension du modèle linéaire qui prend en compte la variabilité liée aux individus. Ce modèle est composé d'une partie fixe et d'une partie aléatoire. La partie fixe est identique pour chaque individu et représente l'effet population. La partie aléatoire est propre à chacun des individus et traduit la variabilité liée à chaque sujet.

Interprétation

L'idée est d'extraire le facteur fixe représenté par *delta taux* qui traduirait une tendance (effet de la conjoncture, de la rationalité), et d'expliquer par des effets aléatoires c'est-à-dire les caractéristiques individuelles ou contractuelles (effet structurel) les bruits autour de cette tendance.

Les modèles permettant de manipuler les données corrélées appartiennent à la famille des modèles linéaires mixtes (annexe G, page 131). Le facteur à effets fixes sera alors *delta taux* et les facteurs à effets aléatoires représenteront toutes les caractéristiques des individus et du contrat observables ou non.

Résultat

Le résultat de l'estimation par le modèle mixte est exposé sur le graphique IV.3, de la présente page. Il montre une tendance à la baisse des taux de rachats lorsque les valeurs de *delta taux* augmentent. La sélection des nœuds (annexe G.2, page 134) a permis de garder la valeur de $\gamma = \mathbf{0,63}$ déterminante pour le déclenchement de la baisse effective des rachats.

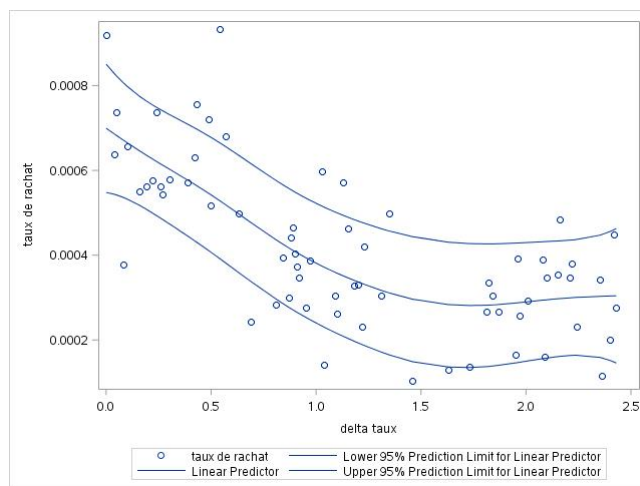


FIGURE IV.3 – Lissage des taux de rachat en fonction de *delta taux* par le modèle mixte

Ceci permet de prendre en considération l'effet d'atténuation des taux de rachats. La dispersion des observations révélant le comportement aléatoire des adhérents conduit à un intervalle de confiance assez éloigné de la courbe lissée.

Le comportement des adhérents est supposé non linéaire comme vu précédemment dans la détermination de *delta taux*. La courbe de rachat dynamique proposée par l'ACPR est plane à partir de la valeur β .

Problème

Or le graphique IV.3 ne montre pas clairement cet effet. Il faut alors avoir recours à un problème d'optimisation dans le but de déterminer à partir de quelle valeur de *delta taux* il est possible de constater une indifférence à l'augmentation des taux.

Ceci consiste à résoudre analytiquement ou numériquement un problème de minimisation de la fonction sur les valeurs de *delta taux* supérieures à 0,63.

IV.3 Estimation de RC_{min}

Hypothèse

La convergence vers un plateau d'indifférence à la baisse des taux à partir d'un point est modélisée à l'aide d'un modèle quadratique avec palier (annexe H, page 135). Ce point détermine le seuil de stabilisation de la baisse des rachats conjoncturels. Il est possible d'écrire un tel modèle sous la forme suivante :

$$E[Y|x] = \begin{cases} \alpha + \beta x + \gamma x^2 & \text{si } x < x_0 \\ c & \text{si } x \geq x_0 \end{cases} \quad (\text{IV.2})$$

Dans l'équation du modèle, les paramètres α , β et γ sont les coefficients du segment quadratique et la valeur c définit le plateau de la fonction moyenne.

L'idée d'opter pour ce type de modèle plutôt que pour une estimation de la pente par un modèle linéaire repose sur l'objectif de la découverte d'un seuil à partir duquel les valeurs de *delta taux* n'ont plus d'effet sur la décision des rachats. Du plus il conforte l'hypothèse d'un taux de rachat conjoncturel unique lorsque les valeurs de *delta taux* dépassent un certain seuil.

Résultat

La figure IV.4, de la présente page, permet d'observer le résultat du problème de minimisation. Sa structure suit parfaitement la courbe lissée puisqu'elle a une forme quadratique et définit un plateau à partir de la valeur $\gamma=1,68\%$. Les deux seuils γ et δ sont par conséquent calculés. Il ne reste désormais que la valeur de RC_{min} à obtenir.

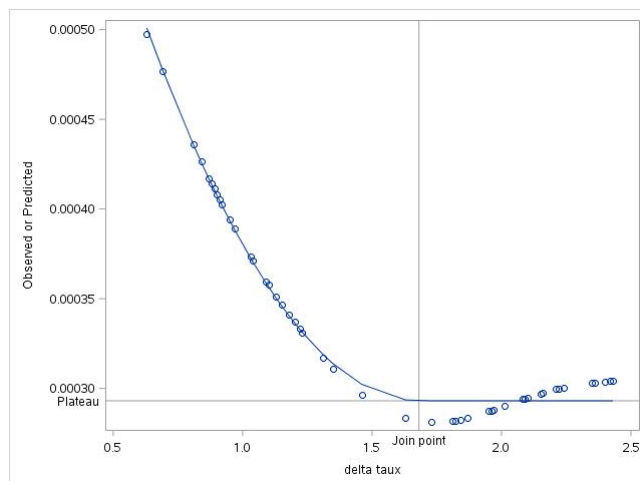


FIGURE IV.4 – Détermination du seuil γ à partir duquel les rachats conjoncturels sont stables

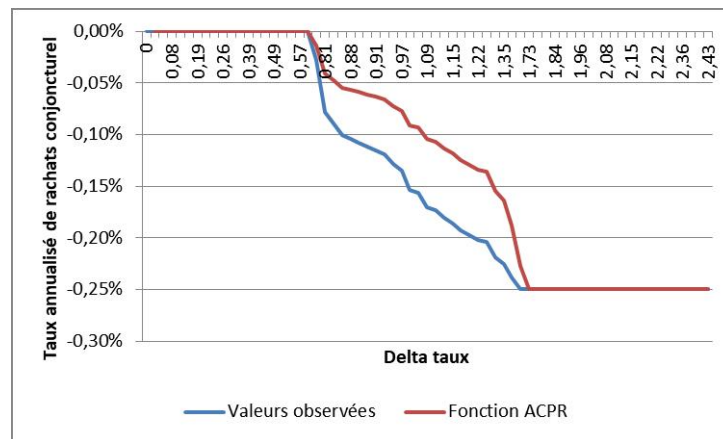
Le principe repose alors sur la différence entre le taux de rachat observé en γ et celui observé en δ . Cette valeur est annualisée pour être exploitée dans le modèle de projection du département ALM.

La valeur de RC_{min} est de **0,25%**. Elle est très faible par rapport à celle proposée par l'ACPR, qui était de 6%. Mais elle reflète davantage le comportement des taux de rachats de la Mutuelle qui sont bien en deçà de la valeur de la fonction *Min*.

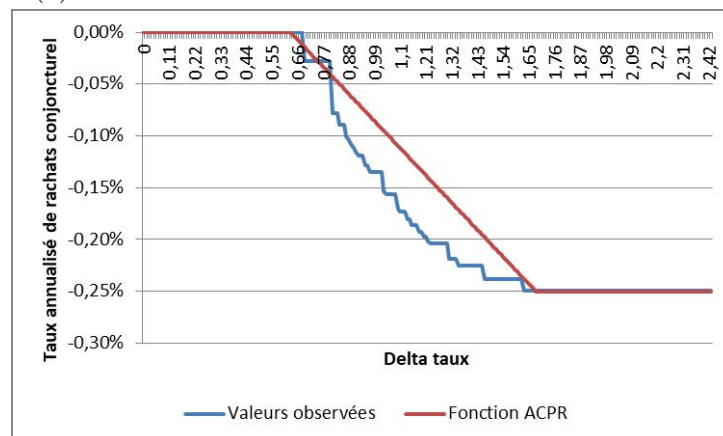
IV.4 Conclusion

La formule de la courbe proposée par l'ACPR est linéaire alors que la tendance observée se présente sous une forme non monotone. Le graphique IV.5, de la présente page, détaille les prédictions observées contre les valeurs des paramètres dans la loi ACPR qui est une approximation linéaire. La différence de leur forme repose sur le fait que sa construction repose sur peu d'observations.

L'écart entre les deux courbes est supposé non significatif à la vue de la valeur de RC_{min} . De plus le but est de s'appropriier les paramètres de la fonction ACPR et non de modifier son incidence sur le taux de rachat structurel.



(a) Formule ACPR et LFM sur les données observées



(b) Formule ACPR et LFM sur des données complètes

FIGURE IV.5 – Taux annualisés des rachats conjoncturels observés et formule ACPR en fonction de delta taux

Loi conjoncturelle : partie gauche

Il existe très peu d'observations de la chronique de *delta taux* présentant des valeurs négatives. Le problème de n'avoir pas pu être confronté à un environnement de taux obligataires élevés conduit, par conséquent, à se reposer sur des simulations de contextes économiques dans lesquels un concurrent viendrait proposer un taux de rémunération plus attractif à l'adhérent de la Mutuelle.

L'idée est tirée de la modélisation des remboursements de prêts octroyés par les établissements bancaires ou organismes de crédit. Elle est basée sur le gain réalisé si le prêt est renégocié à un taux plus bas.

Le scénario retenu est pris à l'envers pour étudier le gain réalisé au terme d'un arbitrage total du contrat d'épargne primitif à un taux plus favorable. À partir de ce jeu de données théoriques il est possible d'appliquer le modèle de choix discrets et de déterminer à partir de la courbe **ROC** le premier seuil de déclenchement des rachats conjoncturels.

Les individus seront enfin définis comme insensibles aux variations de *delta taux* si par hypothèse 9 personnes sur 10 auraient dénoué leur contrat initial.

V.1 Modélisation des remboursements de prêt

V.1.1 Hypothèses

La modélisation des remboursements de prêts hypothécaires repose sur les principales hypothèses suivantes. Un créancier s'engage à payer un débiteur par des sommes mensuelles égales jusqu'au terme du prêt. Chaque contrat est caractérisé par le fait que le créancier a le droit de rembourser par anticipation à tout instant.

De plus, cette option est exercée sans pénalité financière. Le crédit est remboursé en raison de la vente du bien sous-jacent ou d'un refinancement en un autre prêt.

Parmi les incitations financières qui contribuent au remboursement anticipé, la plus significative est celle due à un taux d'intérêt plus faible. Cependant le remboursement comporte des coûts qui peuvent décourager le créancier de l'exercer.

V.1.2 Adaptation au rachat d'un contrat d'épargne

L'adhérent en possession d'un contrat d'épargne réfléchit de la manière opposée au détenteur d'un prêt hypothécaire. La proposition d'un taux plus élevé vient inciter celui-ci à racheter son contrat dans l'espoir d'un meilleur rendement.

Toutefois le contrat d'épargne est quant à lui lié à la fiscalité de son ancienneté. Plus l'adhérent possédera longtemps son contrat plus il pourra maximiser les gains de l'accumulation de son épargne grâce aux conséquences de l'ancienneté fiscale.

Le principe de la simulation des données réside dès lors dans la modélisation de la fiscalité du contrat, qui peut, sachant le taux du prélèvement forfaitaire libératoire, contraindre l'adhérent à ne pas racheter son contrat.

Il est également supposé que la décision repose uniquement sur un choix rationnel de maximisation de son utilité à changer de contrat pour un gain certain.

V.2 Modélisation des rachats

La simulation d'un contrat d'épargne reposant uniquement sur un fonds Euro a été réalisée dans le but de connaître l'impact de la proposition d'un taux de rendement plus élevé, sur la stratégie d'investissement. Elle est réalisée dans un environnement déterministe et non stochastique par simplicité.

V.2.1 Le contrat alternatif

L'objectif est de projeter un contrat initial C_0 caractérisé par un taux d'intérêt garanti i_0 sur un horizon de placement T défini en $t = 0$.

L'adhérent a ensuite la possibilité de racheter son contrat avec des pénalités fiscales et contractuelles à un temps t compris entre $t = 1$ et $t = T - 1$ par un contrat garantissant un taux d'intérêt garanti i_1 tel que $i_1 > i_0$.

L'horizon de placement T est compris entre 2 et 15 ans et est commun aux contrats C_0 et C_1 . La figure V.1, de la présente page, démontre le processus de rachat du contrat initial.

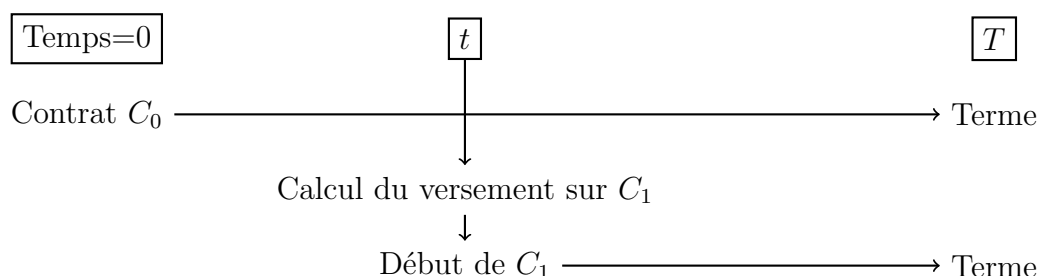


FIGURE V.1 – Simulation d'un rachat du contrat initial en t

Il s'avère que ce genre de contrat alternatif est difficile à trouver sur le marché bien que le contrat C_0 existe déjà. Cela concerne certains contrats avec une garantie de taux portée par un ou plusieurs versements réalisés par le passé.

Les valeurs de *delta taux* sont représentées par la différence $i_0 - i_1$. Le taux i_0 varie entre 0% et 5% par pas de 0,1%. Cette hypothèse repose sur les valeurs des taux techniques

de tarification observés au sein du portefeuille. Les valeurs simulées de *delta taux* varient de 0% à 8% par pas de 0,1%. L'étude repose sur le nombre de 348 705 simulations.

V.2.2 Frais d'entrée

Les frais d'entrée sur le contrat C_0 et le contrat concurrent C_1 sont identiques et définis par la valeur g sans rapport avec les sommes investies. Ils représentent ceux pratiqués par la Mutuelle.

L'hypothèse s'appuie sur le fait que la compagnie concurrente prélèverait autant de chargement que la Mutuelle. Elle paraît raisonnable et prudente puisque la Mutuelle propose des tarifs dans la fourchette basse du marché.

De plus ces frais d'entrée sur le contrat alternatif pourraient être considérés comme le coût de la recherche d'un meilleur rendement.

V.2.3 Fiscalité

Les pénalités de rachat représenteront uniquement le régime fiscal de l'assurance vie. Ce n'est qu'à l'occasion d'un retrait (rachat) définitif, total ou partiel, que les intérêts sont taxés, en fonction de la date du retrait et du choix de l'épargnant.

Les coûts de rachat sont modélisés selon la fiscalité en vigueur en 2016 sans la prise en compte des impacts de la loi *Flat Tax*. Il est considéré que ces effets mettront certainement quelques années avant d'être observés.

Dans le cadre d'un contrat d'assurance-vie, les produits du capital placé sont imposables. Deux possibilités sont offertes à l'adhérent en cas de rachat de son contrat afin de déclarer les intérêts : soit intégrer les gains au barème de l'impôt sur le revenu, soit opter pour le prélèvement forfaitaire libératoire.

Dans le but de rester dans un cadre simple et peut-être général, seuls les prélèvements forfaitaires libératoires sans abattement seront intégrés dans le calcul du montant des gains.

Les prélèvements sociaux alors notés ps sont prélevés sur les intérêts attribués au contrat par les taux i_0 et i_1 selon l'ancienneté révolue suivante :

- inférieure à 3 ans : 35%
- entre 4 et 7 ans : 15%
- plus de 8 ans : 7,5%

V.2.4 Placements investis

Le montant investi varie dans un ensemble de cinq valeurs prises arbitrairement $\{100, 300, 1\,000, 15\,000, 30\,000\}$ noté V_0^0 .

La deuxième valeur correspond par exemple au montant minimum de versement sur un des contrats d'épargne de la Mutuelle. Les trois dernières correspondent aux valeurs de quantile observées dans le portefeuille.

La valeur de 100€ a été retenue pour obtenir une sensibilité fine.

V.2.5 Chronologie du rachat

La figure V.2 de la présente page reprend la même chronologie que celle exposée à la figure V.1 page 53, mais intègre les notions mathématiques des valeurs des contrats à différents moments de la simulation.

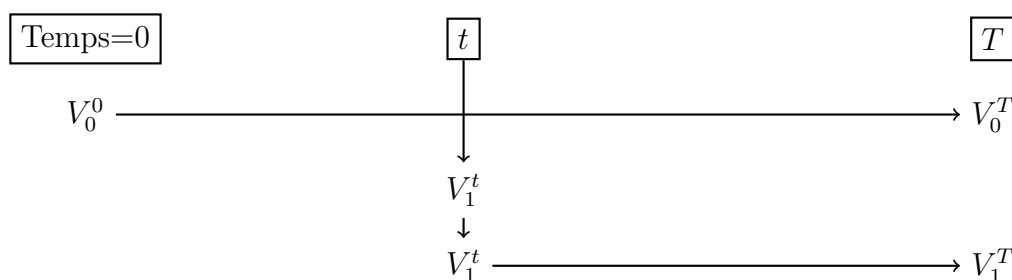


FIGURE V.2 – Simulation d'un rachat du contrat initial en t avec les notions mathématiques

V.2.6 Valeur du rachat

La valeur au terme T de la simulation i du contrat d'épargne C_0 est la suivante :

$$V_0^T = V_0^0 \times (1 - g) \times \left(1 + i_0 \times (1 - ps_0^T)\right)^T$$

La valeur au temps t du placement dans le nouveau contrat C_1 est la suivante :

$$V_1^t = V_0^0 \times (1 - g)^2 \times \left(1 + i_0 \times (1 - ps_0^t)\right)^t$$

La valeur terminale au temps T du contrat d'épargne C_1 est la suivante :

$$V_1^T = V_1^t \times \left(1 + i_1 \times (1 - ps_1^{T-t})\right)^{T-t}$$

Finalement le résultat au temps T de la décision de rachat est la suivante :

$$\text{si } V_1^T + 1 > V_0^T \text{ alors rachat}$$

L'adhérent rachète son contrat pour un gain supérieur à 1€ qui se présente comme l'hypothèse la plus prudente.

V.3 Résultats intermédiaires

La simulation pour la valeur de 100€ est représentée sur la figure V.3, page suivante. Le temps de retour sur investissement décroît au fur et à mesure que *delta taux* décroît, c'est-à-dire i_1 augmente par rapport à i_0 .

En s'appuyant sur la forme de la courbe proposée par l'ACPR, un premier palier peut être observé pour lequel les adhérents seraient indifférents à une modification des taux proposés jusqu'à 2%.

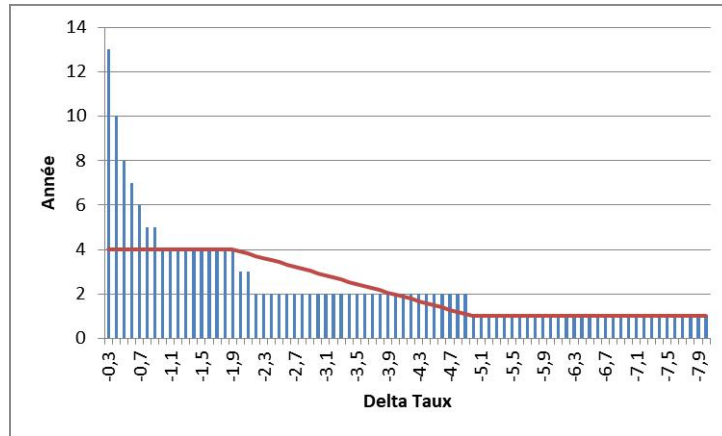


FIGURE V.3 – Temps de retour sur investissement d’une valeur de 100€ en fonction de *delta taux*

Ensuite la hausse des rachats conjoncturels débiterait à partir de cette valeur jusqu’à celle de 5% où tous les adhérents qui souhaiteraient racheter leur contrat exerceraient leur option.

V.4 Retour sur le modèle de choix discrets

Le critère du temps de retour sur investissement est une première approche de l’impact de la hausse des taux sur les rachats. Toutefois il peut être intéressant d’utiliser toutes ces simulations dans un modèle à choix discrets.

V.4.1 Coefficient de corrélation de Matthews

En effet une variable modélisant le rachat peut être écrite de façon binaire à la manière du chapitre précédent. La variable réponse se présente alors sous la forme suivante :

$$y_i = \begin{cases} 1 & \text{si } V_{i1}^T + 1 > V_{i0}^T \\ 0 & \text{sinon} \end{cases} \quad (\text{V.1})$$

Problème du choix de la mesure de qualité

La construction du modèle ne reposant pas cette fois-ci sur la maximisation du l'**AUC**, d'autres critères statistiques sont disponibles quant à la détermination du seuil optimal.

Pour rappel, les catégories de la matrice de confusion sont les suivantes : vrais négatifs (TN), vrais positifs (TP), faux positifs (FP) et faux négatifs (FN).

Le critère de la précision (Accuracy) s'écrit :

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (\text{V.2})$$

Le critère du score F1 (F1 score) s'écrit :

$$\text{F1 score} = \frac{2 \times TP}{2 \times TP + FP + FN} \quad (\text{V.3})$$

Cependant, même si ces deux critères sont largement utilisés en statistiques, tous deux sont trompeurs car ils ne considèrent pas complètement la taille des quatre classes de la matrice de confusion dans leur élaboration.

Une solution

Une mesure de performance proposée est le coefficient de corrélation de Matthews (1975). Le coefficient MCC peut être vu comme la discrétisation du coefficient de corrélation de Pearson pour les variables binaires. Cette mesure est largement utilisée en Bio-informatique comme une mesure de performance.

Il est calculé à partir des valeurs de la matrice de confusion :

$$\text{MCC} = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP) \times (TP + FN) \times (TN + FP) \times (TN + FN)}} \quad (\text{V.4})$$

En considérant la proportion de chaque classe de la matrice de confusion dans cette formule, ce score est élevé seulement si le classifieur fonctionne bien sur les éléments négatifs et positifs.

D'autre part, vérifier le coefficient de corrélation de Matthews permet de constater que le modèle n'attribue pas aléatoirement les probabilités. Il alerte l'utilisateur sur le mauvais fonctionnement de son modèle.

Résultat

Dès les scénarios des versements supérieurs ou égaux à 300€ la valeur du coefficient MCC est de 53%. Le seuil correspondant s'élève quant à lui à 1,20. Le coefficient de la précision donne un seuil de 1 et le score F1 de 1,10. Pour indication, le calcul de l'indice de Youden fournit une valeur de 3,10 pour *delta taux*.

La valeur **1,20%** est donc retenue pour le paramètre β .

V.4.2 Probabilité estimée et seuil optimum

Le seuil à partir duquel les rachats conjoncturels cessent de s'ajouter aux rachats structurels est déterminé à partir des probabilités estimées des différents scénarios modélisés.

Scénarios

Le tableau V.1, de la présente page, montre par exemple pour le scénario 100€ qu'il y a une chance sur deux que le rachat ait lieu pour une différence de taux atteignant 1,90.

	Sc 100€	Sc 300€	Sc 1k€	Sc 15k€	Sc 30k€
$\hat{\pi} = 50\%$	1,90	1,70	1,60	1,60	1,60
$\hat{\pi} = 75\%$	3,40	3,00	2,80	2,80	2,80
$\hat{\pi} = 90\%$	4,90	4,30	4,10	4,00	4,00
$\hat{\pi} = 95\%$	5,90	5,10	4,90	4,80	4,80

TABLE V.1 – Seuils optimums selon les divers scénarios en fonction de la prévalence de l'événement de rachat

La détermination du seuil α peut reposer sur l'hypothèse $\hat{\pi} > 95\%$. Ainsi 95% des observations ont une prédiction de l'événement de rachat. Les 5% restant sont caracté-

sés par un ensemble d'individus indifférents à la présence de nouveaux concurrents plus attractifs sur le marché.

La moyenne des quatre derniers scénarios, c'est-à-dire de 300€ à 30k€, donne une valeur de **4,90** qui est aussi la valeur à 90% du scénario de sensibilité à 100€.

Conclusion

Par conséquent si plus 9 individus sur 10 sont prêts à racheter leur contrat pour un écart de taux de 4,90 points et pour une mise de 100€ alors par hypothèse le dernier prendra inmanquablement cette décision.

Loi structurelle et impacts

Solvabilité II

Ce chapitre traite de la manière dont a été déterminée la loi de rachat conjoncturelle. Il aborde pour terminer les impacts de la nouvelle loi dynamique sur les calculs prudentiels Solvabilité II.

VI.1 Loi structurelle

La loi de rachat dynamique est composée d'une partie structurelle qu'il faut désormais tenter de déterminer.

Elle peut dépendre des caractéristiques du contrat et de l'adhérent comme cela a été démontré dans le modèle de choix discrets complet.

Pour rappel, le taux de rachat observé semble être expliqué par les variables structurelles admises comme l'ancienneté du contrat et l'âge du souscripteur.

VI.1.1 Hypothèse

Le département ALM a mené plusieurs études qui ont néanmoins mis en évidence qu'une loi de rachat par ancienneté n'a qu'un impact négligeable dans le calcul des indi-

cateurs prudents. Un taux unique est donc évalué pour représenter la taux de rachat structurel.

Ce dernier est calculé mensuellement de la manière suivante :

$$\begin{aligned}
 RS(.) &= RT(R, TA, .) - RC(R, TA) \\
 &= RT(R, TA, .) - \text{Provisions Mathématiques} \times \text{dynamique}
 \end{aligned}$$

où la variable *dynamique* représente les valeurs de la loi de rachat conjoncturel.

Cette écriture permet de retirer du taux de rachat observé l'influence de la conjoncture à laquelle est confronté l'adhérent. Le graphique VI.1, de la présente page, fait apparaître la déformation des taux de rachats observés par *delta taux* sur la période étudiée.

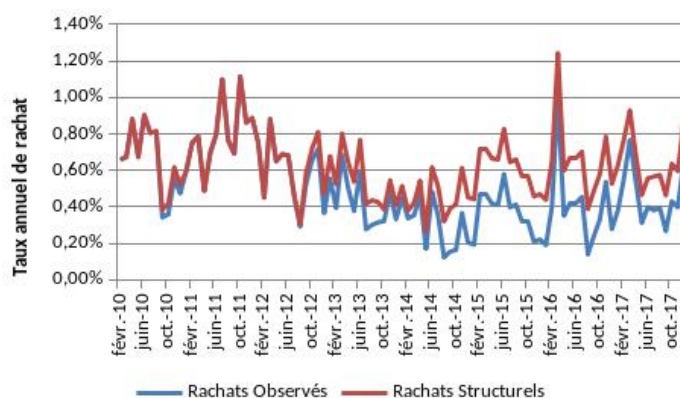


FIGURE VI.1 – Taux de rachat total Euro observé et taux de rachat structurel théorique

VI.1.2 Observation

Les taux de rachat structurel sont plus élevés que les taux observés puisque les valeurs de $R - TA$ n'interviennent que pour atténuer ces derniers. Le graphique VI.2, page ci-contre, rend compte plus exactement de l'influence de la loi conjoncturelle.

La période de 2010-2011 serait caractérisée principalement par des rachats structurels car la variable *delta taux* est la plupart du temps inférieure à 0,63. Les retraits réalisés par les adhérents à cet intervalle de temps n'auraient eu pour objectif que de répondre à un besoin de capital.

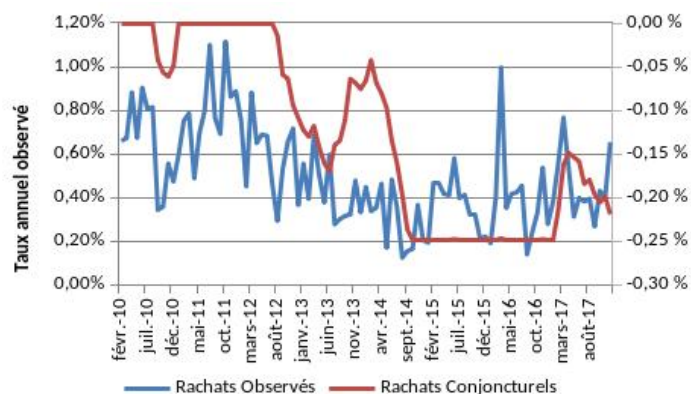


FIGURE VI.2 – Taux de rachat total Euro observé contre taux de rachat conjoncturel

La Mutuelle n’a pas suivi rapidement la baisse des taux obligataires en proposant toujours des taux de participation attractif et a dans ce cas réussi à retenir une partie de son portefeuille. Cet effet est apparent entre octobre 2014 et janvier 2017 puisque la valeur des rachats conjoncturels serait à son minimum : RC_{Min} .

VI.1.3 Justification

L’écriture du rachat structurel par différence peut être considérée comme une hypothèse forte mais toute la difficulté de cette étude a été de travailler avec un portefeuille dont le but est de se constituer un capital à transmettre.

Pour appuyer cette hypothèse les valeurs du coefficient de corrélation avec le taux de rachat observé sont exposées dans le tableau VI.1, de la présente page.

Rachats	2010-2017	2015-2017
Structurels	89,46%	98,02%
Conjoncturels	60,45%	20,03%

TABLE VI.1 – Coefficients de corrélation avec le taux de rachat observé

Le taux de rachats structurels est très fortement corrélés au taux de rachat observé, encore plus ces trois dernières années. Il semble que la principale raison de rachat soit un besoin de liquidité.

Il serait alors possible de formuler le taux de rachat structurel par différence entre le taux observé et le taux de rachat conjoncturel.

VI.1.4 Évaluation

L'historique employé pour la construction des lois de rachat au sein de la Mutuelle est de trois années. Le graphique VI.3, de la présente page, met en évidence les taux de rachats représentant l'ensemble des natures de rachats observées (totaux plus partiels).

Il démontre une stabilité au moins ces trois dernières années. L'hypothèse d'un historique de trois ans semble raisonnable. La construction de la loi structurelle est basée sur les règlements des années 2015, 2016 et 2017 en y incorporant le montant des provisions de sinistres à payer constaté au 31 décembre 2017.

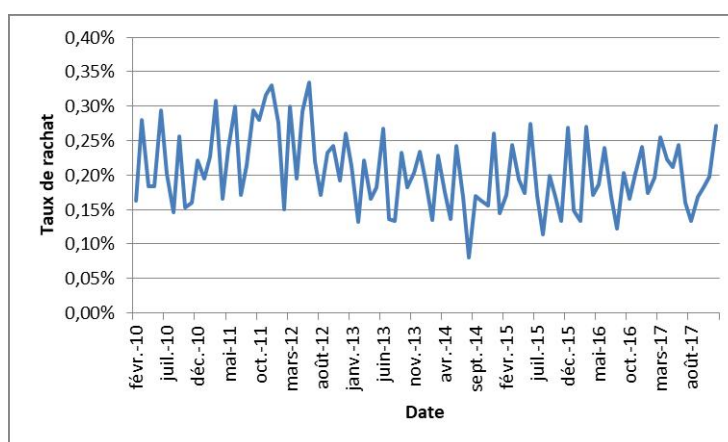


FIGURE VI.3 – Taux de rachat global observé de 2010 à 2017

Le calcul des taux de rachats partiels Euro/UC et totaux UC est le même que celui employé dans la construction de la loi conjoncturelle, hormis que les montants sont sommés sur les trois années retenues. La formule du taux de rachat mensuel est la suivante :

$$\text{Taux de rachat mensuel } M = \frac{\text{Montant racheté du mois } M}{\text{Provisions Mathématiques du mois } M - 1}$$

Le modèle de projection des flux futurs repose sur des pas annuels et non mensuels. La transformation du taux mensuel est la suivante :

$$\text{Taux de rachat annuel} = 1 - \prod_{M=1}^{12} \left(1 - \frac{\text{Montants rachetés } M}{\text{Provisions Mathématiques } M-1} \right)$$

VI.2 Calculs prudentiels sous Solvabilité II

Depuis le 1^{er} janvier 2016, la réglementation Solvabilité II impose entre autres aux compagnies d'assurances et aux mutuelles de calculer un taux de couverture en évaluant leurs engagements par une approche économique de leurs provisions techniques.

VI.2.1 La valeur économique

La Mutuelle définit dans ce cadre général la valeur économique (Best Estimate ou **BE**) du portefeuille correspondant à la moyenne pondérée par leur probabilité des flux de trésorerie futurs, compte tenu de la valeur temporelle de l'argent. Cette dernière est estimée sur la base de la courbe des taux sans risque pertinente¹.

Le **BE** est alors la valeur actuelle attendue des flux de trésorerie futurs². La Mutuelle a recours à une modélisation stochastique des risques en prenant en compte les interactions Actif/Passif et les clauses de participation aux bénéfices incluses dans les contrats. Le **BE** requiert nombres d'hypothèses dans le but d'une évaluation adéquate du portefeuille du passif : table de mortalité, politique de participation aux bénéfices/revalorisation, lois de rachat, *etc.*

VI.2.2 Le capital de solvabilité requis

Le Capital de Solvabilité Requis ou SCR (en anglais Solvency Capital Requirement) représente le montant de fonds propres estimé nécessaire pour absorber une perte provoquée par un choc exceptionnel.

Il correspond en d'autres termes au montant de capital minimum dont la Mutuelle doit disposer dans le but de limiter la probabilité de ruine à un an à 0,5%.

1. Il est important de noter que l'utilisation d'une courbe des taux est fondamentale en assurance vie. Ces taux servent de référence dans la projection de l'actif de l'assureur grâce aux générateurs de scénarios économiques, mais également pour l'actualisation des engagements de ce dernier dans le calcul des provisions techniques.

2. Article R. 351-2 du Code des Assurances, transposition en droit français de l'article 77 de la Directive Solvabilité II

Dès lors que l'organisme ne couvrira plus son SCR, le superviseur devra établir un plan de redressement en concertation avec ce dernier. Le SCR est basé sur le profil de risque de l'organisme.

Le capital de solvabilité requis est calculé au sein de la Mutuelle selon l'approche standard.

VI.2.3 Le taux de couverture

Le taux de couverture de la marge de solvabilité est égal au rapport entre la marge de solvabilité effective de l'assureur et l'exigence minimale de marge de solvabilité calculée sur la base de ses engagements et des risques qu'il doit couvrir. Il est défini autrement comme le rapport entre les fonds propres éligibles et le capital de solvabilité requis.

Le ratio de la marge de solvabilité sur le SCR détermine le ratio de solvabilité et doit naturellement être supérieur à 100%, pour satisfaire aux exigences de solvabilité. Autrement dit, les capitaux propres doivent couvrir l'exigence de capital réglementaire.

VI.3 Analyse des variations

Les impacts ont été étudiés à partir d'un scénario central, reposant notamment sur la loi conjoncturelle notée *Min* proposée par l'ACPR.

VI.3.1 Méthodologie

Le scénario central correspondant au portefeuille complet des contrats en cours au 31 décembre 2017 et des hypothèses non économiques au 31 décembre 2016.

La méthode consiste à observer la valeur du SCR à partir d'une analyse des variations de chaque paramètre de la loi dynamique. Pour rappel les éléments de la loi sont présentés dans le tableau VI.2 page suivante.

Loi	α	β	γ	δ	RC_{min}	RC_{max}
Central	-6,00%	-2,00%	1,00%	2,00%	-6,00%	20,00%
LFM	-4,90%	-1,20%	0,63%	1,68%	-0,25%	20,00%

TABLE VI.2 – Loi dynamique ACPR et loi dynamique LFM

Plus exactement, cette analyse consiste à modifier chaque paramètre, dont l'impact souhaite être observé, les uns à la suite des autres. Les quatre grandes étapes se présentent de la manière suivante :

1. Scénario central
2. Impact de la loi structurelle
3. Impact des paramètres déclencheurs α , β , γ et δ
4. Impact RC_{min} puisque RC_{max} n'a pas été modifié

VI.3.2 Résultats

Les impacts de la loi structurelle puis des nouveaux paramètres de la loi conjoncturelle sont présentés dans le tableau VI.3, de la présente page :

Étape détaillée	Taux de couverture
0 - Central	150,54%
1 - Impact loi structurelle	156,14%
2 - Impact RC_{min}	171,37%
3 - Impact γ	171,05%
4 - Impact δ	170,98%
5 - Impact α	170,96%
6 - Impact β	170,57%

TABLE VI.3 – Impacts de la nouvelle loi dynamique sur le taux de couverture au 31 décembre 2017

VI.3.3 Quelques explications

Loi structurelle

L'étape de la loi structurelle a un impact de +5,6 points de pourcentage sur le taux de couverture. Il est expliqué par :

- baisse du BE rachat : la somme des rachats totaux et partiels des hypothèses de 2016 sont supérieurs à ceux modélisés avec la loi construite en 2017 principalement dû au fait de la valeur prise par $RC_{Min} = 6\%$ les années précédentes,
- augmentation du BE mortalité : moins de prestations à payer en début de projection y compris Participation aux Bénéfices (PB). Ceci entraîne par conséquent un surplus de PB à verser pour les contrats restants lors des années suivantes,
- baisse du BE frais : les taux de rachat partiel en 2016 sont en moyenne supérieurs au taux de rachat partiel en 2017.

Ce résultat correspond avec celui démontré dans Mazzanti (2012). L'auteur constate, dans ses travaux réalisés au sein de la Mutuelle, que l'impact d'une modification à la baisse d'un point de pourcentage sur les taux de rachats structurels conduit à une augmentation sensible du taux de couverture.

L'augmentation du SCR marché est dû à la baisse des engagements en portefeuille qui entraîne plus d'exposition aux risques de marché. Le SCR souscription vie diminue via le SCR rachat avec le choc à la hausse. Étant donné que les taux de rachat sont plus bas, un choc de rachat à la hausse a moins d'impact dans ce cas là. Au final le SCR diminue.

Rachat conjoncturel minimal

L'étape de la valeur de RC_{min} a un impact de +15,2 points de pourcentage sur le taux de couverture. Compte tenu du faible niveau des taux et du taux servi, les premières années de projection sont dans une grande majorité des scénarios sur la partie basse du graphe des rachats conjoncturels. En prenant $RC_{min} = -0,25\%$ au lieu de -6% cela fait augmenter les rachats en début de projection. La PB distribuée en début de projection est alors plus importante.

La PB distribuée sur les contrats restants lors des années suivantes devient moins importante, cela explique une baisse du BE de mortalité. La hausse des rachats entraîne une baisse du nombre de contrats restants et donc une baisse du BE de frais. Avec RC_{min} presque nul, le portefeuille est beaucoup moins exposé à une baisse des taux, ce qui entraîne une baisse du SCR.

VI.3.4 Illustration

À titre d'illustration, le graphique VI.4, de la présente page, met en évidence l'évolution sur 50 ans de la différence entre le taux servi R et le taux attendu TA selon la valeur du quantile du *Best Estimate* au 31 décembre 2017.

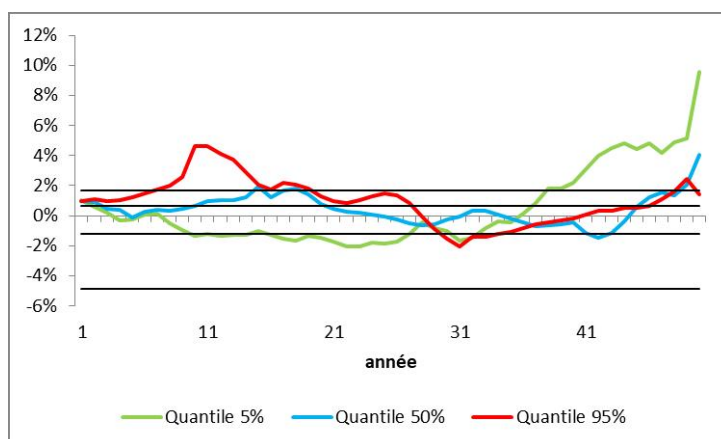


FIGURE VI.4 – Évolution de la différence $R - TA$ selon la valeur du quantile du *Best Estimate* au 31 décembre 2017

Rappel : $\alpha = -4,90\%$, $\beta = -1,20\%$, $\gamma = 0,63\%$ et $\delta = 1,68\%$

Les valeurs des *Best Estimate* ont été classées à la suite de la projection stochastique des hypothèses du scénario central. La courbe correspondant au quantile 95% montre que la différence est plus grande que la valeur δ : les rachats structurels sont diminués au maximum par RC_{min} . Cet évènement se produit très tôt dans la projection, dès 7 ans, ce qui a un impact important sur le *Best Estimate*. Ceci met en évidence l'impact de la valeur de RC_{min} .

Le *Best Estimate* moyen est impacté par la valeur de RC_{min} à partir de la 11^e année pendant 10 ans. Ceci démontre que la moitié des projections est affectée par cette nouvelle valeur.

Les différentes valeurs de $R - TA$ se situent peu au-delà de β et jamais après α . Ceci explique le faible impact sur le taux de couverture lors de l'analyse des changements.

VI.3.5 Résumé

Les paramètres déclencheurs de la hausse ou de la baisse des rachats conjoncturels ont un impact très limité sur le taux de couverture.

Le résumé des variations est présenté sur le graphique VI.6, de la présente page. L'effet de la nouvelle loi dynamique est attribué pour 5,6% à la loi structurelle et pour 14,4% à la loi conjoncturelle.

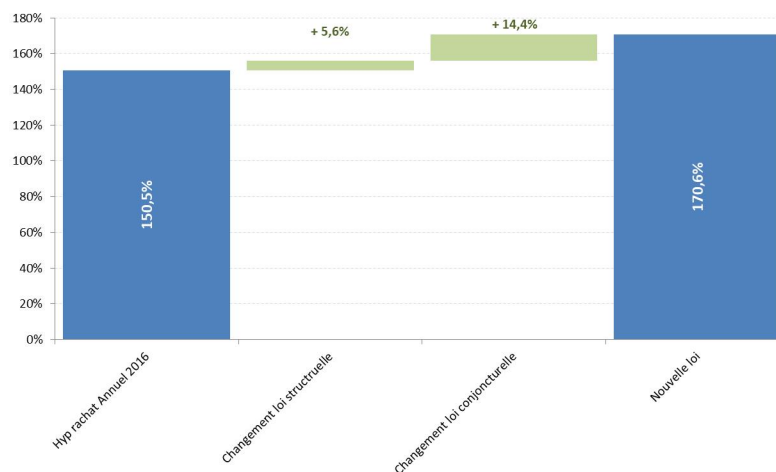


FIGURE VI.6 – Analyse des changements sur le taux de couverture entre 2016 et 2017

Conclusion

L'ensemble des données recueillies à partir du portefeuille LFM présume l'existence d'un lien entre le rachat total d'un contrat d'assurance-vie et l'évolution perçue des marchés obligataires par les adhérents. La période étudiée étant courte, seules les conséquences d'une baisse des taux ont pu être analysées statistiquement sachant l'hypothèse forte d'un taux adverse reposant principalement sur le TME.

L'étude de l'influence des émotions et des facteurs personnels propres à chaque individu, dans la prise de décisions économiques, est complexe à modéliser car les choix ne sont pas toujours rationnels. L'outil des modèles à choix discrets a été mis à profit dans le but de mieux comprendre les mécanismes décisionnels qui poussent les adhérents à racheter leur contrat d'épargne. Ce modèle, grâce à certains indicateurs de classification, a donc permis de découvrir de quelle manière l'environnement économique pourrait influencer la prise de décision d'un adhérent.

Un des problèmes rencontrés est l'inobservation d'un environnement de hausse des taux. Ce risque est pourtant le plus à craindre et pour cela doit être traité de manière cohérente. La solution qui a été présentée est basée sur des simulations. La réalisation d'un tel scénario serait désormais intéressante dans le but de pouvoir confronter, comme toujours, la théorie avec la réalité.

Les résultats obtenus et présentés dans ce mémoire entraînent un ajustement de la loi conjoncturelle. Les nouveaux paramètres sont présentés dans le tableau VI.4, de la présente page, en comparaison de la loi *Min* utilisée jusqu'alors :

Loi	α	β	γ	δ	RC_{min}	RC_{max}
Min	-6,00%	-2,00%	1,00%	2,00%	-6,00%	20,00%
LFM	-4,90%	-1,20%	0,63%	1,68%	-0,25%	20,00%

TABLE VI.4 – Loi dynamique ACPR et loi dynamique LFM

La loi structurelle repose désormais sur cette nouvelle loi, ce qui entraîne sur la période observée une majoration de taux des rachats observés. La revue de la loi de rachat dynamique conduit le taux de couverture à mieux représenter les risques auxquels la Mutuelle doit faire face. L'impact de la revue du comportement des rachats est de 20 points de pourcentage au 31 décembre 2017.

La construction de la loi des rachats structurels n'a, dans un premier temps, pas été réalisée par ancienneté. Les futurs travaux mettront en évidence ce paramètre qui doit être particulièrement suivi avec la mise en vigueur de la réforme de la fiscalité de l'assurance-vie, la *Flat tax*.

Le souhait de La France Mutualiste d'attirer une population plus jeune pourrait modifier les résultats obtenus. L'assurance-vie considérée comme une réserve d'argent au sens large (y compris retraite) aujourd'hui, pourrait être caractérisée davantage par l'hypothèse de rendement demain. Frey (2016) énonce que ce motif peut être représentatif des taux de rachat plus élevés que le marché. La méthodologie présentée dans ce mémoire pourra être, à ce moment là, un point de départ dans le suivi des mécanismes de rachat des contrats en assurance-vie de façon à accompagner la transformation de la Mutuelle.

Bibliographie

- ACPR (2013). Orientations nationales complémentaires aux spécifications techniques pour l'exercice 2013 de préparation à solvabilité ii. Technical report, Autorité de contrôle prudentiel et de résolution. 13
- Afsa, C. (2016, 01). *Le modèle Logit, Théorie et application* (Méthodologie statistique ed.). Institut National de la Statistique et des Études Économiques. Document de travail.
- Ahado, F., E. Ferrand-Tenot, S. Slama, and C. Yang (2017). Assurance vie en france et environnement de taux bas. *Analyses et Synthèses* (78).
- Akerlof, G. and R. Shiller (2009). *Animal Spirits : How Human Psychology Drives the Economy, and Why It Matters for Global Capitalism* (1 ed.). Princeton University Press. 10
- Anderson, C. J., J. Verkuilen, and T. Johnson (2010). Applied generalized linear mixed models : Continuous and discrete data. Technical report, University of illinois at urbana-champaign. 34, 101, 102
- Banque Centrale Européenne (2007). Potential impact of solvency ii on financial stability. Technical report, European Central Bank.
- Beliveau, B. (1981). *Two Aspects of Market Signalling*. Ph. D. thesis, Yale University, New Haven, Connecticut.
- Ben-Akiva, M. and S. R. Lerman (1985). *Discrete Choice Analysis : Theory and Application to Travel Demand*. MIT Press Series. 35, 104
- Bluhm, W. F. (1982). Cumulative anti-selection theory. *Transactions of society of actuaries* 34, 215–246.
- Boughorbel, S., F. Jarray, and M. El-Anbari (2017, 06). Optimal classifier for imbalanced data using matthews correlation coefficient metric. *PLOS ONE* 12(6), 1–17.

- Briere-Giroux, G., J.-F. Huet, R. Spaul, A. Staudt, and D. Weinsier (2010). Predictive modeling for life insurers : Application of predictive modeling techniques in measuring policyholder behavior in variable annuity contracts. *Towers Watson*. 19
- CEIOPS (2010). Qis5 technical specifications. Technical report, European Commission. 13, 14
- Cerchiara, R. R., M. Edwards, and A. Gambini (2008). Generalized linear models in life insurance : Decrements and risk factor analysis under solvency ii. *Actuarial Approach for Financial Risks*. 17, 19, 102
- Charpentier, A., E. Flachaire, and A. Ly (2018, May). Économétrie & Machine Learning. working paper or preprint. 19
- Chevalier, P. (2013). Statistique c. *Minerva* 12(1). 42
- Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society* 34(2), 187–220. 96
- Dagsvik, J. K. (2004, 8). *Random Utility Models for Discrete Choice Behavior, An Introduction*. Statistics Norway, Research Department.
- Dar, A. and C. Dodds (1989). Interest rates, the emergency fund hypothesis and saving through endowment policies : Some empirical evidence for the u.k. *The Journal of Risk and Insurance* 56(3), 415–433. 18
- Duguet, E. (2008). *Econométrie des Variables Qualitatives* (Version 5 ed.).
- Frey, L. (2016). Éclairages de l'enquête patrimoine sur les comportements de rachat en assurance-vie. *Analyses et Synthèses* (59). 22, 72
- Frey, L. and G. Capitaine (2017). Revalorisation 2016 des contrats d'assurance-vie et de capitalisation - engagements à dominante épargne et retraite individuelle. *Analyses et Synthèses* (84). 22
- Goncalves, L., A. Subtil, M. Rosario Oliveira, and P. Bermudez (2014, 03). Roc curve estimation : An overview. 12, 1–20. 114
- Gonen, M. (2006). Receiver operating characteristic (roc) curves. *SAS SUGI Proceedings* 210(31). 117
- Guszcza, J. (2005). Cart from a to b. Technical report, Deloitte and Touche LLP. 34, 102
- Haberman, S. and A. E. Renshaw (1996, 01). Generalized linear models and actuarial science. 45, 407. 19

- Hanley, J. A. and B. J. McNeil (1982). The meaning and use of the area under a receiver operating characteristic (roc) curve. *Radiology* 143, 29–36. 117
- INSEE (2011). L’assurance-vie en 2010 : Une composante majeure du patrimoine des ménages. *INSEE Première* (1361). 8, 9, 24
- INSEE (2015). Le patrimoine des ménages début 2015 : Repli des valeurs risquées au profit des produits sécurisés. *INSEE Première* (1574). 8
- Institut des Actuaire (2016). Groupe de travail best estimate liabilities vie : Exemples de pratiques actuarielles applicables au marché français. Technical report, Institut des Actuaire. 17, 23, 34, 91
- Izrael, D., A. A. Battaglia, D. C. Hoaglin, and M. P. Battaglia (2003). Sas macros and tools for working with weighted logistic regression models that use survey data. *SAS SUGI Proceedings* 275(28). 118
- Kaplan, E. L. and P. Meier (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association* 53(282), 457–481. 95
- Kim, H.-M., Y. Yasui, and I. Burstyn (2005). Attenuation in risk estimates in logistic and cox proportional-hazards models due to group-based exposure assessment strategy. *Department of Public Health Sciences, The University of Alberta, Canada*. 102
- Kuo, W., C. Tsai, and W. Chen (2003). An empirical study on the lapse rate : The cointegration approach. *Journal of Risk and Insurance* 70(3), 489–508. 11, 18, 79
- Loisel, S. and X. Milhaud (2011). From deterministic to stochastic surrender risk models : impact of correlation crises on economic capital. *European Journal of Operational Research* 214(2), 348–357. 11, 18, 79
- Lorent, B. (2008). Raisons fondamentales d’une régulation prudentielle du secteur des assurances. *Centre Emile Bernheim*.
- Matthews, B. (1975). Comparison of the predicted and observed secondary structure of t4 phage lysozyme. *Biochimica et Biophysica Acta (BBA) - Protein Structure* 405(2), 442 – 451. 58
- Mazzanti, A. (2012, 2). La formule standard (version qis5), leviers et/ou incertitudes pour des garanties épargne et retraite. Master’s thesis, Ecole nationale d’assurance. 68
- McFadden, D. (1973). Conditional logit analysis of qualitative choice behaviour. In P. Zarembka (Ed.), *Frontiers in Econometrics*, pp. 105–142. New York, NY, USA : Academic Press New York. 37, 107

- Milhaud, X., S. Loisel, and V. Maume-Deschamps (2011). Surrender triggers in life insurance : what main features affect the surrender behavior in a classical economic context ? *Bulletin Français d'Actuariat* 11(22), 5–48. 18
- Moody's Global Credit Research (2015). European insurers face credit-negative quantitative easing program. •.
- Nelder, J. A. and R. W. M. Wedderburn (1972). Generalized linear models. *Journal of the Royal Statistical Society* 135(3), 370–384. 93
- Outreville, J. (1990). Whole-life insurance lapse rates and the emergency fund hypothesis. *Insurance : Mathematics and Economics* 9(4), 249 – 255. 18
- Parienty, A. (2009). Animal spirits (in the material world). *L'éco en questions*. 10
- Pepe, M. (2000, 03). Receiver operating characteristic methodology. 95, 308–311. 117
- Powers, D. M. W. (2007). Evaluation : From precision, recall and f-factor to roc, informedness, markedness and correlation. Technical report, Flinders University of South Australia. 124
- Rakotomalala, R. *Pratique de la Régression Logistique, Régression Logistique Binaire et Polytomique* (Version 2.0 ed.). Université Lumière Lyon 2.
- Rodriguez, G. (2007). Lecture notes on generalized linear models. Technical report, Princeton University. 97
- Ruppert, D. (2002). Selecting the number of knots for penalized splines. *Journal of Computational and Graphical Statistics* 11(4), 735–757. 134
- Ruppert, D., M. P. Wand, and R. J. Carroll (2003). *Semiparametric Regression*. Cambridge University Press. 134
- SAS** Institute Inc. (2011). *SAS/STAT Software, Version 9.3 User's Guide*. Cary, NC.
- Toulet, A. (2017, 3). Modélisation et couverture du risque derachat total en epargne individuelle. Master's thesis, Ecole Nationale de la Statistique et de l'Administration Economique.
- Youden, W. J. (1950). Index for rating diagnostic tests. *Cancer* 3(1), 32–35. 44, 123

Généralités

A.1 Loi du 4 août 1923

Les Sociétés et Unions approuvées de Sociétés de Secours Mutuels exclusivement composées d'anciens combattants de la Guerre 1914/1918, de veuves, d'orphelins et d'ascendants de militaires morts pour la France au cours de la guerre et celles qui auront organisé, au profit de leurs membres participants, une Caisse Autonome fonctionnant dans les conditions du décret du 25 mars 1901, portant règlement d'administration publique pour l'application de la loi du 1^{er} avril 1898, bénéficieront sous réserve de l'article 28 de la loi du 1^{er} avril 1898, d'une subvention de l'État égale au quart des sommes provenant des ressources sociales affectées au cours de l'année précédente à la constitution des retraites des membres participants sus indiqués, à condition que ces derniers, pour bénéficier de leur pension, soient tenus d'effectuer des versements de retraites pendant dix ans au moins.

A.2 RMC : particularités

Le produit RMC est assorti d'une majoration de l'État dans les conditions prévues à l'article L. 222-2 du Code de la Mutualité (opérations d'assurance sur la vie relevant de la branche 20). L'État verse une majoration dont le taux varie de 12,5% à 60%. Il

est déterminé en fonction de l'âge, du conflit auquel l'adhérent a participé et de la date d'obtention de la Carte du Combattant ou du Titre de Reconnaissance de la Nation. Pour les Victimes de Guerre, c'est la date de délivrance de la mention « Mort pour la France à titre militaire » qui sera prise en considération. En outre, la retraite est revalorisée périodiquement par l'État pour conforter le pouvoir d'achat, grâce aux majorations légales des rentes viagères fixées, chaque année, par arrêté ministériel.

Le montant des versements est libre. Ils peuvent intervenir à n'importe quel moment. Selon les disponibilités, l'adhérent effectue plusieurs versements dans l'année ou opte pour le prélèvement automatique. Les versements sont entièrement déductibles de l'impôt sur le revenu de l'adhérent quant ils servent à financer une rente viagère inférieure au plafond de rente majorée. La rente servie est de même exonérée d'impôt sur le revenu et de prélèvements sociaux (CSG¹ et CRDS²) pour le montant n'excédant pas le plafond fixé. Il est possible pour un adhérent à la retraite de percevoir sa rente viagère tout en continuant de cotiser. Ainsi, il pourra déduire le montant versé sur le contrat et être exonéré d'impôts sur la rente dans la limite du plafond de rente majoré.

A.3 Le risque de rachat sous toutes ces formes

L'incertitude liée à l'estimation des taux de rachat peuvent impacter différents calculs comme :

- **SCR** et **MCR** : les montants des flux de prestations affectent ces valeurs. Ce capital devrait être détenu en partie en actif sans risque. Le coût de détention équivaut aux coûts d'opportunité c'est-à-dire aux gains qu'ils pourraient être retirés en réalisant d'autres placements.
- **Prix de l'assurance** : la déviation des taux de rachats estimés impacte les futurs flux de trésorerie. Le cas se produit lorsque la majeure partie des frais d'administration et d'acquisition sont à la charge de l'assureur au début du contrat. La conséquence d'un rachat inattendu peut engendrer des problèmes de rentabilité.

1. Contribution sociale généralisée.

2. Contribution pour le remboursement de la dette sociale.

- **Évaluation des liquidités** (Kuo et al. (2003)) : la liquidité d'un produit est essentielle lorsque la date du dénouement, comme le rachat, est incertaine. Dans la mesure où le rachat est possible, la couverture du portefeuille devrait être souple dans une certaine mesure. La liquidité des produits a un prix, ce qui veut dire que le coût de couverture est lié au taux de rachat.
- **Évaluation de la rentabilité** (Loisel and Milhaud (2011)) : il est important de savoir dans quelle mesure les flux futurs de trésorerie sont cohérents avec le plan stratégique.

Les conséquences d'une erreur d'estimation des taux de rachat peuvent mettre l'assureur en difficulté face à ces engagements. Il est donc nécessaire de s'efforcer à les estimer au mieux.

A.4 Modélisation d'un contrat racheté

Par exemple, la modélisation se traduit comme suit : l'adhérent a été observé pour la dernière fois en août 2015 car il a racheté totalement son contrat au mois de septembre 2015 mais sa décision a été prise avec le taux du TME de 1,02%.

Date	Epargne acquise	Rachat	TME
30/06/2015	1000€		1,23%
31/07/2015	1000€		1,15%
31/08/2015	1000€	1000€	1,02%

TABLE A.1 – Exemple d'un contrat racheté

A.5 Sur-exposition de certains mois

Les graphiques A.1 et A.2, page suivante, démontrent que la construction de la variable *delta taux* n'est pas issue d'une agglomération de certains mois. En effet, la saisonnalité des rachats aurait pu être à l'origine de la forme en escalier de la courbe des taux de rachat.

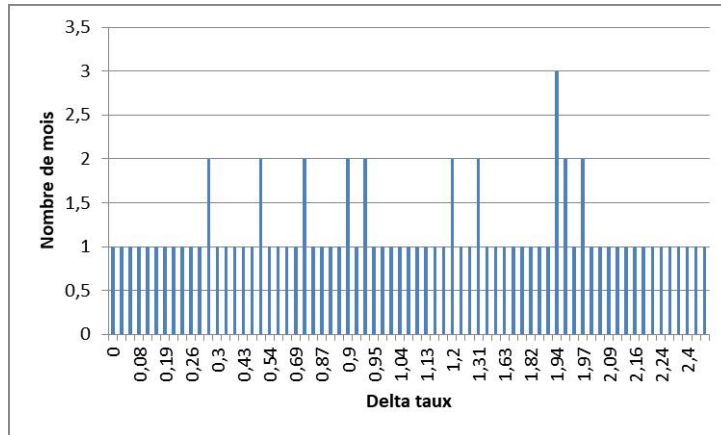


FIGURE A.1 – Nombre de mois accumulés par *delta taux*

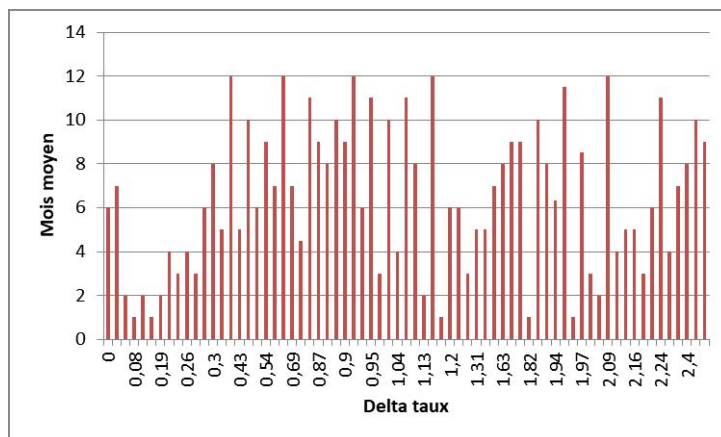


FIGURE A.2 – Mois moyen par *delta taux*

A.6 La RMC (Retraite Mutualiste du Combattant) favorise le développement de La France Mutualiste

La Première Guerre mondiale bouleverse complètement l'organisation de la Mutuelle. Dès 1915, une nouvelle caisse de prévoyance pour les familles des morts au combat est créée. Au lendemain du conflit, l'État vote la loi du 4 août 1923 par laquelle est accordé aux anciens combattants le droit de se constituer une retraite dans des conditions privilégiées. C'est dans ce contexte que La France Mutualiste est fondée en 1925. Caisse autonome de retraite des anciens combattants, elle gère la Retraite Mutualiste du Combattant. Considérant son succès auprès des anciens combattants, c'est l'ensemble de La Boule de Neige qui prend le nom actuel de La France Mutualiste en 1927. Ce nom n'est autre que celui de la revue fondée par Léon Guillot en 1900 à destination de toutes les sociétés « ayant pour but d'améliorer la situation matérielle et morale des travailleurs ».

En 1932, La France Mutualiste devient une Union des sociétés mutualistes d'anciens combattants et victimes de guerre. Elle a pour rôle à la fois de fédérer les sociétés affiliées et de gérer la caisse autonome. Les activités héritées de La Boule de Neige sont confiées à une nouvelle mutuelle, la « Retraite Mutualiste », elle-même affiliée à La France Mutualiste. A compter de cette date et pendant plus de 50 ans, La France Mutualiste se consacre exclusivement à la RMC.

Tout au long de son histoire, La France Mutualiste a vécu au rythme de ces ouvertures de droits à la RMC. Si au départ, la RMC a été conçue pour les Poilus, progressivement, chaque génération d'anciens combattants a pu en bénéficier : ceux de la Seconde Guerre mondiale en 1950, d'opérations extérieures menées en 1954 (guerre du Rif, campagne de Syrie), d'Indochine et de Corée en 1964, et enfin d'Afrique du Nord en 1972. Depuis 1995, les titulaires du titre de reconnaissance de la nation sont autorisés à souscrire à la RMC. Elle est ainsi accessible aux militaires ayant participé aux opérations extérieures. En 1992, La France Mutualiste crée une gamme de produits d'assurance vie complémentaires à la RMC ouverte aux familles des anciens combattants. En 2002, la France Mutualiste

devient « Mutuelle nationale » et, ouvre deux ans plus tard ses produits au grand public. Aujourd'hui, La France Mutualiste perpétue l'esprit mutualiste qui ne cesse de l'animer depuis sa fondation. La Mutuelle entretient par un dialogue intergénérationnel, l'esprit « Boule de Neige » préconisé par son fondateur. Elle travaille à l'élargissement de son offre pour mieux répondre aux attentes de ses adhérents.

A.7 Mesures pour encadrer les taux garantis

L'Arrêté³ du 7 juillet 2010 portant modification des modalités de garanties d'un taux minimum par les entreprises d'assurance vise à mieux encadrer les taux garantis dans le but de :

- Introduction d'une règle garantissant que la collectivité des assurés ne finance pas le supplément de rendement garanti à certains clients :
 - Le montant minimal de participations aux bénéfices attribué à la collectivité des assurés est augmenté du coût que représente l'offre de taux garantis supérieurs au taux moyen servi à cette collectivité. Concrètement, un assureur qui garantit des taux plus élevés à certains assurés (par exemple 4% sur 20% de son encours, pour un taux moyen servi prévu de 3,5%) devra prendre sur ses ressources propres pour augmenter la rémunération de la collectivité des assurés (le taux moyen servi sera au final de 3,6%).
- Prise en compte de l'environnement des marchés obligataires dans la détermination du niveau de taux maximal pouvant être garanti contrat par contrat :
 - **En cas de stabilité des marchés obligataires**, les assureurs peuvent garantir 110% du taux moyen servi lors des deux derniers exercices. Par exemple, en 2010, sur la base d'un rendement moyen de 3,75% en 2009 et 2008⁴, un assureur peut garantir un taux maximum de 4,12% soit un niveau proche de la

3. Journal officiel de la république française numéro 0174 du 30 juillet 2010 page 14111 texte numéro 13 modifiant le Code des assurances (articles A. 132-2, A. 132-3 et A. 331-5).

4. Taux moyen servi en 2008 égal à 3,90% (source ACAM) et hypothèse d'une baisse de 0,3% du taux servi en 2009.

perspective de rémunération d'un contrat en euros dont le rendement bénéficie de la diversification des actifs.

- **En cas de baisse des taux significative**, l'assureur doit en revanche limiter ses offres commerciales à un taux ayant pour référence 110% du taux moyen des emprunts d'État (TME). L'existence d'une telle règle aurait limité dans le passé récent les garanties offertes par les assureurs. Ainsi, en fin d'année 2005, les assureurs ne pouvaient obtenir que difficilement un rendement compris entre 3% et 3,5% en investissant les primes encaissées sur les marchés obligataires. Or, la règle alors en vigueur les autorisait à garantir des taux de l'ordre de 4,5%. Dans la même configuration de marché et après la réforme, la garantie ne pourra excéder un taux de 3,75%, plus en phase avec les perspectives de revalorisation du contrat.
 - **En cas de hausse des taux significative**, l'assureur peut dynamiser son offre commerciale en proposant un taux ayant pour référence 90% du TME. Ainsi, dans une situation le TME atteindrait rapidement 5%, les assureurs pourraient garantir un taux de 4,5%, prudent par rapport aux conditions d'investissement des primes.
- Introduction d'une enveloppe globale limitant l'ensemble des propositions de taux garantis de l'assureur au regard de ses performances financières passées et des engagements contractuels déjà pris :
- L'enveloppe est égale au produit de (i) 80% du taux moyen de rendement des actifs de l'assureur et de (ii) son encours en début d'année, duquel sont déduits les intérêts techniques qu'il devra verser.

Modèle logistique

B.1 Notations

L'objectif est de prédire les valeurs prises par la variable aléatoire Y définie dans $\{y_1, y_2, \dots, y_K\}$. Pour la régression logistique binaire, Y prend uniquement deux modalités $\{+, -\}$ (ou $\{1, 0\}$ pour simplifier). Nous disposons d'un échantillon Ω de taille n . La valeur prise par Y pour un individu ω est notée $Y(\omega)$.

Le fichier comporte J descripteurs $\{X_1, X_2, \dots, X_J\}$. Le vecteur de valeurs pour un individu ω s'écrit $(X_1(\omega), X_2(\omega), \dots, X_J(\omega))$.

Dans le cadre binaire, pour un individu donné, sa probabilité a priori d'être positif s'écrit $P[Y(\omega) = +] = p(\omega)$. Lorsqu'il ne peut y avoir d'ambiguïtés, nous la noterons simplement p .

Lorsque l'échantillon est issu d'un tirage aléatoire dans la population, sans distinction des classes d'appartenance, si n_+ est le nombre d'observations positives dans Ω , p peut être estimée par $\frac{n_+}{n}$. On parle de "schéma de mélange".

La **probabilité a posteriori** d'un individu ω d'être positif c.-à-d. sachant les valeurs prises par les descripteurs est notée $P[Y(\omega) = +/X(\omega)] = \pi(\omega)$. Ici également, lorsqu'il ne peut y avoir de confusions, nous écrirons π . Ce dernier terme est très important. En effet, c'est la probabilité que l'on cherche à modéliser en apprentissage supervisé.

Le LOGIT d'un individu ω s'écrit

$$\ln \left[\frac{\pi(\omega)}{1 - \pi(\omega)} \right] = X(\omega) \times a \quad (\text{B.1})$$

avec $X(\omega) = (1, X_1(\omega), X_2(\omega), \dots, X_J(\omega))$, la première composante ($X_0(\omega) = 1, \forall \omega$) symbolise la constante; $a' = (a_0, a_1, \dots, a_J)$ est le vecteur des paramètres.

Enfin, toujours pour alléger l'écriture, nous omettrons le terme ω lorsque cela est possible.

B.2 Principe et estimation

En apprentissage supervisé, l'objectif est de prédire et/ou expliquer une variable catégorielle Y à partir d'une collection de descripteurs $X = (X_1, X_2, \dots, X_J)$. Il s'agit en quelque sorte de mettre en évidence l'existence d'une liaison fonctionnelle sous-jacente (en anglais, *underlying concept*) de la forme

$$Y = f(X, \alpha) \quad (\text{B.2})$$

entre ces variables.

La fonction $f(\cdot)$ est le modèle de prédiction, on parle aussi de classifieur; α est le vecteur des paramètres de la fonction, on doit en estimer les valeurs à partir des données disponibles.

Dans le cadre de la discrimination binaire, nous considérons que la variable dépendante Y ne prend que 2 modalités : positif "+" ou négatif "-". Nous cherchons à prédire correctement les valeurs de Y , mais nous pouvons également vouloir quantifier la propension (la probabilité) d'un individu à être positif (ou négatif).

Un aspect fondamental de l'apprentissage supervisé est de pouvoir au mieux évaluer la qualité de la modélisation. Simplement cela consiste à mesurer la qualité de la prédiction autrement dit la capacité du modèle à prédire correctement dans la population Ω^{pop} . Elle est définie comme l'**erreur théorique** qui est interprétée comme la **probabilité de mal**

classer un individu :

$$\varepsilon = \frac{1}{\text{card}(\Omega^{\text{pop}})} \sum_{\omega} \Delta[Y(\omega), f(X(\omega))] \quad (\text{B.3})$$

où Δ est une fonction indicatrice qui, pour un individu ω donné, prend la valeur 1 lorsque la prédiction ne concorde pas avec la vraie valeur ; elle prend la valeur 0 lorsque le modèle prédit à bon escient.

On confronte ainsi les vraies valeurs prises par la variable dépendante dans la population et les prédictions du modèle. Dans le cas idéal, toutes les prédictions sont correctes, l'erreur théorique est égal à 0. L'autre extrême serait que le modèle se trompe systématiquement, dans ce cas le taux serait égal à 1. Mais en réalité, il est plus judicieux de prendre comme borne haute le classement au hasard. Lorsque les classes sont équi-distribuées c.-à-d. les proportions de positifs et de négatifs sont identiques dans la population, nous obtiendrions un taux d'erreur théorique égal à 0,5. Le classifieur doit faire mieux.

Le classifieur bayésien est celui qui répond de manière optimale aux spécifications ci-dessus. Pour un individu ω , il s'agit de calculer les probabilités conditionnelles (probabilité a posteriori)

$$P[Y(\omega) = y_k | X(\omega)] \quad (\text{B.4})$$

pour chaque modalité y_k de Y .

On affecte à l'individu la modalité la plus probable y_{k^*} c.-à-d.

$$y_{k^*} = \arg \max_k P[Y(\omega) = y_k | X(\omega)] \quad (\text{B.5})$$

On associe donc l'individu à la classe la plus probable compte tenu de ses caractéristiques $X(\omega)$.

Cette approche est optimale au sens de l'erreur théorique. Mais un problème apparaît tout de suite : comment estimer correctement ces probabilités conditionnelles ?

Pour rendre calculable la quantité $P(Y = y_k | X)$, il nous faudra donc introduire une ou plusieurs hypothèses sur les distributions. Nous sommes dans le cadre des méthodes dites "paramétriques" (ou "semi-paramétriques"). Elles semblent plus contraignantes par

rapport aux méthodes dites non-paramétriques qui, elles, procèdent à l'estimation des probabilités sans jamais introduire des hypothèses sur les distributions (ex. les arbres de décision, la méthode des plus proches voisins, etc.). En effet, lors du traitement d'un problème réel, il faudrait en toute rigueur s'assurer de la crédibilité des hypothèses avant de pouvoir mettre en œuvre la technique.

En pratique, ce n'est pas nécessaire. On se rend compte que les méthodes paramétriques sont souvent robustes. Elles restent opérationnelles même lorsque l'on s'écarte assez fortement des hypothèses qui les sous-tendent. L'idée la plus importante à retenir finalement est que les hypothèses pèsent sur la forme de la frontière induite pour distinguer les classes dans l'espace de représentation. La régression logistique par exemple produit un séparateur linéaire, c'est la principale information qu'il faut retenir.

Avant de décrire les hypothèses introduites dans la régression logistique, reconsidérons la probabilité conditionnelle $P(Y = y_k|X)$:

$$\begin{aligned} P(Y = y_k|X) &= \frac{P(Y = y_k) \times P(X|Y = y_k)}{P(X)} \\ &= \frac{P(Y = y_k) \times P(X|Y = y_k)}{\sum_k P(Y = y_k) \times P(X|Y = y_k)} \end{aligned} \quad (\text{B.6})$$

Dans le cas à deux classes, nous devons comparer simplement $P(Y = +|X)$ et $P(Y = -|X)$. Formons-en le rapport,

$$\frac{P(Y = +|X)}{P(Y = -|X)} = \frac{P(Y = +)}{P(Y = -)} \times \frac{P(X|Y = +)}{P(X|Y = -)} \quad (\text{B.7})$$

La règle de décision devient

$$\text{Si } \frac{P(Y = +|X)}{P(Y = -|X)} > 1 \text{ Alors } Y = + \quad (\text{B.8})$$

Le véritable enjeu réside donc dans l'estimation du rapport de probabilité $\frac{P(Y=+|X)}{P(Y=-|X)}$. La régression logistique introduit l'hypothèse fondamentale suivante :

$$\ln \left[\frac{P(Y = +|X)}{P(Y = -|X)} \right] = b_0 + b_1 X_1 + \dots + b_J X_J \quad (\text{B.9})$$

La régression logistique peut être décrite d'une autre manière. Pour un individu ω , on appelle transformation LOGIT de $\pi(\omega)$ l'expression

$$\ln \left[\frac{\pi(\omega)}{1 - \pi(\omega)} \right] = a_0 + a_1 X_1 + \cdots + a_J X_J \quad (\text{B.10})$$

La quantité $\frac{\pi}{1-\pi} = \frac{P(Y=+|X)}{P(Y=-|X)}$ exprime un **odds** c.-à-d. un rapport de chances. Par exemple, si un individu présente un odds de 2, cela veut dire qu'il a 2 fois plus de chances d'être positif que d'être négatif.

Posons $C(X) = a_0 + a_1 X_1 + \cdots + a_J X_J$, nous pouvons revenir sur π avec la fonction logistique

$$\begin{aligned} \pi &= \frac{e^{C(X)}}{1 + e^{C(X)}} \\ &= \frac{1}{1 + e^{-C(X)}} \end{aligned} \quad (\text{B.11})$$

B.3 Estimation des paramètres par la maximisation de la vraisemblance

Pour estimer les paramètres de la régression logistique par la méthode du maximum de vraisemblance, nous devons tout d'abord déterminer la loi de distribution de $P(Y|X)$.

Y est une variable binaire définie dans $\{+, -\}$, (ou $\{1, 0\}$ pour simplifier les écritures). Pour un individu ω , on modélise la probabilité à l'aide de la loi binomiale $\mathcal{B}(1, \pi)$, avec

$$P[Y(\omega)|X(\omega)] = \pi(\omega)^{y(\omega)} \times (1 - \pi(\omega))^{1-y(\omega)} \quad (\text{B.12})$$

Cette modélisation est cohérente avec ce qui a été dit précédemment, en effet :

- Si $y(\omega) = 1$, alors $P[Y(\omega) = 1|X(\omega)] = \pi$;
- Si $y(\omega) = 0$, alors $P[Y(\omega) = 0|X(\omega)] = 1 - \pi$;

La vraisemblance (en anglais likelihood) d'un échantillon Ω s'écrit

$$L = \prod_{\omega} \pi(\omega)^{y(\omega)} (1 - \pi(\omega))^{1-y(\omega)} \quad (\text{B.13})$$

Pour alléger l'écriture, nous utiliserons pour la suite

$$L = \prod_{\omega} \pi^y (1 - \pi)^{1-y} \quad (\text{B.14})$$

N'oublions pas que la vraisemblance correspond à la probabilité d'obtenir l'échantillon Ω à partir d'un tirage dans la population. Elle varie donc entre 0 et 1. La méthode du maximum de vraisemblance consiste à produire les paramètres $a = (a_0, a_1, \dots, a_J)$ de la régression logistique qui rendent maximum la probabilité d'observer cet échantillon.

Pour faciliter les manipulations, on préfère souvent travailler sur la log-vraisemblance (*log-likelihood*)

$$LL = \sum_{\omega} y \times \ln \pi + (1 - y) \times (1 - \ln \pi) \quad (\text{B.15})$$

Le logarithme étant une fonction monotone, le vecteur \mathbf{a} qui maximise la vraisemblance est le même que celui qui maximise la log-vraisemblance. Cette dernière en revanche varie entre $-\infty$ et 0.

Puisque \hat{a} est un estimateur du maximum de vraisemblance, il en possède toutes les propriétés :

- Il est asymptotiquement sans biais ;
- Il est de variance minimale ;
- Il est asymptotiquement gaussien.

Ces éléments, notamment le dernier, seront très importants pour l'inférence statistique (intervalle de confiance, test de significativité, etc.).

La log-vraisemblance est une fonction convexe, il existe donc une solution unique \hat{a} cependant elle n'est pas disponible à partir d'une solution analytique. L'algorithme de Newton-Raphson est une des méthodes numériques les plus utilisées pour optimiser la log-vraisemblance. Il démarre avec une initialisation quelconque du vecteur de paramètre \mathbf{a} ; pour passer de l'étape (i) à l'étape $(i + 1)$, il se rapproche de la solution finale \hat{a} en utilisant la formule suivante

$$a^{i+1} = a^i - \left(\frac{\partial LL}{\partial a \partial a^T} \right)^{-1} \times \frac{\partial LL}{\partial a} \quad (\text{B.16})$$

Modèle à choix discret

L'outil le plus communément employé parmi les articles académiques, en ce qui concerne la modélisation des rachats, est le modèle logistique ou autrement dit le modèle logit.

La première idée lors du traitement d'une variable binaire est pourtant simplement le modèle Binomial. Ce dernier ne considère cependant pas le cas des censures et des troncatures dans l'observation des données, ainsi qu'il doit être fait d'après l'Institut des Actuaire (2016). Alors dans ce cas, il est possible de s'appuyer sur le modèle de durée de survie de Cox. Nonobstant, à cause de la structure longitudinale des données, il sera démontré par la suite que ces deux modèles conduisent tous deux au modèle logistique appartenant à la famille des Modèles Linéaires Généralisés.

Ces derniers seront alors présentés afin de poser les principales hypothèses de leur maniement. En supposant que le comportement d'un individu n'est pas nécessairement en accord avec le modèle standard de la rationalité parfaite¹, il apparaît finalement plus opportun de se reporter vers les modèles à choix discrets. Ils sont basés sur la théorie de l'utilité aléatoire relativement à des alternatives discrètes. Ils ont comme objectif l'esti-

1. L'homo oeconomicus :

- est capable de maximiser sa satisfaction en utilisant au mieux ses ressources : il maximisera son utilité (et non pas son profit)
- sait analyser et anticiper le mieux possible la situation et les événements du monde qui l'entoure afin de prendre les décisions permettant cette maximisation.

mation (la prévision) d'un choix à partir de la connaissance d'un ensemble d'alternatives mutuellement exclusives et collectivement exhaustives.

C.1 Le modèle binomial

La variable rachat est codée sous une forme binaire en faisant l'hypothèse qu'elle ne peut prendre seulement que deux valeurs appartenant à l'ensemble $\{0, 1\}$. La variable réponse est définie comme suit :

$$y_i = \begin{cases} 1 & \text{si le rachat intervient} \\ 0 & \text{sinon} \end{cases}$$

L'observation y_i est vue comme une réalisation d'une variable aléatoire Y_i prenant les valeurs 1 et 0 avec respectivement les probabilités π_i et $1 - \pi_i$. La distribution des probabilités Y est appelée une distribution de *Bernouilli*² de paramètre π_i et peut être écrite sous une forme compacte :

$$\Pr\{Y_i = y_i\} = \pi_i^{y_i} (1 - \pi_i)^{1-y_i}$$

pour $y_i = 0, 1$. Pour $y_i = 1$ alors la probabilité est égale à π_i , et pour $y_i = 0$ elle vaut $1 - \pi_i$. Un calcul direct permet de démontrer que la valeur moyenne théorique et la variance sont :

$$\begin{aligned} \mathbf{E}(Y_i) &= \mu_i = \pi_i, \text{ et} \\ \mathbf{Var}(Y_i) &= \sigma_i^2 = \pi_i(1 - \pi_i) \end{aligned}$$

Il faut noter que la moyenne et la variance dépendent de la probabilité sous-jacente π_i . N'importe quel facteur affectant la probabilité impacte non seulement la moyenne mais aussi la variance des observations.

2. Décrite pour la première fois par Isaac Newton en 1676 et démontrée pour la première fois par le mathématicien suisse Jakob Bernoulli en 1713, la loi binomiale est l'une des distributions de probabilité les plus fréquemment rencontrées en statistique appliquée.

Dans le cas général avec n_i valeurs de y_i et en posant l'hypothèse que ces n_i observations sont indépendantes et ont toutes la même probabilité π_i de se réaliser, alors la distribution de Y_i est binomiale avec les paramètres π_i et n_i :

$$Y_i \sim B(n_i, \pi_i)$$

La fonction de distribution des probabilités de Y_i est par conséquent :

$$\Pr\{Y_i = y_i\} = \binom{n_i}{y_i} \pi_i^{y_i} (1 - \pi_i)^{n_i - y_i}$$

Pour $y_i = 0, 1, \dots, n_i$. Alors $\pi_i^{y_i} (1 - \pi_i)^{n_i - y_i}$ est la probabilité d'obtenir y_i succès et $n_i - y_i$ échecs. Le coefficient de combinaison définit le nombre de fois d'obtenir y_i succès parmi n_i essais.

L'espérance et la variance de Y_i sont :

$$\mathbf{E}(Y_i) = \mu_i = n_i \pi_i, \text{ et}$$

$$\mathbf{Var}(Y_i) = \sigma_i^2 = n_i \pi_i (1 - \pi_i)$$

Or il peut être démontré que la loi $B(n_i, \pi_i)$ appartient à la famille des lois exponentielles de Nelder and Wedderburn (1972) et peut être réécrite sous la forme :

$$f_i(y_i) = \binom{n_i}{y_i} \pi_i^{y_i} (1 - \pi_i)^{n_i - y_i}$$

En passant au log

$$\log f_i(y_i) = y_i \log(\pi_i) + (n_i - y_i) \log(1 - \pi_i) + \log \binom{n_i}{y_i}$$

En regroupant les termes en y_i alors

$$\log f_i(y_i) = y_i \log\left(\frac{\pi_i}{1 - \pi_i}\right) + n_i \log(1 - \pi_i) + \log \binom{n_i}{y_i}$$

L'expression a une forme exponentielle générale

$$\log f_i(y_i) = \frac{y_i \theta_i - b(\theta_i)}{a_i(\phi)} + c(y_i, \phi)$$

Le coefficient de y_i peut être vu comme le paramètre canonique³ dont la valeur est le logit de π_i

$$\theta_i = \log \left(\frac{\pi_i}{1 - \pi_i} \right)$$

En résolvant l'équation pour π_i alors

$$\pi_i = \frac{e^{\theta_i}}{1 + e^{\theta_i}}, \text{ alors } 1 - \pi_i = \frac{1}{1 + e^{\theta_i}}$$

En posant $\log(1 - \pi_i) = -\log(1 + e^{\theta_i})$, il est possible d'identifier la fonction $b(\theta_i)$ comme

$$b(\theta_i) = n_i \log(1 + e^{\theta_i})$$

Le dernier terme $c(y_i, \phi)$ est une fonction de y_i et non de π_i donc

$$c(y_i, \phi) = \log \binom{n_i}{y_i}$$

Finalement $a_i(\phi) = \phi$ et donc $\phi = 1$.

La moyenne et la variance se calculent en différenciant $b(\theta_i)$ selon θ_i

$$\mu_i = b'(\theta_i) = n_i \frac{e^{\theta_i}}{1 + e^{\theta_i}} = n_i \pi_i$$

En différenciant encore une fois

$$\sigma_i^2 = a_i(\phi) b''(\theta_i) = n_i \frac{e^{\theta_i}}{(1 + e^{\theta_i})^2} = n_i \pi_i (1 - \pi_i)$$

3. La forme canonique d'une fonction est une forme d'écriture paramétrique de l'équation de la fonction. On dit que la forme canonique d'une fonction est porteuse de sens puisqu'elle donne de l'information sur l'allure du graphique de la fonction. On parle aussi de la forme transformée.

Cette section démontre qu'à partir d'une modélisation simple de la probabilité, qui équivaut à la proportion, de rachat dans le portefeuille, le cas du modèle de régression logistique peut être naturellement considéré et appliqué aux observations de l'étude.

C.2 Modèle de durée de survie

Le modélisation du temps de présence au sein d'un jeu de données peut être réalisée à partir des modèles de durée de survie. L'événement terminal est ici la survenue de la décision de rachat. L'estimation de la fonction de survie par l'estimateur de Kaplan and Meier (1958) est inappropriée au jeu de données employé dans cette étude car les observations recueillies sont longitudinales. En effet plusieurs observations sont disponibles pour un même individu à travers le temps. Cependant il existe une extension aux modèles de hasards proportionnels en temps discret conduisant également dans ce cas au modèle logistique.

C.2.1 Risque en temps discret et survie

Les modèles de survie ont la capacité de décrire tout événement en rapport avec un changement d'état binaire. Dans le cas de cette étude l'événement est le rachat total. Le mot survie et ses termes associés seront utilisés par généralité d'un événement terminal afin d'exprimer le temps de présence au sein du portefeuille. Sachant T être une variable aléatoire discrète prenant les valeurs $t_1 < t_2 < \dots$ avec les probabilités :

$$f(t_j) = f_j = \Pr\{T = t_j\}$$

La fonction de survie est définie au temps t_j comme la probabilité d'être vivant au moins jusqu'au temps t_j . Elle est le complément de la fonction de répartition cumulée :

$$S(t_j) = S_j = \Pr\{T \geq t_j\} = \sum_{k=j}^{\infty} f_k$$

Le taux de défaillance ou fonction d'aléa au temps t_j est défini comme la probabilité conditionnelle de décès à ce temps sachant qu'il a survécu jusqu'à ce temps, alors :

$$\lambda(t_j) = \lambda_j = \Pr\{T = t_j | T \leq t_j\} = \frac{f_j}{S_j}$$

Il faut noter qu'en temps discret le taux de défaillance est une probabilité conditionnelle plutôt qu'un taux. Cependant, le résultat exprimant le taux de défaillance comme un ratio de la densité à la fonction de survie est toujours valide.

Un résultat intéressant en temps discret est que la fonction de survie au temps t_j peut-être écrite en termes de taux de défaillance à tous les temps précédents t_1, \dots, t_{j-1} comme :

$$S_j = (1 - \lambda_1)(1 - \lambda_2) \dots (1 - \lambda_{j-1})$$

Autrement dit, ce résultat établit que pour survivre au temps t_j l'individu doit survivre à t_1 , puis survivre à t_2 sachant qu'il a survécu à t_1 et ainsi de suite, finalement survivre à t_{j-1} jusqu'à cet instant. Ce résultat est analogue à celui liant la fonction de survie en temps continu à la fonction d'aléa cumulé aux temps précédents.

C.2.2 Survie en temps discret et régression logistique

Cox (1972) propose une extension au modèle de risque proportionnel au temps discret en travaillant avec les risques relatifs ou *odds* de décès à chaque moment t_j sachant la survie jusqu'à ce point. Il proposa ce modèle :

$$\frac{\lambda(t_j | \mathbf{x}_i)}{1 - \lambda(t_j | \mathbf{x}_i)} = \frac{\lambda_0(t_j)}{1 - \lambda_0(t_j)} \exp\{\mathbf{x}b\} \quad (\text{C.1})$$

Ici $\lambda(t_j | \mathbf{x}_i)$ est le risque au temps t_j pour un individu avec des covariables \mathbf{x}_i , $\lambda_0(t_j)$ est fonction de risque instantané au temps t_j , et $\exp\{\mathbf{x}b\}$ est le risque relatif associé aux valeurs des covariables \mathbf{x}_i . En passant au log, il est facile d'obtenir un modèle sur le *logit*

des probabilités proportionnelles de décès au temps t_j sachant la survie à ce point

$$\text{logit}\lambda(t_j|\mathbf{x}_i) = \alpha_j + \mathbf{x}_i b \quad (\text{C.2})$$

où $\alpha_j = \text{logit}\lambda_0(t_j)$ est le *logit* de la fonction de risque instantané et $\mathbf{x}_i' b$ est l'effet des covariables sur le *logit* du hasard. Le modèle traite essentiellement le temps comme un facteur discret en introduisant un paramètre α_j pour chaque moment possible de décès t_j . L'interprétation des paramètres β associés avec les autres covariables suit pour l'essentiel le même principe que la régression logistique.

En fait, l'analogie avec la régression logistique va plus loin : le modèle à risques proportionnels en temps discret peut être ajusté en utilisant une régression logistique sur un jeu de données générées comme ceci : un individu i meurt ou est censuré à un certain moment $t_{j(i)}$. Des indicateurs de décès d_{ij} sont ensuite générés qui prennent la valeur 1 si l'individu i décède au temps j et 0 sinon, en multipliant l'individu pour chaque temps discret de t_1 à $t_{j(i)}$. A chacun de ces indicateurs il est associé un vecteur de covariables \mathbf{x}_i et un label j identifiant le moment. Le modèle à risques proportionnels (C.2) peut alors convenir en traitant les d_{ij} comme des observations de *Bernoulli* indépendantes avec une probabilité donnée par le risque λ_{ij} pour un individu à ce moment là.

Plus généralement, ces observations peuvent être groupées avec des valeurs de covariables identiques. Si d_{ij} est le nombre de décès et n_{ij} le nombre total d'individus avec les valeurs des covariables observées à un moment donné t_j alors d_{ij} est considéré comme binomial avec les paramètres n_{ij} et λ_{ij} , où ce dernier satisfait le modèle de risques proportionnels.

La preuve de ce résultat est donnée par Rodriguez (2007).

C.3 Le modèle logistique

Le traitement d'une variable décrivant une réponse binaire peut être issu d'une modélisation par une loi binomiale ou un modèle de durée. Ces deux méthodes conduisent au modèle logistique qui présente l'avantage d'estimer une probabilité π_i dépendant d'un

vecteur de covariables observées \mathbf{x}_i . Cette partie présente la structure systématique (terme approfondi à la section C.4 page 101) qui spécifie le vecteur μ en fonction de paramètres inconnus β ou vecteur des coefficients de régression. L'idée la plus simple serait d'écrire π_i comme une fonction linéaire des covariables :

$$\pi_i = \mathbf{x}'_i \beta \tag{C.3}$$

Ce modèle (C.3) est quelquefois appelé *modèle linéaire de probabilités*. Il est souvent estimé à l'aide de la méthode des moindres carrés ordinaires (en anglais *ordinary least squares*).

Le problème avec l'utilisation d'un tel modèle se retrouve avec la probabilité π_i sur la partie gauche. Elle doit être prédite à partir des valeurs 0 ou 1, mais les covariables $\mathbf{x}'_i \beta$ de la partie droite peuvent modéliser n'importe quelles valeurs réelles. Il n'y a alors aucune garantie que les valeurs prédites soient bien comprises dans un ensemble binaire $\{0, 1\}$ sans que quelques restrictions soient imposées sur les coefficients.

Une solution peu complexe au problème est de transformer la probabilité pour retirer la restriction sur les valeurs estimées et modéliser la transformation comme une fonction linéaire des covariables. La probabilité π_i peut être écrite comme un rapport ou *odds* :

$$\text{odds}_i = \frac{\pi_i}{1 - \pi_i}$$

définie comme le ratio de la probabilité sur son complémentaire ou le ratio des cas favorables sur les cas non favorables. Ensuite en passant au logarithme, le calcul du *logit* devient :

$$\eta_i = \text{logit}(\pi_i) = \log\left(\frac{\pi_i}{1 - \pi_i}\right)$$

qui a l'effet de retirer la restriction sur la valeur plancher. Ce point peut être vu en s'apercevant que si la probabilité tend vers 0, le rapport tend vers zéro et le *logit* approche la valeur $-\infty$. A l'autre extrême si la probabilité s'approche de 1, le rapport tend vers $+\infty$ ainsi que le *logit*. En fin de compte, les *logits* représentent les probabilités de l'ensemble $\{0, 1\}$ sur une ligne de nombres réels, voir la figure C.1 page ci-contre.

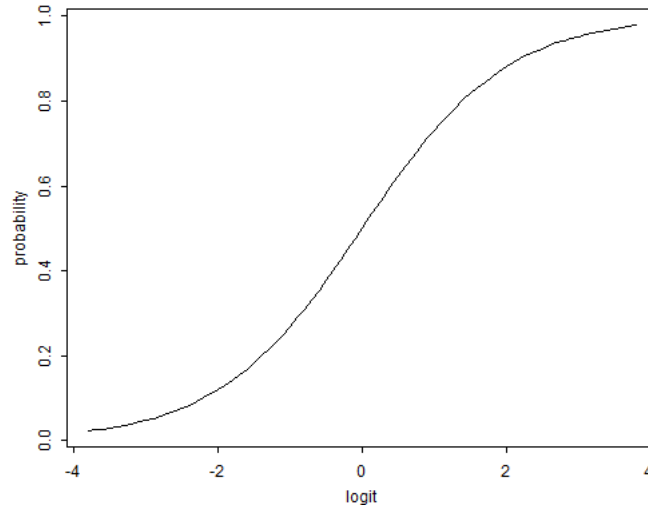


FIGURE C.1 – La transformation logit

Les *logits* peuvent aussi être définis par la moyenne $\mu_i = n_i \pi_i$ qui est le *log* du ratio des succès estimés μ_i sur les échecs estimés $n_i - \mu_i$. Le résultat est le même car le dénominateur binomial n_i est supprimé lorsque le calcul du *odds* est réalisé.

Soient k observations indépendantes y_1, \dots, y_k , et que la i^e observation est la réalisation d'une variable aléatoire Y_i . Elle suit une loi binomiale par hypothèse :

$$Y_i \sim \mathcal{B}(n_i, \pi_i) \tag{C.4}$$

de paramètres n_i et de probabilité π_i . Ceci définit la structure stochastique du modèle. Si le *logit* de la probabilité sous-jacente π_i est une fonction linéaire des covariables :

$$\text{logit}(\pi_i) = \mathbf{x}'_i \beta \tag{C.5}$$

où \mathbf{x}_i est un vecteur de covariables et β un vecteur des coefficients de régression. Ceci définit la structure systématique du modèle.

Le modèle défini aux équations (C.4) et (C.5) est un modèle linéaire généralisé avec une réponse binomiale et un lien *logit*.

Les coefficients de régression β peuvent être interprétés comme ceux des modèles linéaires, en gardant à l'esprit que la partie gauche est un *logit* plutôt qu'une moyenne.

Alors β_j représente la modification du *logit* de la probabilité associée lorsque la j^e variable s'accroît d'une unité et toutes les autres restant constantes.

En passant l'équation (C.5) à l'exponentielle le *odds* de la i^e variable est données par :

$$\frac{\pi_i}{1 - \pi_i} = \exp\{\mathbf{x}'_i \beta\} \quad (\text{C.6})$$

Cette expression définit un modèle multiplicatif pour les *odds*. Par exemple si la j^e variable varie d'une unité pendant que toutes les autres restent constantes le *odds* devrait être multiplié par $\exp\{\beta_j\}$. Alors ce dernier est souvent dénommé un *odds ratio*.

En résolvant pour π_i l'équation (C.5), la probabilité peut s'écrire sous la forme :

$$\pi_i = \frac{\exp\{\beta_j\}}{1 + \exp\{\beta_j\}} \quad (\text{C.7})$$

Si la partie de gauche est une échelle de probabilités, la partie de droite est quant à elle une fonction non linéaire des covariables et il n'existe pas une manière simple d'interpréter l'effet sur la probabilité d'une variation d'une unité d'une covariable, les autres ne changeant pas. Une approximation peut être déduite en prenant les dérivées selon x_j . En utilisant la règle du quotient ⁴ :

$$\frac{d\pi_i}{dx_{ij}} = \beta_j \pi_i (1 - \pi_i) \quad (\text{C.8})$$

Par conséquent, l'effet de la j^e variable sur la probabilité π_i dépend du coefficient β_j et de la valeur de la probabilité. Ce produit peut être évalué comme la moyenne de l'échantillon, la proportion de cas présentant l'événement. Le résultat est une approximation de l'effet de la covariable près de la moyenne.

Le modèle logistique ainsi défini il est intéressant de passer en revue le cadre méthodologique de la famille à laquelle il appartient : le modèle linéaire généralisé.

4. Soient $f(x)$ et $g(x)$ deux fonctions, alors la dérivée du quotient

$$\left(\frac{f(x)}{g(x)}\right)' = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}$$

C.4 Modèle Linéaire Généralisé : généralités

Après avoir exposé les raisons pour lesquelles il est opportun d'utiliser un modèle logistique dans cette étude, il faut maintenant s'assurer des hypothèses prises. La régression logistique appartient à la famille des modèles linéaires généralisés (GLM) qui se décomposent comme il a été présenté auparavant en trois éléments. Ils sont d'après Anderson et al. (2010) :

- **une partie aléatoire** : c'est la variable réponse Y , elle doit être identifiée et sa distribution de probabilité être spécifiée. Ici Y est le taux de rachat.
- **une partie systématique** : c'est un ensemble de variables explicatives X dont les coefficients correspondant sont les β . Une observation est notée x_{ij} où i est la i^e observation et j la j^e variable dans l'équation. Les variables explicatives entrent de manière linéaire, par exemple $\beta_0 + x_{11}\beta_1 + x_{12}\beta_2 + \dots + x_{1J}\beta_J$ où β_0 représente une constante.
- **une fonction lien** : c'est un aspect important du modèle. Elle est utilisée pour spécifier la relation entre la valeur attendue de la partie aléatoire et de la partie systématique. Le GLM est écrit par l'équation suivante : $g(Y) = X\beta + \varepsilon$ ou d'une autre manière $Y = g^{-1}(X\beta + \varepsilon)$. $g(\cdot)$ représente la fonction de lien et ε représente le terme d'erreur qui suit une loi $\mathcal{N}(0, \sigma^2)$.

En général il y a deux méthodes différentes dans les travaux académiques pour modéliser les taux de rachat. La différence repose sur la fonction de lien et est essentiellement dépendante de l'agrégation utilisée pour les données recueillies.

C.4.1 Niveau contrat - variable binaire

Dans beaucoup de cas la variable des rachats est observée contrat par contrat et prend la valeur 0 ou 1, symbolisant le rachat ou le non rachat. Ce résultat binaire rend certaines méthodes inadéquates.

Les académiciens choisissent habituellement une des fonctions de lien avec ce type de variable réponse :

- la fonction Logit : $g(Y) = \ln\left(\frac{Y}{1-Y}\right)$,
- le modèle Probit : $g(Y) = \phi^{-1}(Y)$ où ϕ^{-1} est la fonction de distribution inverse de la loi Normale standard,
- Log-linéaire/Poisson : $g(Y) = \ln(Y)$.

L'application d'une de ces fonctions donnera une estimation qui peut être différente de 0 ou 1. C'est pourquoi un seuil est choisi qui permettra de voir si la prédiction est au dessous ou au dessus de celui-ci. Il permettra d'arrondir la valeur prédite à 0 ou 1.

Le modèle de Poisson cité suppose qu'une variation en pourcentage des variables explicatives a un effet multiplicatif et que la variance est égale à la moyenne. Le modèle de Poisson est un choix adéquat afin de prédire des événements extrêmes. Il est utilisé lorsque les taux de rachat sont attendus proches de 0 et lorsque les résultats du modèle sont présentés de manière qualitative plutôt que quantitative (Cerchiara et al. (2008)).

Comme le démontre Anderson et al. (2010), il existe une petite différence dans les résultats entre les fonctions de *logit* et de *probit*, entre autre la fonction *logit* donne des résultats plus précis. Kim et al. (2005) montrent également que la fonction *logit* est généralement meilleure que les modèles de survie existants qui utilisent un seul paramètre pour l'estimation des taux de rachat. A ces éléments la fonction *logit* semble être la méthode la plus adaptée à la modélisation des taux de rachat. Guszczka (2005) a comparé trois modèles avec la régression logistique et a conclu que la régression logistique est meilleure que le modèle **CART** mais moins bonne que le modèle sophistiqué **MARS**.

Un des avantages de la fonction *logit* est qu'elle permet de calculer un rapport de chance ou *odds-ratio* (Anderson et al. (2010)). Il indique la chance d'opter pour un choix plutôt qu'un autre. Le facteur d'échelle augmente l'interprétation et l'application du modèle.

C.4.2 Données agrégées - variable continue

Peu d'articles sur les taux de rachat dans le domaine de l'assurance indique quelle fonction lien est utilisée pour analyser la variable de rachat continue. Les fonctions lien courantes sont :

- régression linéaire : $g(Y) = Y$,
- régression logistique : $g(Y) = \log(Y)$; cette fonction est la même que celle pour des données binaires,
- régression log-log complémentaire : $g(Y) = \ln(-\ln(1 - Y))$.

La différence avec des données binaires est que les valeurs prédites ne sont pas arrondies à 0 ou 1. Choisir une fonction lien dans ce cas peut se faire par le biais d'un diagramme de dispersion, ou une mesure statistique, par exemple le coefficient de corrélation de Pearson. Lorsque le choix est fait d'appliquer un GLM mais qu'il est difficile d'identifier la distribution des données il est courant de commencer sans fonction lien. Le GLM particulier est en fait connu comme le modèle de régression linéaire.

C.5 Du modèle logit au modèle de choix discrets

Selon le mathématicien autodidacte B. Gompertz, deux causes distinctes peuvent provoquer la mort :

- l'une, le hasard, l'accident qui peut priver de la vie une personne en parfaite santé, cause dont l'intensité est indépendante de l'âge,
- l'autre, qu'il nommait *force of mortality*, est l'affaiblissement progressif de l'individu.

À la différence des lois de survie décrivant un fait subi, la construction d'une loi de rachat peut être vue comme le résultat d'un fait générateur à savoir un choix individuel.

Les modèles de choix discrets décrivent le comportement d'un individu face à un ensemble d'actions mutuellement exclusives. Les économistes supposent généralement

que les individus choisissent toujours l'alternative qu'ils préfèrent de façon déterministe. Néanmoins il semble que les choix des individus peuvent fluctuer selon le contexte, c'est pourquoi les modèles de choix discrets adoptent une démarche probabiliste plutôt que déterministe. Les choix des individus sont par conséquent soumis à un certain aléa qui conduit à calculer des probabilités de choix pour chaque alternative et à confronter ces probabilités aux fréquences observées.

Ben-Akiva and Lerman (1985) indique que le choix peut être vu comme le résultat d'un processus séquentiel de prises de décisions suivant les étapes :

1. définition du problème de choix,
2. mise en place d'alternatives,
3. évaluation des caractéristiques des alternatives,
4. choix,
5. réalisation.

Ceci montre qu'un choix n'est pas réduit au seul fait d'une décision à un moment donné mais est plutôt le résultat d'un processus. Les choix sont alors faits sur les caractéristiques ou attributs des alternatives. Dans le but de prendre une décision, l'individu a besoin d'une règle : l'attractivité de l'alternative peut en être une. Elle renvoie souvent à l'utilité. Cette mesure, que le décideur essaie de maximiser, est utilisée dans les modèles récents.

Le terme de comportement rationnel est basé sur les croyances qu'un individu a du gain qu'il en retirerait s'il prenait cette décision. Cela semble évident que différents individus ont différentes croyances. Par conséquent il n'y a pas un seul type de comportement rationnel. C'est pourquoi la rationalité n'est pas très facile à mettre en œuvre pour décrire le comportement de choix individuels. Un autre concept décrit dans la littérature est l'opposé de l'impulsivité : un individu ne prendra pas une décision basée sur son état psychologique. Il suivra un processus cohérent et calculé, ce qui ne veut pas dire que cela ne répondra pas à ses propres objectifs. La rationalité est un concept ambigu et nécessite donc d'introduire un jeu de règles capables de la représenter. Pour simplifier, il suffit de faire l'hypothèse que pour un preneur de décision si l'alternative A est plus rentable que l'alternative B , il choisira tout le temps A .

Le modèle *logit* peut être appliqué à la théorie économique des comportements individuels en posant les hypothèses qui suivent. Soit un individu i pouvant choisir entre deux options possibles, notées 0 et 1. Sa décision s'appuie sur le modèle sous-jacent suivant, appelé modèle d'utilité stochastique additive (en anglais : *additive random utility model – ARUM*) : l'utilité U_{ji} qu'il retire (ou retirerait) de l'option j , où j peut prendre la valeur 0 ou 1, est la somme d'une composante *déterministe* V_{ji} et d'une composante aléatoire u_{ji} :

$$U_{ji} = V_{ji} + u_{ji} \quad (\text{C.9})$$

La première est nommée ainsi car elle est entièrement déterminée ou expliquée par un ensemble de caractéristiques individuelles observées et notées x_i : $V_{ji} = V_j(x_i)$. La forme généralement retenue de la fonction $V_j(x)$ est linéaire en x : $V_j(x_i) = x_i\beta_j$. La seconde composante rassemble les variables inobservées et inobservables qui peuvent jouer sur la décision de l'agent i . Elle est supposée varier de manière aléatoire. En résumé, l'utilité que i retire(raît) de l'option j s'écrit :

$$U_{ji} = x_i\beta_j + u_{ji} \quad (\text{C.10})$$

La règle de décision est alors la suivante : l'individu choisit une des deux options si l'utilité qu'il en retire est supérieure à l'utilité attendue de l'autre option. Si y est la variable binaire repérant l'option choisie (i.e. $y_i = 0$ si l'individu i a choisi l'option 0, et $y_i = 1$ si i a choisi 1), alors :

$$\begin{cases} y_i = 0 & \iff U_{0i} > U_{1i} \\ y_i = 1 & \iff U_{1i} > U_{0i} \end{cases} \quad (\text{C.11})$$

En introduisant les caractéristiques observées x et en remplaçant l'utilité par son expression (C.10) :

$$\begin{aligned}
P(y_i = 1|x) &= P(V_1(x_i) + u_{1i} > V_0(x_i) + u_{0i}) \\
&= P(u_{0i} - u_{1i} < V_1(x_i) - V_0(x_i)) \\
&= P(u_{0i} - u_{1i} < x_i(\beta_1 - \beta_0)) \\
&= P(u_{0i} - u_{1i} < x_i\beta)
\end{aligned} \tag{C.12}$$

où $\beta = \beta_1 - \beta_0$. Finalement :

$$P(y_i = 1|x) = G(x_i\beta) \tag{C.13}$$

où G est la fonction de répartition de la loi $u_0 - u_1$. Si u_0 et u_1 suivent la loi dite *type I extreme-value* ou loi de Gumbel⁵, dont la fonction de densité s'écrit $f(u) = e^{-u} \exp[-e^{-u}]$, alors il s'agit de l'expression du modèle logit déjà rencontrée :

$$G(x_i\beta) = \frac{1}{1 + e^{-x_i\beta}} \tag{C.14}$$

Ce cadre théorique est adapté aux cas où les deux options entre lesquelles l'individu doit trancher ont elles-mêmes des caractéristiques qui font partie des critères de décision.

Le cas où des caractéristiques des options entrent dans les critères de décision, se formalise de la manière ci-après. Soit z_{ji} , pour $j = 0$ ou 1 , les caractéristiques pour l'individu i . En supposant qu'elles agissent de manière additive sur l'utilité, l'expression (C.10) devient :

$$U_{ji} = x_i\beta_j + z_{ji}\gamma + u_{ji} \tag{C.15}$$

5. La loi de Gumbel est une loi de probabilité continue dont les applications à l'économie sont réservées à quelques domaines particuliers, notamment la finance de marché et les assurances. Un modèle fréquentiel très souvent utilisé pour décrire le comportement statistique des valeurs extrêmes est la distribution statistique de Gumbel.

En appliquant toujours la même règle de décision (C.11), la probabilité de prendre l'option 1 s'écrit maintenant :

$$P(y_i = 1|x_i, z_{0i}, z_{1i}) = G[x_i(\beta_1 - \beta_0) + (z_{1i} - z_{0i})\gamma] = \frac{1}{1 + e^{-x_i\beta - (z_{1i} - z_{0i})\gamma}} \quad (\text{C.16})$$

Formalisé ainsi, le modèle est parfois appelé modèle logit conditionnel selon la dénomination que lui a donné McFadden (1973) (conditional logit model). Il contient en effet, en plus des caractéristiques individuelles x , des variables - les z_j - qui varient avec l'option proposée. De plus, et surtout, les z_j sont les variables du modèle à privilégier dans l'analyse. Les caractéristiques individuelles sont introduites d'abord pour prendre en compte l'hétérogénéité observée des individus. L'utilisation de ce type de modèle est d'estimer l'impact sur les comportements d'une modification des z_j .

Dans le cas de l'étude sur les comportements de rachats, les caractéristiques individuelles sont les variables x de l'âge, de l'ancienneté, *etc.* énoncées plus haut et les caractéristiques économiques les z_j . z_0 est la valeur du TME et z_1 la valeur du taux servi $N - 1$.

En définitive, la régression logistique comprend les covariables structurelles âge, encours, ancienneté, *etc.* et une unique covariable conjoncturelle *delta taux*. Or cette dernière demeure la plus sensible car ses variations influencent le taux de rachat structurel. Par conséquent, il semble raisonnable de privilégier dans un premier temps la construction de cette variable puis de déterminer quelle différence de taux met en évidence la prise de décision du rachat.

Afin de construire la variable *delta taux* qui reflétera au mieux le comportement des individus, la méthode utilisée est définie à partir de la qualité du modèle. La rationalité est différente d'un individu à l'autre, ce qui entraîne des décalages entre la prise de décision et celle qui aurait dû être observée au moment opportun. Il faut par conséquent déterminer une variable *delta taux* qui tentera de rassembler tous ces comportements plus ou moins rationnels en fonction du contexte économique. Le premier modèle comporte alors uniquement la covariable expliquant les rachats conjoncturels.

Évaluation du modèle à choix discrets

Il existe différentes manières d'évaluer l'efficacité d'un modèle de prédiction, par exemple :

- confronter les valeurs observées de la variable dépendante Y avec les prédictions \hat{Y} ,
- comparer les vraies valeurs des probabilités π avec celles prédites par le modèle $\hat{\pi}$.

En effet, la régression logistique permet de fournir une bonne approximation de cette quantité et peut se révéler très utile lorsque le souhait est de classer les individus selon leurs degrés de positivité ou introduire d'autres calculs ultérieurement (exemple intégrer les coûts de mauvais classement).

D.1 Évaluation de la régression

Les résultats de la régression logistique permettent de construire une matrice de confusion (en anglais *classification table*). Elle confronte toujours les valeurs observées de la variable dépendante avec celles prédites, puis comptabilise les bonnes et les mauvaises prédictions. Son intérêt est de permettre à la fois d'appréhender la quantité de l'erreur (le

taux d'erreur) et de rendre compte de la structure de l'erreur (la manière dont le modèle se trompe).

$Y \times \hat{Y}$	$\hat{+}$	$\hat{-}$	Total
+	a	b	a+b
-	c	d	c+d
Total	a+c	b+d	n=a+b+c+d

TABLE D.1 – Matrice de confusion - Forme générique

Dans un problème à deux classes (+ contre - ou 1 contre 0), à partir de la forme générique de la matrice de confusion D.1, de la présente page, plusieurs indicateurs peuvent être déduits pour informer sur la concordance entre les valeurs observées et les valeurs prédites. Les principaux ratios sont les suivants :

- **a** sont les vrais positifs c'est-à-dire les observations qui ont été classées positives et qui le sont réellement,
- **c** sont les faux positifs c'est-à-dire les individus classés positifs et qui sont en réalité des négatifs,
- de la même manière, **b** sont les faux négatifs et **d** sont les vrais négatifs,
- le **taux d'erreur** est égal au nombre de mauvais classement rapporté à l'effectif total c'est-à-dire :

$$\varepsilon = \frac{b + c}{n} = 1 - \frac{a + d}{n} \quad (\text{D.1})$$

Il estime la probabilité de mauvais classement du modèle.

- le **taux de succès** correspond à la probabilité de bon classement du modèle, c'est le complémentaire à 1 du taux d'erreur :

$$\theta = \frac{a + d}{n} = 1 - \varepsilon \quad (\text{D.2})$$

- la **sensibilité** (ou le *rappel*, ou encore le *taux de vrais positifs* [TVP]) indique la capacité du modèle à retrouver les positifs :

$$S_e = \text{Sensibilité} = \text{TVP} = \text{Rappel} = \frac{a}{a + b} \quad (\text{D.3})$$

- la **précision** indique la proportion de vrais positifs parmi les individus qui ont été classés positifs :

$$précision = \frac{a}{a + c} \quad (D.4)$$

Elle estime la probabilité d'un individu d'être réellement positif lorsque le modèle le classe comme tel.

- la **spécificité**, à l'inverse de la sensibilité, indique la proportion de négatifs détectés :

$$S_p = \text{Spécificité} = \frac{d}{c + d} \quad (D.5)$$

- le **taux de faux positifs** (TFP) correspond à la proportion de négatifs qui ont été classés positifs c'est-à-dire :

$$\text{TFP} = \frac{c}{c + d} = 1 - \text{Spécificité} \quad (D.6)$$

- la **F-Mesure** est très utilisée en recherche d'information. Elle synthétise (moyenne harmonique) le rappel et la précision, l'importance accordée à l'une ou à l'autre est paramétrable avec β :

$$F_\beta = \frac{(1 + \beta^2) \times \text{rappel} \times \text{Précision}}{\beta^2 \times \text{Précision} + \text{Rappel}} \quad (D.7)$$

Lorsque

- $\beta = 1$ est la valeur usuelle, on accorde la même importance au rappel et à la précision, la **F-Mesure** devient :

$$F_{\beta=1} = \frac{2 \times \text{Rappel} \times \text{Précision}}{\text{Précision} + \text{Rappel}} \quad (D.8)$$

- $\beta < 1$, plus d'importance est accordée à la précision par rapport au rappel. Une valeur fréquemment utilisée est $\beta = 0,5$.
- $\beta > 1$, plus d'importance est accordée au rappel par rapport à la précision. Une valeur fréquemment rencontrée est $\beta = 2$.

La **F-Mesure** est une moyenne harmonique entre le rappel et la précision, en effet elle peut s'écrire de la manière suivante

$$F = \frac{1}{\alpha \frac{1}{Precision} + (1 - \alpha) \frac{1}{Rappel}} \quad (\text{D.9})$$

où $\beta^2 = \frac{1-\alpha}{\alpha}$.

D.2 La courbe ROC

La courbe **ROC** (en anglais *Receiver Operating Characteristic*) est un outil très riche. Son champ d'application dépasse largement le cadre de l'apprentissage supervisé¹. Elle est par exemple très utilisée en épidémiologie. Elle présente dans le cas présent des caractéristiques très intéressantes pour l'évaluation et la comparaison des performances des classifieurs :

- elle propose un outil graphique qui permet d'évaluer et de comparer globalement le comportement des classifieurs,
- elle est indépendante des coûts de mauvaise affectation. Elle permet par exemple de déterminer si un classifieur surpasse un autre, quelle que soit la combinaison de coûts utilisée,
- elle est opérationnelle même dans le cas des distributions très déséquilibrées,
- enfin, un indicateur synthétique peut lui être associé : le critère **AUC** (aire sous la courbe, en anglais *area under curve*).

La courbe **ROC** met en relation le taux de vrais positifs **TVP** (la sensibilité, le rappel) et le taux de faux positifs **TFP** ($\text{TFP} = 1 - \text{Spécificité}$) dans un graphique nuage de points (exemple de la figure D.1 page ci-contre).

1. L'apprentissage supervisé désigne un type d'algorithme d'apprentissage automatique qui utilise un ensemble de données connues (appelé ensemble de données d'apprentissage) pour réaliser des prévisions. L'ensemble de données d'apprentissage contient des données d'entrée et des valeurs de réponses. À partir de ces éléments, l'objectif de l'algorithme d'apprentissage supervisé consiste à créer un modèle capable de réaliser des prévisions sur les valeurs de réponse pour un nouvel ensemble de données. Un ensemble de données de test est souvent utilisé pour valider le modèle.

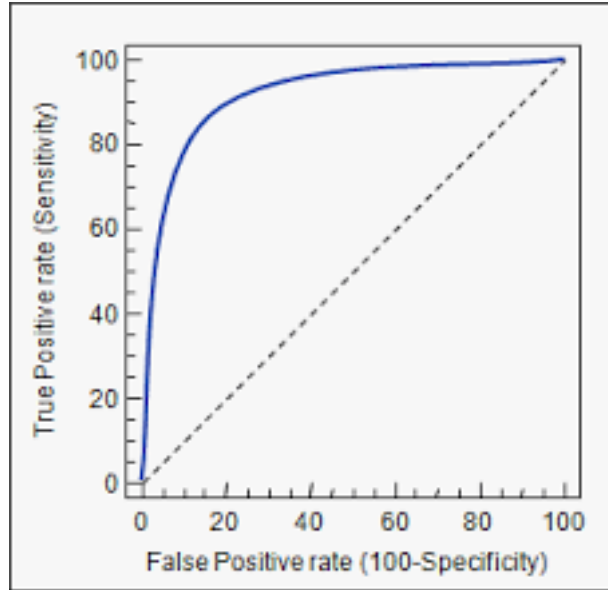


FIGURE D.1 – La courbe **ROC**

Habituellement, $\hat{\pi}$ est comparé à un seuil $s = 0,5$ pour effectuer une prédiction \hat{y} . La matrice de confusion permet d’extraire les deux indicateurs précités. La courbe **ROC** généralise cette idée en faisant varier s sur tout le continuum des valeurs possibles entre 0 et 1. Pour chaque configuration, la matrice de confusion peut être obtenue et les valeurs de **TVP** et **TFP** calculées.

Dans la pratique, il n’est pas nécessaire de construire explicitement la matrice de confusion et il est possible de procéder ainsi :

- calculer le score $\hat{\pi}$ de chaque individu à l’aide du modèle de prédiction,
- trier le fichier selon un score décroissant,
- considérant qu’il n’y ait pas d’ex-æquo chaque valeur du score peut être potentiellement un seuil s . Pour toutes les observations dont le score est supérieur ou égal à s , pour les individus dans la partie haute du tableau, il est possible de comptabiliser le nombre de positifs $n_+(s)$ et le nombre de négatifs $n_-(s)$. Alors $\text{TVP} = \frac{n_+(s)}{n_+}$ et $\text{TFP} = \frac{n_-(s)}{n_-}$.
- la courbe **ROC** correspond au graphique de nuage de points qui relie les couples **(TVP, TFP)**. Le premier point est forcément $(0,0)$, le dernier est $(1,1)$.

La courbe **ROC** peut être caractérisée en calculant la surface située sous la courbe. C'est le critère **AUC**. Elle exprime la probabilité de placer un individu positif devant un négatif. Ainsi, dans le cas d'une discrimination parfaite, les positifs sont sûrs d'être placés devant les négatifs ce qui implique que **AUC** = 1. A contrario, si le classifieur attribue des scores au hasard, il y a autant de chances de placer un positif devant un négatif que l'inverse. La courbe **ROC** se confond ainsi avec la première bissectrice **AUC** = 0,5. C'est la situation de référence, le classifieur doit faire mieux.

D.2.1 La courbe ROC par la méthode binomiale

La courbe empirique, celle obtenue en associant la sensibilité avec 1-spécificité, est largement utilisée parce qu'elle ne pose aucune hypothèse sur la distribution des prédictions individuelles. Mais il peut être intéressant de travailler à partir d'une courbe lissée. Une manière habituelle d'atteindre ce résultat est de revenir au modèle Binomial vu à la section C.1, page 92.

Goncalves et al. (2014) posent X et Y deux variables aléatoires indépendantes, respectivement évaluant la mesure de diagnostic de test pour une population réalisant le rachat ($D=0$) et une population ne le réalisant pas ($D=1$). Sans perte de généralités, et pour un seuil approprié s , le résultat de test est positif si il est plus grand que c , négatif autrement. Sachant F et G les fonctions de distribution de X et Y respectivement. La sensibilité du test est donnée par $Se(s) = 1 - G(s)$, et la spécificité définie comme $Sp(s) = F(s)$. Un exemple est donné à la figure D.2, page ci-contre.

La courbe **ROC** est définie par le graphique de $Se(s)$ versus $1 - Sp(s)$ pour $-\infty < s < +\infty$, ou par équivalence comme le tracé de :

$$ROC(t) = 1 - G(F^{-1}(1 - t))$$

pour $t \in [0, 1]$, où $F^{-1}(1 - t) = \inf \{x \in \mathbb{R} : F(x) \leq 1 - t\}$.

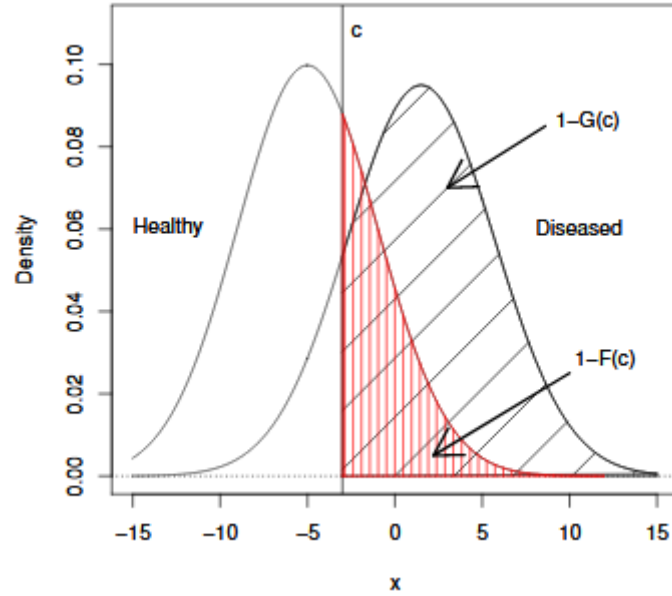


FIGURE D.2 – Exemple : Distribution des mesures de diagnostic de test pour une population en bonne santé et une population malade

La courbe **ROC** est croissante et invariante sous n'importe quelle transformation monotone croissante des variables X et Y . La valeur de l'**AUC** est donnée par :

$$\text{AUC} = \int_0^1 \text{ROC}(u) du$$

L'estimation paramétrique de la courbe **ROC** est un domaine de recherche actif et plusieurs propositions pour F et G peuvent être considérées. La plus répandue est donnée par le modèle binomial.

Le modèle binomial fait l'hypothèse que les individus présentant un rachat suivent une distribution normale de moyenne μ_1 et de variance σ_1^2 puis que les individus qui ne rachètent pas suivent aussi une distribution normale de moyenne μ_0 et de variance σ_0^2 . En utilisant la notation des courbes **ROC** empirique, $G(s) = \Phi((\mu_0 - s)/\sigma_0)$ il apparaît que le seuil s peut être écrit comme une fonction de x : $s = \mu_0 - \sigma_0\Phi^{-1}(x)$. Dès lors un seuil s correspond à la sensibilité et la forme fonctionnelle de la courbe **ROC** peut s'écrire de

la façon suivante :

$$F(t) = \Phi\left(\frac{\mu_1 - s}{\sigma_1}\right) = \Phi\left(\frac{\mu_1 - \mu_0 + \sigma_0\Phi^{-1}(x)}{\sigma_1}\right) = \Phi\left(a + b \times \Phi^{-1}(x)\right) \quad (\text{D.10})$$

où $a = \frac{\mu_1 - \mu_0}{\sigma_1}$ et $b = \frac{\sigma_0}{\sigma_1}$ sont souvent associés à des paramètres binomiaux. Ils sont quelque fois appelés ordonnée à l'origine et pente. L'aire sous la courbe du modèle binomial a une expression qui peut être écrite sous une forme fermée :

$$\text{AUC} = \Phi\left(\frac{a}{\sqrt{1 + b^2}}\right) \quad (\text{D.11})$$

Pour estimer un modèle binomial à partir d'un jeu de données il suffit d'évaluer les moyennes et les variances séparément. Ceci est réalisé grâce à la procédure **PROC MEANS** du logiciel **SAS** en demandant un traitement par classe. Toutefois, cette opération peut aussi être obtenue à l'aide de **PROC NLMIXED**² comme ceci :

```
1  proc nlmixed data=toto;
2  parameters m1=0 m0=0 s1=1 s0=1;
3  if rachat=1 then m=m1;else if rachat=0 then m=m0;
4  if rachat=1 then s=s1**2;else if rachat=0 then s=s0**2;
5  a=(m1-m0)/s1;
6  b=s0/s1;
7  model delta_taux ~ normal(m,s);
8  estimate 'a' a;
9  estimate 'b' b;
10 estimate 'AUC' probnorm(a/sqrt(1+b**2));
11 run;
```

2. La procédure **PROC NLMIXED** permet de manipuler des modèles mixtes non linéaires c'est-à-dire des modèles où les effets systématiques et aléatoires entrent de façon non linéaire. Elle permet de spécifier des distributions conditionnelles à partir des données (sachant l'effet aléatoire) non plus sous une forme standard (normale, binomiale, Poisson) ou comme une distribution générale mais sous un code.

L'avantage de l'utilisation de la procédure **PROC NLMIXED** s'établit via la requête **ESTIMATE** puisque les estimations sont obtenues directement. À partir de cette méthode d'approximation, Gonen (2006) obtient dans son article une aire sous la courbe **ROC** empirique de 87%, un peu plus élevée que la valeur de 86% avec la courbe **ROC** binomiale. L'intervalle de confiance est très similaire (82%, 92%) contre (81%, 92%). Il constate par conséquent que l'analyse binomiale conduit à des résultats similaires à celle de l'analyse empirique de la courbe. Elle apparaît, par conséquent, être un lissage pertinent de la courbe **ROC**.

D.2.2 La courbe ROC par la valeur c

Le logiciel **SAS** présente dans ses sorties des critères globaux d'adéquation. Quatre statistiques sont ainsi présentes :

- somme des fréquences = $\sum_i f_i$,
- somme des pondérations = $\sum_i w_i$,
- log vraisemblance = $\sum_i f_i w_i \log(\hat{\pi}_i)$,
- taux de mauvaise classification = $\frac{\sum_i 1\{F_Y_i \neq I_Y_i\} f_i}{\sum_i f_i}$.

où F_Y_i est la valeur de la variable réponse Y_i et I_Y_i est sa valeur prédite.

Il est démontré dans Hanley and McNeil (1982) qu'une mesure équivalente à la valeur de **AUC** est la statistique c , que la procédure **PROC LOGISTIC** calcule (pour des données non pondérées) et affiche dans ses sorties.

L'aire sous la courbe **ROC** permet d'obtenir un indicateur de la précision du diagnostic de test. Pepe (2000) fait remarquer que l'**AUC** peut être interprétée comme la probabilité que le test est davantage le résultat d'individu malade pris au hasard est plus illustratrice d'une maladie qu'un individu non malade pris au hasard.

Cette interprétation ou équivalence, traitée aussi par Hanley and McNeil (1982), porte l'attention sur les distributions du résultat du test chez les personnes malades ou non malades. Si les deux distributions sont clairement distinctes, la probabilité sera proche de 1 ; mais si elles sont centrées sur la même valeur, la probabilité sera de $1/2$. Dans un contexte de régression logistique le résultat de test est la probabilité de prédire l'événement.

La définition de la valeur \mathbf{c} fait intervenir les paires d'observations concordantes et discordantes. Il faut commencer par former toutes les paires pour lesquelles dans le premier cas il s'agit d'un événement et dans le second d'un non événement. Le nombre total de paires est noté t .

Une paire est dite concordante si un événement observé sera aussi classé par le modèle comme un événement.

Une paire est dite discordante si un événement observé sera classé par le modèle comme un non événement.

Si le modèle ne parvient pas à classer l'observation la paire sera dite liée (en anglais *tied*). Le nombre de paires concordantes est noté nc et le nombre de paires discordantes nd . Alors le nombre de paires liées est $t - nc - nd$. Dans un échantillon non pondéré la formule de c est la suivante :

$$c = \frac{nc + 0,5 \times (t - nc - nd)}{t} \quad (\text{D.12})$$

La procédure **PROC LOGISTIC** du logiciel **SAS** prend en compte les observations pondérées mais ne les considère pas ou ne les utilise pas convenablement, dans le calcul de la valeur de \mathbf{c} . Le but est alors de le calculer à travers certaines étapes comme il est fait dans Izrael et al. (2003). Le principe est de déterminer chaque élément nc , nd et t de la formule (D.12).

Si $\hat{\pi}_i$ est plus petit que $\hat{\pi}_j$ alors la paire est discordante et leur pondération $w_i \times w_j$ est ajoutée au poids total des paires discordantes (nd).

Enfin, si la paire n'est ni concordante ni discordante le produit de $w_i \times w_j$ est ajouté au poids total des paires liées. Sachant W_E la somme totale pondérée des événements et W_N la somme totale pondérée des non événements, le nombre total pondéré de paires est calculé comme $t = W_E \times W_N$. Basé sur la somme totale des pondérations soit les $E \times N$ comparaisons, le logiciel **SAS** calcule \mathbf{c} par la formule (D.12) de la présente page.

D.3 Résultats du modèle complet

Le tableau D.2, de la présente page, présente les résultats³ des coefficients de la régression logistique, les β . Il faut cependant noter que le logiciel **SAS** ne parvient pas à calculer une valeur correcte de la variance des estimateurs dans le cas pondéré. Néanmoins la valeur des β reste interprétable.

Paramètre	Libellé	Estimation	Erreur type
Intercept		-7,2635	0,0559
lib_age	a) inf 30ans	-0,2348	0,1083
lib_age	b) 30ans - 49ans	0,1846	0,0616
lib_age	c) 50ans - 59ans	0,0786	0,0622
lib_age	e) sup 70ans	0,1899	0,0442
lib_anc	a) 0an - 3ans	0,1368	0,0501
lib_anc	b) 4ans - 7ans	-0,3014	0,0499
lib_anc	d) 12ans - 15ans	0,0340	0,0417
lib_anc	e) 16ans - 19ans	0,1489	0,0487
lib_anc	f) sup 20ans	-0,2514	0,0734
top_tmg	a) TMG nul	0,1639	0,0399
top_tmg	b) TMG non nul	-0,0625	0,0369
<i>delta_taux</i>		-0,4692	0,0288

TABLE D.2 – Estimations des coefficients de régression par l’analyse du maximum de vraisemblance du modèle logistique pondéré

L’estimation du paramètre **a) inf 30 ans** présente une valeur de -0,2348, ce qui revient à dire que le log-odds-ratio d’un rachat versus un non rachat si l’adhérent a moins de 30 ans versus entre 60 et 69 ans est de -0,2348.

La valeur $\exp(-0,2348) = 0,79$ est l’estimation du odds-ratio. Celui-ci peut être interprété de la manière suivante : les adhérents de moins de 30 ans ont moins tendance à racheter leur contrat que ceux ayant un âge compris entre 60 et 69 ans. Le modèle prévoit alors que les tranches d’âge b), c) et e) ont tendance à plus racheter que la tranche d).

Une prédisposition supérieure à racheter de la tranche **b) 30 ans - 49 ans** pourrait s’expliquer par une mauvaise vente ou un projet qui se réaliserait à court terme et dont

3. La tranche d’âge d) 60 ans - 69 ans, celle de l’ancienneté c) 8 ans - 11 ans et l’indicatrice c) TMG > Taux_servi ont été omises dans la sortie du logiciel **SAS** du fait qu’elles ont été implémentées comme des valeurs références.

le produit d'épargne serait un support d'attente rémunéré. Cette hypothèse pourrait être liée au cas d'un premier achat immobilier suite à une libéralité faite par testament.

Le faible odds-ratio de la tranche d'anciennetés **b) 4 ans - 7 ans** provient du l'effet de la fiscalité des rachats. Le modèle prévoit que plus longtemps l'adhérent détient son contrat moins il n'a d'intention de racheter après le pic de l'ancienneté **e) 16 ans - 19 ans**. Ces résultats démontrent donc une certaine fidélité des adhérents à la Mutuelle.

Une valeur du TMG semble freiner la décision de rachat de son contrat, ce qui semble cohérent sachant que la Mutuelle rémunère de moins en moins ses encours sur le fonds Euro.

Pour finir avec l'interprétation des paramètres du modèle, le coefficient de -0,4692, qui représente un odds-ratio de 0,63, devant la covariable *delta taux* démontre que plus l'écart entre le taux servi et le taux attendu est important moins les individus sont enclin à exercer leur option.

Le tableau D.3, de la présente page, présente les différentes valeurs de l'**AUC** selon le modèle étudié. Le modèle complet et le modèle structurel possèdent une valeur plus faible que dans le cas non pondéré. Ceci est causé par la construction d'une nouvelle courbe de *delta taux* en montant. Cette baisse n'est pas très inquiétante car le but de cette étude est bien de maximiser le classifieur pour le modèle conjoncturel et non de prédire au plus exact un comportement complet.

	Zone	Erreur Type	95% Intervalle de confiance de Wald	
Modèle	0,6163	0,00554	0,6055	0,6272
Conjoncturel	0,5813	0,00570	0,5702	0,5925
Structurel	0,5722	0,00548	0,5615	0,5829

TABLE D.3 – Valeurs de l'aire sous la courbe ROC du modèle logistique complet pondéré et marginaux

La courbe **ROC** du modèle conjoncturel à la figure D.3, page ci-contre, sera utilisée afin de déterminer un seuil à partir duquel les taux de rachats structurels sont atténués par les taux de rachats conjoncturels.

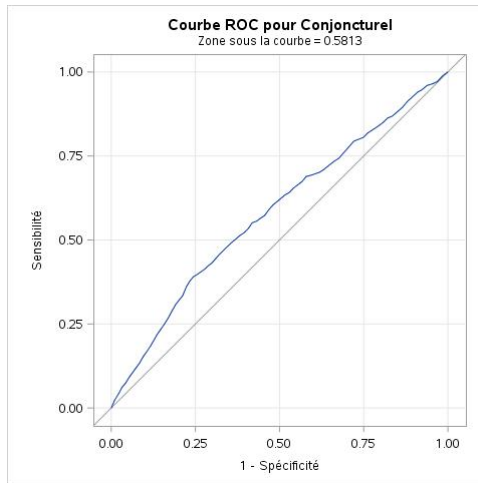


FIGURE D.3 – Courbe ROC du modèle conjoncturel (variable *delta taux*)

Indice de Youden

Youden (1950) a suggéré une statistique J dans le but de synthétiser la performance d'un test de diagnostic lors d'une régression logistique. Sa valeur est comprise entre -1 et 1 . Il est égal à 0 lorsque le classifieur donne la même proportion de résultats positifs pour un groupe avec et sans maladie. Ainsi le modèle n'est pas discriminant. Une valeur de 1 indique qu'il n'y a pas de faux positifs ou de faux négatifs, ce qui signifie que le classifieur est parfait.

L'indice de Youden est bien connu en biostatistique¹ mais moins en apprentissage supervisé. Il s'écrit sous la forme suivante :

$$\begin{aligned} J &= S_e + S_p - 1 \\ &= \frac{a}{a+b} + \frac{d}{c+d} - 1 \\ &= TVP - TFP \end{aligned} \tag{E.1}$$

Son mérite est de caractériser le classifieur selon la sensibilité et la spécificité. Il est déterminant dans le calcul du seuil car maximiser l'équation (IV.1), page 44, revient alors à maximiser l'équation (E.1) de l'indice de Youden.

Cette statistique mesure la distance du seuil au dessus de la diagonale qui représente le modèle qui n'apporte aucune information à la classification. L'utilisation d'un seul

1. La biostatistique est un champ scientifique constitué par l'application de la science statistique à la biologie et à la médecine.

indicateur n'est généralement pas recommandé. Mais cet indice symbolise le terme de la connaissance, qui représente la probabilité d'une décision informée (opposée à une décision prise au hasard), et prend en compte toutes les prédictions.

En effet, Powers (2007) généralise cet indice au cas multinomial en invoquant l'effet de la connaissance (en anglais *informedness*). Pour représenter ce que l'auteur a voulu dire, il faut prendre l'exemple d'un parieur/bookmaker. Le but d'un bookmaker est d'anticiper la valeur de ses gains ou de ses pertes en fonction des probabilités de réalisation de ces événements. Powers (2007) définit alors le concept de connaissance. Il représente l'avantage qu'a un parieur lorsque des gains sont réalisés. Une tarification juste basée sur des probabilités mène à un jeu à somme nulle, c'est-à-dire que deviner n'apportera aucun gain à long terme, tandis qu'un parieur avec une certaine connaissance gagnera tout le temps. La connaissance est la probabilité qu'un parieur a de réaliser un pari informé et qui expliquera qu'avec le temps son avantage est avéré. En d'autres termes ce n'est pas le résultat d'un jeu de hasard. Il propose la définition suivante :

DÉFINITION 1

La connaissance est la manière dont un prédicteur est informé lors d'une situation spécifique, et caractérise la probabilité d'une prédiction basée sur des informations par rapport aux conditions de sa réalisation (contre chance).

La connaissance est une bonne mesure de la cohérence de la prédiction car elle combine les indicateurs d'efficacité sur le niveau de la proportion des observations bien classées. Pour obtenir la valeur de l'indice de Youden il faut cependant récupérer les valeurs des paramètres du modèle logistique contenant uniquement la covariable *delta taux* (tableau (E.1), de la présente page).

Paramètre	Estimation	Erreur type	Pr > Khi-2
Intercept	-7,3555	0,0344	<,0001
<i>delta taux</i>	-0,3953	0,0266	<,0001

TABLE E.1 – Estimations par l'analyse du maximum de vraisemblance des paramètres du modèle conjoncturel pondéré

La table de sortie du logiciel **SAS** n'affiche pas les valeurs de *delta taux* mais seulement le résultat de la régression. Il faut alors inverser le modèle logistique afin de retrouver les

valeurs prises en entrée par **SAS**. A partir des valeurs de la probabilité estimée il est facile de calculer le logit grâce à la formule :

$$\text{logit}(\hat{\pi}_i) = \log\left(\frac{\hat{\pi}_i}{1 - \hat{\pi}_i}\right)$$

Ce qui revient à :

$$\text{Cutoff} = \frac{\text{logit}(\hat{\pi}_i) + 7,3555}{-0,3953}$$

Le logiciel **SAS** fournit les coordonnées de la courbe **ROC** c'est-à-dire S_e et $1 - S_p$. Il faut alors calculer la spécificité pour revenir sur l'indice de Youden.

```
1 data cutoff; set logistic;
2 logit=log(_prob_/(1-_prob_));
3 Cutoff=(logit+7,3555)/(-0,3953);
4 Sensitivity=_sensit_;
5 Specificity=1-_1mspec_;
6 YJ=Sensitivity+Specificity-1;
7 run;
```


Méthodes de lissage

L'observation agrégée des données permet de construire des taux de rachats en montrant en fonction uniquement de la variable conjoncturelle *delta taux*. Le principe qui va être utilisé repose sur l'idée de l'existence d'une tendance qui pourrait être extraite à partir des méthodes de lissage. Le paramètre du déclenchement de la baisse des rachats étant estimé avec le modèle logistique, il reste à déterminer le seuil à partir duquel les adhérents ne sont plus sensibles aux variations de *delta taux*. La question repose sur l'existence d'une valeur de cet indicateur pour laquelle la courbe des taux de rachats conjoncturels se stabiliserait.

Les méthodes d'estimation par moindres carrés pénalisés permettent un bon ajustement aux données et évitent de violentes ou rapides variations dans leurs restitutions. Leur principe repose sur le compromis fait entre la fidélité de la restitution des données observées et une nature lisse de leur ajustement. L'objectif est de trouver une bonne adéquation du lissage tout en introduisant une variation locale moyenne limitée.

Soit \mathbf{x}_i un vecteur d'une variable prédictive de dimension d issu d'une matrice $n \times d$ notée \mathbf{X} , \mathbf{z}_i un vecteur d'une variable prédictive de dimension p et y_i les observations associées avec $(\mathbf{x}_i, \mathbf{z}_i)$. En supposant que la relation entre \mathbf{z}_i et y_i est linéaire mais la relation entre \mathbf{x}_i et y_i est inconnue, l'ajustement des données est réalisé avec le modèle semi paramétrique suivant :

$$y_i = f(\mathbf{x}_i) + \mathbf{z}_i\beta + \varepsilon_i$$

où f est une fonction inconnue supposée suffisamment lisse, ε_i , $i = 1, \dots, n$, sont indépendants, les erreurs aléatoires de moyenne 0, et β est un vecteur de paramètres de dimension p .

Le modèle est scindé en deux parties. Les $\mathbf{z}_i\beta$ sont la partie paramétrique du modèle dont les \mathbf{z}_i sont appelées les covariables. La partie $f(\mathbf{x}_i)$ est celle dénommée non paramétrique dont les \mathbf{x}_i sont les variables lissées.

La méthode des moindres carrés ordinaires estime \mathbf{x}_i et β en minimisant la quantité :

$$\frac{1}{n} \sum_{i=1}^n (y_i - f(\mathbf{x}_i) - \mathbf{z}_i\beta)^2$$

L'ensemble des fonctions $f(\mathbf{x})$ est assez étendu : il est toujours possible de trouver une fonction f qui interpole les observations. Dans le but d'avoir une estimation assez proche des données et de permettre un certain degré de lissage, il suffit d'avoir recours à la méthode des moindres carrés pénalisés :

$$S_\lambda(f) = \frac{1}{n} \sum_{i=1}^n (y_i - f(\mathbf{x}_i) - \mathbf{z}_i\beta)^2 + \lambda J_2(f)$$

où $J_2(f)$ est la pénalité sur la forme de f . Elle est définie dans la plupart des cas comme l'intégrale de la racine carrée de la dérivée seconde de f .

Le premier terme mesure l'ajustement et le second mesure le niveau de lissage associé avec f . Le terme λ est le paramètre de lissage qui détermine le compromis entre le lissage et l'ajustement. Un β élevé signifie un ajustement brutal tandis qu'un β faible insiste davantage sur l'ajustement.

La fonction estimée f_λ peut être représentée comme une combinaison d'une séquence de fonctions de bases. Elle peut s'écrire de la manière suivante :

$$\hat{f}_\lambda(\mathbf{x}_i) = \theta_0 + \sum_{j=1}^d \theta_j \mathbf{x}_{ij} + \sum_{j=1}^p \delta_j B_j(\mathbf{x}_{ij})$$

où B_j est une fonction de base, qui dépend de la position des \mathbf{x}_i , et $\theta = \{\theta_0, \dots, \theta_d\}$ et $\delta = \{\delta_0, \dots, \delta_p\}$ sont les coefficients à estimer.

Pour λ fixé, les coefficients (θ, δ, β) peuvent être estimés en résolvant un système $n \times n$. Le paramètre de lissage est choisi en minimisant la fonction générale de validation croisée (GCV) : $\hat{\mathbf{y}} = \mathbf{A}(\lambda)\mathbf{y}$. alors $\mathbf{A}(\lambda)$ est la matrice de lissage et la fonction $GCV(\lambda)$ est définie comme :

$$GCV(\lambda) = \frac{(1/n)\|(\mathbf{I} - \mathbf{A}(\lambda))\mathbf{y}\|^2}{\left[(1/n)tr(\mathbf{I} - \mathbf{A}(\lambda))\right]^2}$$

L'utilisation des méthodes de lissage semble être une bonne approche dans la modélisation de la tendance. Cependant l'hypothèse sur laquelle repose l'estimation des coefficients de lissage est un effet fixe. Il semblerait ainsi plus approprié de travailler avec un effet aléatoire qui représenterait l'effet structurel. En fait il est assez difficile de connaître pour quelles valeurs de *delta taux* l'ampleur de l'impact des caractéristiques de l'adhérent ou du contrat influe sur l'effet d'une tendance conjoncturelle.

Modèle mixte

La structure de la base de données est construite à partir de la collecte d'observations répétées et définit alors la nature d'une étude longitudinale. Pour chaque individu considéré, les informations à son sujet sont recueillies à différents instants autrement dit répétées dans le temps. Le problème rencontré dans leur traitement statistique est la dépendance entre les variables.

L'agrégation des données permet de construire un taux de rachat en montant. L'idée est d'extraire le facteur fixe représenté par *delta taux* qui traduirait une tendance (effet de la conjoncture), et d'expliquer par des effets aléatoires c'est-à-dire les caractéristiques individuelles ou contractuelles (effet structurel) les bruits autour de cette tendance. Les modèles permettant de manipuler les données corrélées appartiennent à la famille des modèles linéaires mixtes. Le facteur à effets fixes sera alors *delta taux* et les facteurs à effets aléatoires seront toutes les caractéristiques des individus et du contrat observables ou non.

G.1 Forme générique et estimation

La forme générique des modèles mixtes linéaires est la suivante :

$$E(\mathbf{y}|\mathbf{u}) = g(\mathbf{X}\beta + \mathbf{Z}\mathbf{u}), \mathbf{u} \sim (\mathbf{0}, \mathbf{G})$$

où \mathbf{y} est le vecteur réponse, \mathbf{X} et \mathbf{Z} les matrices d'incidences, β est le vecteur des effets fixes et \mathbf{u} le vecteur des effets aléatoires. g est une fonction scalaire de lien. Le cas particulier du modèle linéaire est :

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{Z}\mathbf{u} + \varepsilon, \begin{bmatrix} \mathbf{u} \\ \varepsilon \end{bmatrix} \sim \left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{G} & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{bmatrix} \right) \quad (\text{G.1})$$

Un cas particulier de (G.1) est le modèle mixte gaussien où

$$\begin{bmatrix} \mathbf{u} \\ \varepsilon \end{bmatrix} \sim N \left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{G} & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{bmatrix} \right)$$

La log-vraisemblance de $(\beta, \mathbf{G}, \mathbf{R})$ est la suivante :

$$\ell((\beta, \mathbf{G}, \mathbf{R})) = -\frac{1}{2} \{n \log(2\pi) + \log |\mathbf{V}| + (\mathbf{y} - \mathbf{X}\beta)^\top \mathbf{V}^{-1} (\mathbf{y} - \mathbf{X}\beta)\} \quad (\text{G.2})$$

où

$$\mathbf{V} \equiv \text{Cov}(\mathbf{y}) = \mathbf{Z}\mathbf{G}\mathbf{Z}^\top + \mathbf{R} \quad (\text{G.3})$$

Alors l'estimation par la méthode du Maximum de Vraisemblance (en anglais *Maximum Likelihood* (ML)) de $(\beta, \mathbf{G}, \mathbf{R})$ est la valeur qui maximise la partie droite de l'expression. L'estimation de β est alors :

$$\hat{\beta} = (\mathbf{X}^\top \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{V}^{-1} \mathbf{y} \quad (\text{G.4})$$

Une alternative à l'estimation par Maximum de Vraisemblance de \mathbf{G} et \mathbf{R} est le Maximum de Vraisemblance Restreint (en anglais *Restricted Maximum Likelihood* (REML)). L'estimation de \mathbf{u} :

$$\hat{\mathbf{u}} = \mathbf{G}\mathbf{Z}^\top \mathbf{V}^{-1} (\mathbf{y} - \mathbf{X}\beta) \quad (\text{G.5})$$

Les modèles mixtes peuvent ne pas être utilisés uniquement dans le cas de groupement d'observations. Le lissage d'un diagramme de dispersion correspond à l'ajustement d'un modèle mixte. La manière dont ces derniers réalise le lissage peut être démontrée de la

manière suivante. Soit :

$$y_i = \beta_0 + \beta_1 x_i + \sum_{k=1}^K u_k (x_i - \kappa_k)_+ + \varepsilon_i$$

avec la fonction :

$$(x_i - \kappa_k)_+ = \begin{cases} 0, & x_i \leq \kappa_k \\ x_i - \kappa_k, & x_i > \kappa_k \end{cases} \quad (\text{G.6})$$

représentant une ligne de cassure avec un point de jonction, ou nœud, à κ_k . Si $u_k \stackrel{ind.}{\sim} N(0, \sigma_u^2)$ et les matrices d'incidences suivantes :

$$\mathbf{X} = [1x_i]_{1 \leq i \leq n}, \mathbf{Z} = [(x_i - \kappa_k)_+]_{1 \leq i \leq n \text{ et } 1 \leq k \leq K} \quad (\text{G.7})$$

et $\beta = [\beta_0, \beta_1]^\top$, $\mathbf{u} = [u_1, u_K]^\top$ alors cela revient au modèle linéaire mixte :

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{Z}\mathbf{u} + \varepsilon, \begin{bmatrix} \mathbf{u} \\ \varepsilon \end{bmatrix} \sim N \left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \sigma_u^2 \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \sigma_\varepsilon^2 \mathbf{I} \end{bmatrix} \right) \quad (\text{G.8})$$

En allant plus loin, le lien peut se faire avec la régression pénalisée. Pour σ_u^2 et σ_ε^2 donné :

$$\begin{bmatrix} \hat{\beta} \\ \hat{\mathbf{u}} \end{bmatrix} = \arg \min_{\beta, \mathbf{u}} (||\mathbf{y} - \mathbf{X}\beta - \mathbf{Z}\mathbf{u}||^2 + \alpha ||\mathbf{u}||^2) \quad (\text{G.9})$$

où $\alpha \equiv \sigma_\varepsilon^2 / \sigma_u^2$ et pour un vecteur \mathbf{v} , $||\mathbf{v}|| \equiv \sqrt{\mathbf{v}^\top \mathbf{v}}$. α est un paramètre de régularisation, souvent nommé hyperparamètre de régularisation. Ceci est un exemple des moindres carrés pénalisés dès que la minimisation de $||\mathbf{y} - \mathbf{X}\beta - \mathbf{Z}\mathbf{u}||^2$ est conditionnée par la pénalité $\alpha ||\mathbf{u}||^2$ imposée sur les coefficients de \mathbf{u} . La solution est la suivante :

$$\begin{bmatrix} \hat{\beta} \\ \hat{\mathbf{u}} \end{bmatrix} = (\mathbf{C}^\top \mathbf{C} + \alpha \mathbf{D})^{-1} \mathbf{C}^\top \mathbf{y} \quad (\text{G.10})$$

où $\mathbf{C} = [\mathbf{XZ}]$ et $\mathbf{D} = \text{diag}(0, 0, 1, \dots, 1)$.

La pénalité dans (G.9) est quadratique en \mathbf{X} :

$$\alpha \|\mathbf{u}\|^2 = \alpha \sum_{k=1}^K u_k^2 \quad (\text{G.11})$$

et provient de la condition de normalité des u_k . Différentes pénalités peuvent être obtenues en modifiant la distribution des u_k . Le lisseur radial utilisé par le logiciel **SAS** est un *approximate low-rank thin-plate spline* comme décrit dans Ruppert et al. (2003).

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{Z}_K \mathbf{u} + \varepsilon, \text{Cov}(\mathbf{u}) = \sigma_u^2 (\Omega_K^{-1/2}) (\Omega_K^{-1/2})^\top \quad (\text{G.12})$$

où $\mathbf{X} = [1 \ x_i \ \dots \ x_i^{m-1}]_{1 \leq i \leq n}$,

$$\mathbf{Z}_K = [|x_i - \kappa_k|^{2m-1}]_{\substack{1 \leq i \leq n \\ 1 \leq k \leq K}} \text{ et } \Omega_K = [|\kappa_k - \kappa_{k'}|^{2m-1}]_{\substack{1 \leq i \leq n \\ 1 \leq k, k' \leq K}} \quad (\text{G.13})$$

En utilisant la transformation $\mathbf{Z} = \mathbf{Z}_K \Omega_K^{-1/2}$ le modèle peut être écrit comme

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{Z}\mathbf{u} + \varepsilon, \begin{bmatrix} \mathbf{u} \\ \varepsilon \end{bmatrix} \sim N \left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \sigma_u^2 \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \sigma_\varepsilon^2 \mathbf{I} \end{bmatrix} \right) \quad (\text{G.14})$$

G.2 Détermination des nœuds

Le positionnement des nœuds est important pour la méthode de lissage car ces derniers influencent la courbure. Certaines méthodes ont été proposées afin de déterminer statistiquement quel est le meilleur choix des observations à retenir. La sélection des nœuds est faite à partir de la règle suivante :

$$\kappa_k = \left(\frac{k+1}{K+2} \right) \text{ième quantile des } x_i, 1 \leq k \leq K \quad (\text{G.15})$$

où Ruppert (2002) démontre que :

$$K = \max(5, \min(\frac{1}{4} \times \text{nombre de } x_i \text{ uniques}, 35)) \quad (\text{G.16})$$

Recherche d'un minimum

H.1 Écriture du problème de minimisation

La convergence vers un plateau d'indifférence à la baisse des taux à partir d'un point est modélisée à partir d'un modèle quadratique avec palier. Ce point détermine le seuil de stabilisation de la baisse des rachats conjoncturels. Il est possible d'écrire un tel modèle sous la forme suivante :

$$E[Y|x] = \begin{cases} \alpha + \beta x + \gamma x^2 & \text{si } x < x_0 \\ c & \text{si } x \geq x_0 \end{cases} \quad (\text{H.1})$$

Dans l'équation du modèle, les paramètres α , β et γ sont les coefficients du segment quadratique et c définit le plateau de la fonction moyenne. L'idée d'opter pour ce type de modèle plutôt que pour une estimation de la pente par un modèle linéaire repose sur l'objectif de la découverte d'un seuil à partir duquel les valeurs de *delta taux* n'ont plus d'effet sur la décision des rachats. Du plus il conforte l'hypothèse d'un taux de rachat conjoncturel unique lorsque les valeurs de *delta taux* sont au-delà d'un certain seuil.

Ce système d'équations se résout à l'aide de techniques d'optimisation continue, différentiable et non linéaire. Les algorithmes de calcul de minimum et plus particulièrement les algorithmes de descente sans contrainte sont utilisés dans ce genre de cas. Les pro-

blèmes de type moindres carrés peuvent être alors résolus grâce à l'optimisation non linéaire à travers la famille d'algorithmes de type Gauss-Newton.

La condition d'existence de la dérivée est la suivante :

$$x_0 = -\frac{\beta}{2\gamma} \quad (\text{H.2})$$

L'objectif est alors de pouvoir ajuster un modèle de régression non linéaire et d'estimer ses paramètres par moindres carrés non linéaires. Leur estimation repose sur un processus itératif qui débute à partir de valeurs de départ : α_0 , β_0 et γ_0 . Les valeurs initiales ont été obtenues en utilisant une régression linéaire sur les données observées à partir des variables *delta taux* et de son carré.

L'estimation par moindres carrés non linéaires nécessite de trouver les valeurs dans l'espace des paramètres qui minimise la somme des carrés des résidus. L'analyse requiert certaines hypothèses :

- la structure de la variable des rachats peut être décomposée additivement entre une fonction moyenne et une composante d'erreurs,
- les erreurs du modèle sont non corrélées, ont une moyenne nulle et sont supposées être homoscédastiques (de variance égale),
- la fonction moyenne consiste en des variables connues et des paramètres inconnus.

Il n'existe pas de solution analytique à l'estimation des paramètres. Le processus est alors itératif. Il peut être réalisé par la recherche d'un minimum local sur la surface de la somme des carrés des résidus.

Un problème d'optimisation s'écrit, à partir de X un sous-ensemble non vide de \mathbb{R}^n , sous la forme :

$$\min f(x) \text{ s.c } x \in X \quad (\text{H.3})$$

- la fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est appelée fonction coût, objectif ou critère,
- l'ensemble X est appelé ensemble ou domaine des contraintes,
- tout point $x \in \mathbb{R}^n$ vérifiant : $x \in X$, est appelé point admissible du problème.

H.2 Solution du problème de minimisation

Chercher une solution au problème avec contraintes revient à chercher un point de minimum local de f dans l'ensemble des points admissibles, au sens de la définition suivante :

- $x_0 \in \mathbb{R}^n$ est un point de minimum local de f sur $X \subset \mathbb{R}^n$ si et seulement si :

$$x_0 \in X \text{ et } \exists \mathcal{V}_{x_0} \text{ un voisinage de } x_0 \text{ tq } \forall x \in \mathcal{V}_{x_0} \cap X, f(x) \geq f(x_0) \quad (\text{H.4})$$

- $x_0 \in \mathbb{R}^n$ est un point de minimum global de f sur X si et seulement si :

$$x_0 \in X \text{ et } \forall x \in X, f(x) \geq f(x_0) \quad (\text{H.5})$$

Les notions de maximum local et global sont définies de façon tout à fait similaire. Les algorithmes d'optimisation reposent sur les éléments suivants :

- calcul des dérivées : l'algorithme aura besoin d'informations sur la fonction objectif f : la valeur numérique de f en un point donné x , et souvent également le gradient $\nabla f(x)$,
- convergence et vitesse de convergence : un algorithme de descentes, est dit convergent si la suite de ses itérés $(x_k)_{k \in N}$ converge vers un point limite x^* , solution du problème $\min_{x \in \mathbb{R}^n} f(x)$,
- critères d'arrêt : un test d'arrêt est choisi pour garantir que l'algorithme s'arrête toujours après un nombre fini d'itérations et que le dernier point calculé soit suffisamment proche de x^* .

A partir du modèle pour la i^{e} observation :

$$Y_i = f(\beta; \mathbf{x}'_i) + \varepsilon_i \quad (\text{H.6})$$

puis écrire simplement la somme des carrés des résidus :

$$\text{SSE}(\beta) = \sum_{i=1}^n (y_i - f(\beta; \mathbf{x}'_i))^2 \quad (\text{H.7})$$

Il est plus instructif de décrire les algorithmes par une approche consistant à minimiser la somme des carrés des résidus et dans leur manière de mettre à jour leur formule. Si $\hat{\beta}^{(u)}$ est la valeur des estimations des paramètres à la u^e itération et $\hat{\beta}^0$ est la valeur de départ, l'algorithme cherche alors les valeurs k et Δ telles que :

$$\hat{\beta}^{(u+1)} = \hat{\beta}^{(u)} + k\Delta$$

$$\text{SSE}(\hat{\beta}^{(u+1)}) < \text{SSE}(\hat{\beta}^{(u)})$$

La méthode de Gauss-Newton :

$$\Delta = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{r}$$

Table des matières

Résumé	i
Table des figures	vii
Liste des tableaux	ix
Introduction	1
I La France Mutualiste et le risque de rachat	3
I.1 La France Mutualiste	3
I.1.1 Présentation de La France Mutualiste	4
I.1.2 Histoire	4
I.1.3 Produits commercialisés	5
I.1.4 Caractéristiques du portefeuille LFM	7
I.2 Les produits d'épargne en France	8
I.2.1 La sécurité du support garanti	8
I.2.2 Un placement pour préparer la retraite?	9
I.3 Le risque de rachat	10
I.3.1 Les motivations incertaines	10
I.3.2 Les conséquences	10
I.3.3 L'encadrement des risques	11
I.4 Mesure du risque rachat	12
I.4.1 Modélisation actuelle des rachats à La France Mutualiste	12
I.4.2 Choix de la fonction de rachats conjoncturels	14
I.5 Problématique	16
II Revue méthodologique et analyse exploratoire	17
II.1 Revue de la littérature	17
II.1.1 Les hypothèses	17

II.1.2	Les modèles proposés	19
II.2	Les données	20
II.2.1	Structure	20
II.2.2	Délai de règlements	20
II.3	L'événement et ses origines	21
II.3.1	Analyse exploratoire	23
II.3.2	Conclusion	27
II.4	Le taux attendu et le taux servi	27
II.4.1	Choix du taux attendu	28
II.4.2	Conséquences	28
II.5	Premier résultat	29
III	Hypothèse du taux d'intérêt	31
III.1	Mesure de l'intensité et du sens de la relation	31
III.2	Du modèle Logistique au modèle de choix discrets	33
III.2.1	Raisons du modèle logistique	33
III.2.2	Puis du modèle de choix discrets	34
III.2.3	Processus d'un choix	35
III.2.4	Comportement rationnel	35
III.3	Le modèle de choix discrets	36
III.3.1	Théorie	36
III.3.2	Évaluation	38
III.4	Application	38
IV	Loi conjoncturelle : partie droite	41
IV.1	Estimation de γ	41
IV.1.1	Résultats du modèle de choix discrets	42
IV.1.2	Validation de l'effet additif	43
IV.1.3	Seuil optimal	44
IV.2	Estimation de δ	45
IV.2.1	Lissage	45
IV.2.2	Modèle mixte	46
IV.3	Estimation de RC_{min}	48
IV.4	Conclusion	50
V	Loi conjoncturelle : partie gauche	51
V.1	Modélisation des remboursements de prêt	52
V.1.1	Hypothèses	52
V.1.2	Adaptation au rachat d'un contrat d'épargne	52

V.2	Modélisation des rachats	53
V.2.1	Le contrat alternatif	53
V.2.2	Frais d'entrée	54
V.2.3	Fiscalité	54
V.2.4	Placements investis	55
V.2.5	Chronologie du rachat	55
V.2.6	Valeur du rachat	56
V.3	Résultats intermédiaires	56
V.4	Retour sur le modèle de choix discrets	57
V.4.1	Coefficient de corrélation de Matthews	57
V.4.2	Probabilité estimée et seuil optimum	59
VI	Loi structurelle et impacts Solvabilité II	61
VI.1	Loi structurelle	61
VI.1.1	Hypothèse	61
VI.1.2	Observation	62
VI.1.3	Justification	63
VI.1.4	Évaluation	64
VI.2	Calculs prudentiels sous Solvabilité II	65
VI.2.1	La valeur économique	65
VI.2.2	Le capital de solvabilité requis	65
VI.2.3	Le taux de couverture	66
VI.3	Analyse des variations	66
VI.3.1	Méthodologie	66
VI.3.2	Résultats	67
VI.3.3	Quelques explications	68
VI.3.4	Illustration	69
VI.3.5	Résumé	70
	Conclusion	71
	Bibliographie	73
	Annexes	77
A	Généralités	77
A.1	Loi du 4 août 1923	77
A.2	RMC : particularités	77
A.3	Le risque de rachat sous toutes ces formes	78
A.4	Modélisation d'un contrat racheté	79

A.5	Sur-exposition de certains mois	79
A.6	La RMC (Retraite Mutualiste du Combattant) favorise le développement de La France Mutualiste	81
A.7	Mesures pour encadrer les taux garantis	82
B	Modèle logistique	85
B.1	Notations	85
B.2	Principe et estimation	86
B.3	Estimation des paramètres par la maximisation de la vraisemblance	89
C	Modèle à choix discret	91
C.1	Le modèle binomial	92
C.2	Modèle de durée de survie	95
C.2.1	Risque en temps discret et survie	95
C.2.2	Survie en temps discret et régression logistique	96
C.3	Le modèle logistique	97
C.4	Modèle Linéaire Généralisé : généralités	101
C.4.1	Niveau contrat - variable binaire	101
C.4.2	Données agrégées - variable continue	103
C.5	Du modèle logit au modèle de choix discrets	103
D	Évaluation du modèle à choix discrets	109
D.1	Évaluation de la régression	109
D.2	La courbe ROC	112
D.2.1	La courbe ROC par la méthode binomiale	114
D.2.2	La courbe ROC par la valeur c	117
D.3	Résultats du modèle complet	119
E	Indice de Youden	123
F	Méthodes de lissage	127
G	Modèle mixte	131
G.1	Forme générique et estimation	131
G.2	Détermination des nœuds	134
H	Recherche d'un minimum	135
H.1	Écriture du problème de minimisation	135
H.2	Solution du problème de minimisation	137