EURIA — EURo Institut d'Actuariat

Mémoire présenté devant le jury de l'EURIA en vue de l'obtention du Diplôme d'Actuaire EURIA et de l'admission à l'Institut des Actuaires

le 20 Septembre 2018

Par : Miguel LEBLOND
Titre : Comment les assureurs peuvent-ils modéliser le comportement de rachat des assurés suite à l'entrée en vigueur de la « Flat Tax » ?

Confidentialité : Non

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

Membre présent du jury de l'Institut des Actuaires :
Nicolas BOURE
Signature :

Entreprise :
FIXAGE
Signature :

Membres présents du jury de l'EURIA :
Romain LAILY
Françoise PENE
Dominique ABGRALL

Directeur de mémoire en entreprise :
Magali ROUJAS
Signature :

Invité :

Signature :

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

Signature du responsable entreprise :

Signature du candidat :

Secrétariat :

Bibliothèque :

EURIA
EURo Institut d'Actuariat
6, avenue le Gorgeu
CS 93837
29238 Brest Cedex 3
T +33 (0)2 98 01 66 55
F +33 (0)2 98 01 66 57
euria@univ-brest.fr

Remerciements

Je souhaite remercier en premier lieu Monsieur Michel PIERMAY, président de FIXAGE, pour m'avoir accueilli au sein de sa société et pour ses conseils lors de la construction et la rédaction de ce mémoire.

Je remercie également tout particulièrement Magali ROUJAS, ma tutrice de stage, pour sa disponibilité, son encadrement et ses suggestions tout au long de mon stage.

Je remercie toute l'équipe de FIXAGE pour leur aide et leur accueil chaleureuse tout au long de ce stage, en particulier mes collègues stagiaires.

J'adresse mes sincères remerciements à l'ensemble du corps enseignant de l'EURIA pour la qualité des cours dispensés durant ces trois années de formation. Je remercie également Sophie BOURDET, ma tutrice universitaire pour son encadrement et ses conseils réguliers.

Enfin, je remercie mes très chers parents pour leur aide précieuse et tous ceux qui ont contribué à ce mémoire.

Résumé

Le gouvernement a mis en place un nouveau prélèvement forfaitaire unique, la « Flat Tax », afin de simplifier et d'harmoniser la fiscalité de l'épargne. Depuis le 1er janvier 2018, il s'applique à tous les revenus mobiliers, en particulier les produits des contrats d'assurance vie. Avec l'ancienne fiscalité, le taux d'imposition appliqué aux produits dépendait de l'ancienneté du contrat. La « Flat Tax » supprime le compartiment fiscal à 4 ans et réduit la différence entre les taux fiscaux des compartiments avant et après 8 ans d'ancienneté (voire la supprime à partir de 150 000 euros versés sur le contrat).

Les assureurs ont la nécessité de modéliser le comportement de rachat des assurés dans l'objectif d'estimer leurs indicateurs économiques. En général, ils utilisaient l'ancienneté fiscale comme facteur explicatif des rachats structurels par l'assuré. La modification des compartiments fiscaux remet en cause cette méthode de modélisation.

Nous avons proposé dans ce mémoire deux méthodes de modélisation des rachats structurels des assurés. Chacune de ces méthodes a été utilisée dans l'estimation du ratio de couverture et de la *Present Value of Future Profits* et comparée aux résultats issus de l'utilisation de la loi de rachats structurels en fonction de l'ancienneté. Nous avons eu recours à un outil de projection pour projeter le portefeuille de contrats (pour l'estimation des indicateurs).

Nous avons tout d'abord présenté une adaptation de la loi de rachats en fonction de l'ancienneté en utilisant une fonction paramétrique caractérisant le gain ou la perte pour les assurés sur les produits de leur contrat générés par la nouvelle fiscalité.

Nous avons eu recours ensuite aux techniques de *machine learning*. Ces méthodes permettent d'apprendre le taux de rachat par individu à partir de plusieurs variables explicatives. Elles sont une bonne alternative à l'utilisation de l'ancienneté comme seul facteur explicatif du rachat structurel. Nous avons retenu trois modèles ensemblistes et un modèle connexionniste pour cette étude : le *Random Forest*, le *Gradient Boosting Machine*, *XGBoost* et les réseaux de neurones à plusieurs couches cachées. La performance des modèles a été étudiée en vue de la prédiction des taux de rachat. Il en ressort que les résultats des modèles ont une erreur assez similaire. Le *Gradient Boosting Machine* est le modèle qui a la prédiction des taux de rachat la plus proche des taux de rachat réels observés sur la base de test.

Nous avons ensuite estimé les indicateurs de solvabilité et de rentabilité à l'aide d'un modèle de projection. Avec les lois construites, le ratio de couverture et la *Present Value of Future Profits* sont dégradés comparés aux indicateurs issus de la loi de rachats en fonction de l'ancienneté. Les taux de rachat prédits sont en moyenne plus élevés que les taux fournis par la loi de rachats en fonction de l'ancienneté.

Mots clés : « Flat Tax », rachats structurels, *machine learning*, modèle de projection, ratio de couverture, *Present Value of Future Profits*.

Abstract

The government has introduced a new flat-rate flat tax levy to simplify and harmonize the taxation of savings. Since January 1, 2018, it applies to all movable income, in particular the products of life insurance contracts. With the old tax system, the tax rate applied to the products depended on the age of the contract. The flat tax eliminates the tax compartment at 4 years and reduces the difference between the tax rates of the compartments before and after 8 years of seniority (or even eliminates it from 150,000 euros paid on the contract).

Insurers have the need to model the repurchase behavior of policyholders in order to estimate their economic indicators. In general, they used the tax seniority as an explanatory factor of structural redemptions by the insured. The modification of the tax compartments calls into question this method of modeling.

In this paper, we proposed two methods of modeling the structural surrender of policyholders. Each of these methods has been used in the estimation of the coverage ratio and *Present Value of Future Profits* and compared to the results from the use of the law of structural redemptions according to seniority. We used a projection tool to project the portfolio of contracts (for estimating indicators).

We first presented an adaptation of the buy-back law according to seniority using a parametric function characterizing the gain or loss for policyholders on the products of their contract generated by the new tax system.

Then, we used the techniques of *machine learning*. From these methods, it is possible to learn the redemption rate per individual from several explanatory variables. They are a good alternative to the use of seniority as the sole explanatory factor for structural buyback. We have selected three set models and a connectionist model for this study : the *Random Forest*, the *Gradient Boosting Machine*, *XGBoost* and the hidden multilayer neural networks. The performance of the models was studied for the purpose of predicting surrender rates. It turns out that the model results have a fairly similar error. The *Gradient Boosting Machine* is the model that has the prediction of the surrender rates closest to the actual redemption rates observed on the test basis.

Moreover, we estimated the solvency and profitability indicators using a projection model. With the laws built, the hedge ratio and the *Present Value of Future Profits* are degraded compared to the indicators resulting from the law of redemptions according to the seniority. The Predicted redemption rates are on average higher than the rates provided by the buy-backs by seniority.

Keywords : Flat Tax, structural redemptions, machine learning, projection model, coverage ratio, Present Value of Future Profits.

Note de synthèse

Contexte de l'étude

Depuis le 1^{er} janvier 2018, le prélèvement forfaitaire unique à 30% (ou « Flat Tax ») est en vigueur pour tous les revenus du capital. Les produits des contrats d'assurance vie, issus des nouveaux versements à compter du 27 septembre 2017, sont concernés par cette réforme. Les compartiments fiscaux sont donc modifiés. Les produits issus des nouveaux versements des contrats de moins de 8 ans d'ancienneté fiscale sont imposés à 12,8% (auxquels sont rajoutés les 17,2% de prélèvements sociaux). Au delà de 8 ans, le taux d'imposition diffère selon que le montant des primes versées et non rachetées par l'assuré est supérieur ou inférieur à 150 000 euros.

Auparavant la fiscalité des contrats d'assurance vie dépendait de l'ancienneté du contrat souscrit. Les assureurs modélisent généralement leur loi de rachats structurels en fonction de l'ancienneté du contrat en vue du calcul de leurs indicateurs économiques (ratio de couverture, *Present Value of Future Profit*, etc). La modification des compartiments fiscaux remet en cause l'utilisation du facteur ancienneté comme seule variable explicative du comportement de rachat structurel des assurés. Comment les assureurs peuvent-ils désormais modéliser le comportement de rachat structurel des assurés ?

Description de la démarche

Dans ce mémoire, nous proposons aux assureurs deux méthodes de modélisation de la loi de rachats structurels. La première consiste à adapter la loi de rachats structurels en fonction de l'ancienneté et la seconde est de modéliser les rachats structurels en utilisant une approche *machine learning* qui permet de prendre en compte d'autres facteurs explicatifs que l'ancienneté. Nous en estimerons ensuite les effets sur les indicateurs de solvabilité (ratio de couverture) et de rentabilité (PVFP).

Pour l'application de ces deux méthodes, nous avons à notre disposition un historique de rachats sur les fonds en euros sur la période 2010 à 2017. Tout d'abord, nous avons calculé les taux de rachat par individu. Nous avons ensuite construit la loi de rachats en fonction de l'ancienneté basée sur l'historique de la période 2010 à 2016. La loi est la suivante :

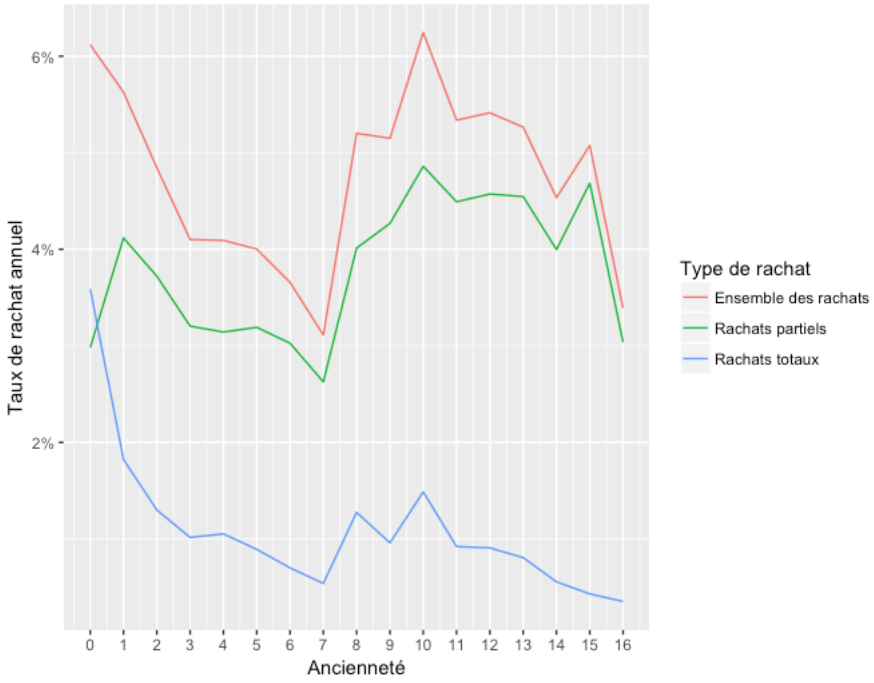

FIGURE 1 – Évolution du taux de rachat par ancienneté du contrat.

L'adaptation de la loi de rachats structurels en fonction de l'ancienneté

La première méthode consiste à prendre compte l'évolution des règles fiscales dans la loi de rachats structurels en fonction de l'ancienneté. Nous avons eu recours à une fonction paramétrique pour adapter cette loi. Cette fonction représente le gain ou la perte pour les assurés (lors d'un rachat) sur les produits que génère la mise en place de la « Flat Tax ». Par exemple, dans le cas d'un contrat dont le montant de primes versées et non rachetées est de 200 000 euros et qui a généré 10 000 euros de produits (nets de prélèvements sociaux), les résultats sont les suivants :

Ancienneté contrat	Produit net (Fiscalité avant)	Produit net (Fiscalité après)	Gain / Perte
0 - 4 ans	6 500 €	8 720 €	34,15%
4 - 8 ans	8 500 €	8 720 €	2,59%
>= 8 ans	9 595 €	9 462,5 €	-1,38%

A partir de l'ancienneté et des caractéristiques des contrats, la fonction paramétrique permet de déterminer le gain ou la perte qu'induit la nouvelle fiscalité. Dans notre situation de rachat, ces valeurs sont détaillées dans la dernière colonne du tableau ci-dessus. Nous proposons d'ajuster la loi de rachats structurels en fonction de l'ancienneté proportionnellement au gain/perte observé pour chaque cas de contrats. Les différents cas sont les suivants :

— Les contrats souscrits après le 27 septembre 2017 avec un montant de primes versées et non rachetées inférieur à 150 000 euros.

— Les contrat souscrits après le 27 septembre 2017 avec un montant de primes versées et non rachetées supérieur à 150 000 euros.

— Les contrats souscrits avant le 27 septembre 2017 avec un montant de primes versées et non rachetées inférieur à 150 000 euros sans effectuer de versements après cette même date.

— Les contrats souscrits avant le 27 septembre 2017 avec un montant de primes versées et non rachetées inférieur à 150 000 euros avec versements après cette même date.

— Les contrats souscrits avant le 27 septembre 2017 avec un montant de primes versées et non rachetées supérieur à 150 000 euros sans effectuer de versements après cette même date.

— Les contrats souscrits avant le 27 septembre 2017 avec un montant de primes versées et non rachetées supérieur à 150 000 avec versements après cette même date.

La loi de rachats structurels adaptée est définie comme suit :

$$Loi_rachat_adaptée(x_i, y_i) = (1 + Fonction_paramétrique(x_i, y_i)) \times Loi_rachat(x_i, y_i)$$

Avec x_i et y_i respectivement l'ancienneté du contrat et les caractéristiques du contrat d'un individu i.

Les deux lois de rachats du portefeuille sont représentées ci-dessous :

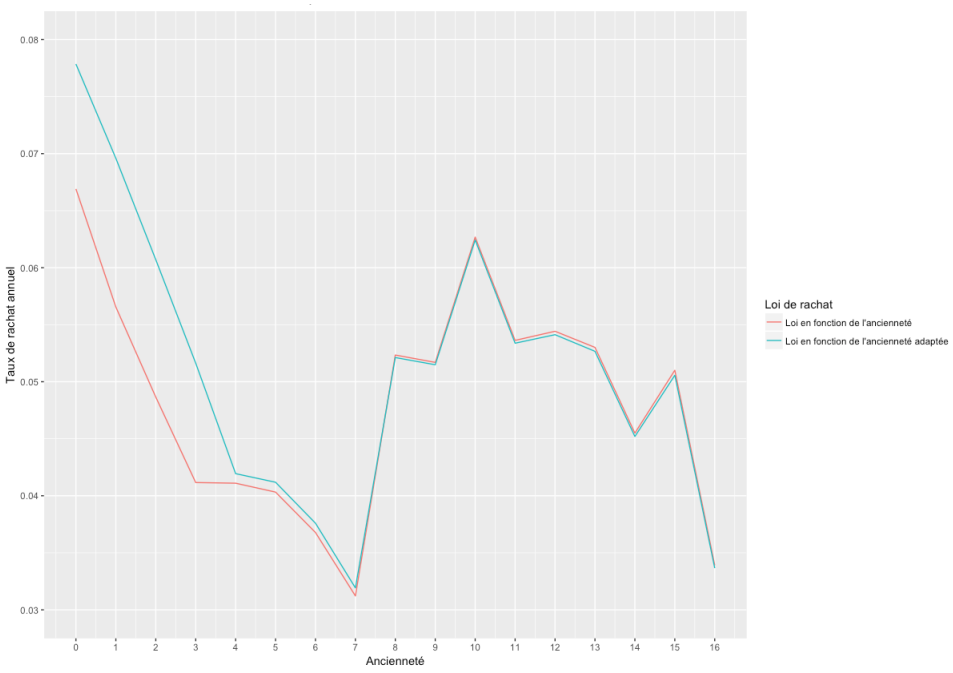

FIGURE 2 – Comparaison des deux lois de rachats.

Cette première méthode repose sur l'hypothèse forte de l'existence d'une relation linéaire entre le rachat et l'évolution du gain ou de la perte des produits généré par la « Flat Tax ».

En considérant que le rachat dépend exclusivement de l'ancienneté du contrat, on peut supposer que les assurés seront plus enclins à racheter entre 0 et 8 ans et moins pour une ancienneté au delà de 8 ans (par rapport à avant).

Les lois *machine learning*

Comme deuxième méthode, nous proposons d'utiliser les techniques de *machine learning* pour prédire le taux de rachat par individu en fonction de plusieurs variables explicatives. La base de rachats à notre disposition contient le sexe, l'âge, l'ancienneté du contrat, le montant de provision mathématique et l'absence ou non de versement et d'arbitrage effectué sur le contrat.

Nous avons retenu deux catégories de modèles :

— Les méthodes ensemblistes qui permettent de transformer un modèle simple tel que les arbres de décision en un modèle plus performant capable de modéliser les liens complexes existants entre la variable à expliquer et les variables explicatives. Nous utiliserons 3 modèles reposant sur ce principe général : *Random Forest*, *Gradient Boosting Machine* et *Xgboost* ;

— Les réseaux de neurones (*Deeplearning*) qui permettent de modéliser de fortes interactions non linéaires entre les variables explicatives et la variable à expliquer.

La démarche retenue pour effectuer la prédiction des taux de rachat est la suivante :

— Le traitement de la base de données ;

— La phase d'échantillonage ;

— La phase d'apprentissage et de calibrage des paramètres ;

— La phase de prédiction.

Nous avons effectué l'apprentissage sur la période 2010 à 2016. Le test est effectué sur l'année 2017. Le taux de rachat moyen réel observé est de 4,6 %. Les résultats sont les suivants :

Modèle	RMSE	Taux de rachat moyen
Random Forest	0,1843	4,53%
Gradient Boosting	0,18416	4,56%
XGBoost	0,18428	4,53%
Deeplearning	0,18431	4,49%

La loi de rachats en fonction de l'ancienneté propose un taux de rachat moyen de 4,84% sur l'année 2017. Les taux de rachat moyens prédits (en fonction des six variables explicatives) sont tous plus faibles que le taux de rachat moyen réel mais sont plus proches du taux réel observé que le taux moyen proposé par la loi de rachats en fonction de l'ancienneté. En ne prenant en compte que l'ancienneté comme variable explicative, le taux de rachat est surestimé.
La RMSE (*Root Mean Square Error*) est l'indicateur qui permet de mesurer la qualité des modèles (à minimiser). Parmi les modèles testés, le *Gradient Boosting Machine* a la meilleure performance en terme d'erreur quadratique moyenne.

Les méthodes ensemblistes sont très performantes pour déterminer le pouvoir explicatif des variables. Le *Gradient Boosting* fournit comme résultat :

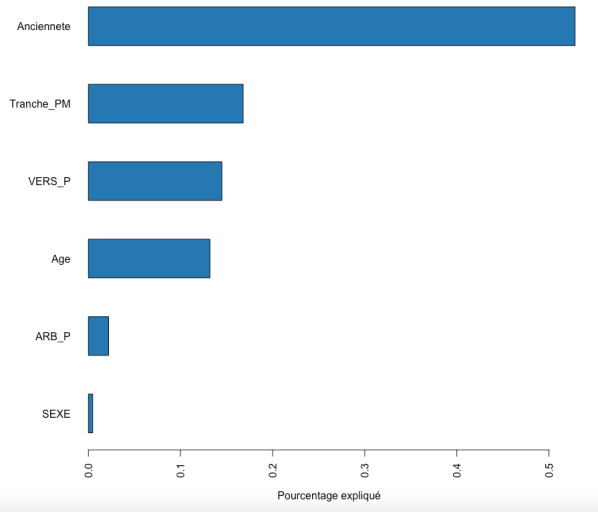

FIGURE 3 – Importance des variables.

La variable ancienneté explique environ 50% de la prédiction du taux de rachat par individu. Le montant de provision mathématique, la présence ou non de versement et l'âge ont de l'importance également dans la prédiction du taux. Cette analyse vient confirmer l'idée de prendre en compte plusieurs variables explicatives dans l'estimation du taux de rachat structurel. De plus, les modèles de *machine learning* ont fourni une meilleure estimation des taux de rachat que la loi en fonction de l'ancienneté sur l'année 2017.

L'impact de la modification des lois de rachats sur les indicateurs économiques

Nous avons ensuite confronté les valeurs des indicateurs économiques issus du modèle de projection en modifiant la loi de rachats structurels. Les lois construites sont comparées à une loi de référence qui est la loi de rachats en fonction de l'ancienneté de l'assureur (sans prise en compte des évolutions fiscales). Par construction des lois en fonction de l'ancienneté, une fois que l'ancienneté des *model points* atteint 16 ans, les taux de rachat sont égaux à 3,4%.

Les résultats issus du modèle de projection en utilisant la loi de référence et la loi de rachats adaptée sont les suivants :

Indicateur	Situation de référence	Loi de rachats adaptée
Best Estimate	18 059 millions €	18 060 millions €
PVFP	757 millions €	755 millions €
Ratio de solvabilité	121,2 %	121,0 %

En utilisant la loi de rachats adaptée, la solvabilité et la rentabilité de l'assureur diminuent légèrement. Les prestations de rachat associées sont plus importantes en début de projection. En effet, la présence des contrats ayant moins de 8 ans d'ancienneté fiscale induit cette distinction. Effectivement, l'actualisation impacte moins les prestations en début qu'en fin de période.

Nous avons ensuite comparé la valeur des indicateurs avec les lois construites par les techniques de *machine learning*. Les résultats obtenus sont donnés ci-dessous :

Indicateur (M €)	Référence	*RF*	*GBM*	*XGB*	*DL*
Best Estimate	18 059	18 185	18 197	18 186	18 191
PVFP	757	631	618	629	624
Ratio de solvabilité	121,2 %	114,3 %	113,7 %	114,2 %	114,0 %

Tous les indicateurs issus de la projection avec les lois *machine learning* sont dégradés par rapport aux indicateurs construits avec la loi de rachats en fonction de l'ancienneté. La solvabilité et la rentabilité diminuent de manière significative. Pour approfondir notre analyse, nous proposons de constater l'évolution du taux de rachat moyen prédit à chaque pas de temps lors de la projection :

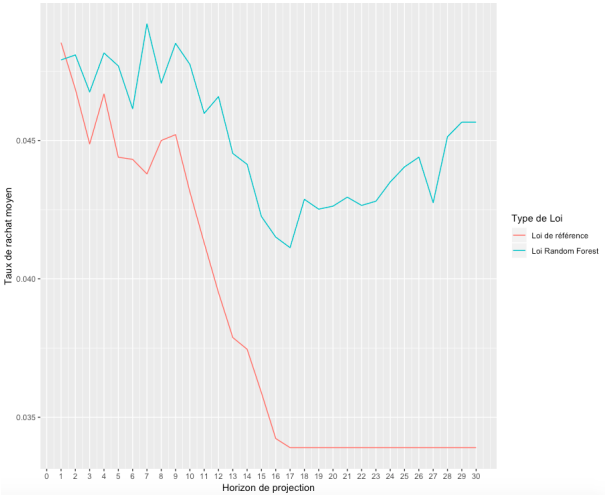

FIGURE 4 – Comparaison des taux de rachat moyens *Random Forest* avec la loi de référence.

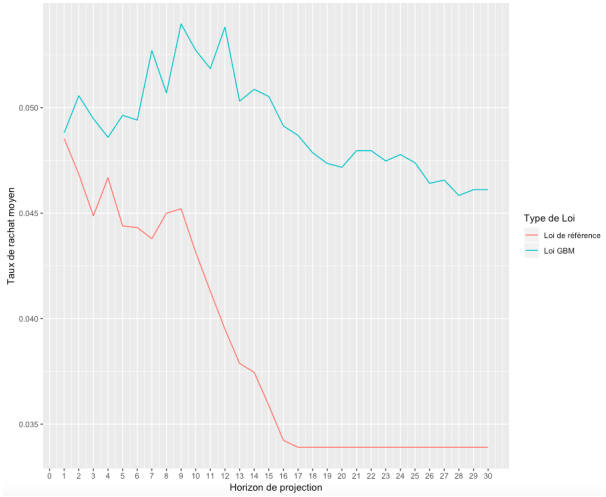

FIGURE 5 – Comparaison des taux de rachat moyens *GBM* avec la loi de référence.

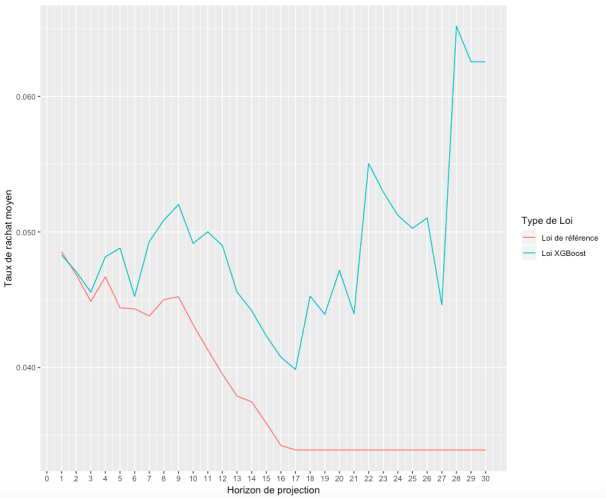

FIGURE 6 – Comparaison des taux de rachat moyens *XGBoost* avec la loi de référence.

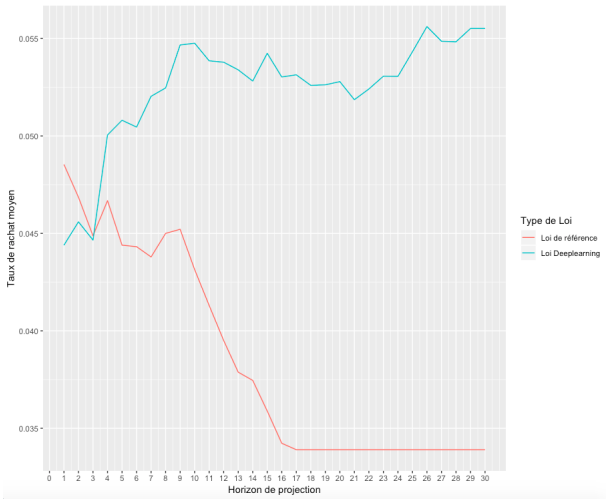

FIGURE 7 – Comparaison des taux de rachat moyens *Deeplearning* avec la loi de référence.

L'évolution des taux de rachat moyens au titre du rachat structurel permet d'expliquer la baisse de la rentabilité. En effet, nous pouvons constater, mis à part pour le modèle *Deeplearning* sur les trois premières années de projection, que les taux de rachat prédits sont supérieurs aux taux de rachat fournis par la loi en fonction de l'ancienneté. Notre portefeuille est sensible à la hausse des rachats donc la rentabilité diminue.

L'intérêt d'utiliser le *machine learning* dans la prédiction des taux de rachat est qu'à chaque pas de temps, les taux sont estimés en prenant en compte l'évolution d'autres variables que l'ancienneté comme l'âge ou le montant de provision mathématique. Nous avons constaté sur les graphiques représentant l'évolution du taux de rachat moyen l'apport du *machine learning*. A partir de l'instant où toutes les anciennetés des contrats du portefeuille ont atteint 16 ans, le taux devient constant jusqu'à la fin de la projection (pour la loi de référence). Les méthodes de *machine learning* permettent de donner un poids à chaque variable. Une fois que tous les contrats du portefeuille ont une ancienneté supérieure à 16 ans, au lieu de donner le même taux de rachat jusqu'à la fin de la projection, les modèles vont se servir également des autres variables pour prédire le taux de rachat et ainsi obtenir des taux différents à chaque pas de temps.

Conclusion

Nous avons proposé dans ce mémoire deux méthodes alternatives à la loi de rachats en fonction de l'ancienneté pour estimer le taux de rachat structurel. L'intérêt de la première méthode est de prendre en compte l'évolution de la fiscalité dans la loi de rachats structurels en fonction de l'ancienneté. La seconde permet de prendre en compte d'autres variables explicatives que l'ancienneté dans l'estimation des taux de rachat. Les lois *machine learning* ont fourni une meilleure estimation des taux de rachat par rapport à la loi de rachats structurels en fonction de l'ancienneté sur l'année 2017. L'approche *machine learning* semble être une bonne alternative pour la construction de la loi de rachats structurels. Enfin, nous avons constaté qu'un changement de loi de rachats structurels pouvait impacter de manière significative les indicateurs économiques d'un assureur. En effet, l'utilisation des lois *machine learning* a induit une baisse de la rentabilité et de la solvabilité de notre assureur.

Une prochaine étape pour compléter nos travaux serait de réaliser une étude sur l'estimation des taux de rachat globaux (structurels et conjoncturels) en prenant en compte, en plus des facteurs structurels, des variables décrivant la conjoncture économique du marché (taux servi, taux de croissance etc).

Mots clés : « Flat Tax », Assurance vie, fiscalité, ratio de couverture, *Present Value of Future Profits*, rachats structurels, *machine learning*, méthodes ensemblistes, réseaux de neurones, modèle de projection.

Summary

Study context

From January 1, 2018, the flat rate levy of 30% (or Flat Tax) is in effect for all capital income. The proceeds of life insurance contracts, resulting from new payments as of September 27, 2017, are affected by this reform. The tax compartments are therefore modified. Revenues from new payments of contracts with less than 8 years of tax maturity are taxed at 12.8% (to which is added 17.2% of social security contributions). Beyond 8 years, the tax rate differs depending on whether the amount of premiums paid and not repurchased by the insured is higher or lower than 150 000 euros.

Previously the taxation of life insurance contracts depended on the age of the contract. Insurers generally model their structural buy-backs based on the age of the contract in order to calculate their economic indicators (coverage ratio, Present Value of Future Profit, etc.). The modification of the tax compartments calls into question the use of the maturity factor as the sole explanatory variable for the structural buyback behavior of the insured. How can insurers now model the structural buyback behavior of policyholders ?

Description of the approach

In this paper, we propose to insurers two methods of modeling the law of structural redemptions. The first is to adapt the law of structural buy-backs according to maturity and the second is to model the structural buybacks using a *machine learning* approach that allows to take into account other explanatory factors that seniority. We will then estimate the effects on the solvency (coverage ratio) and profitability (PVFP) indicators.

For the application of these two methods, we have at our disposal a history of redemptions on funds in euros over the period 2010 to 2017. First, we calculated the redemption rates per individual. We then built the buyout law based on seniority based on the history of the period 2010 to 2016. The law is as follows :

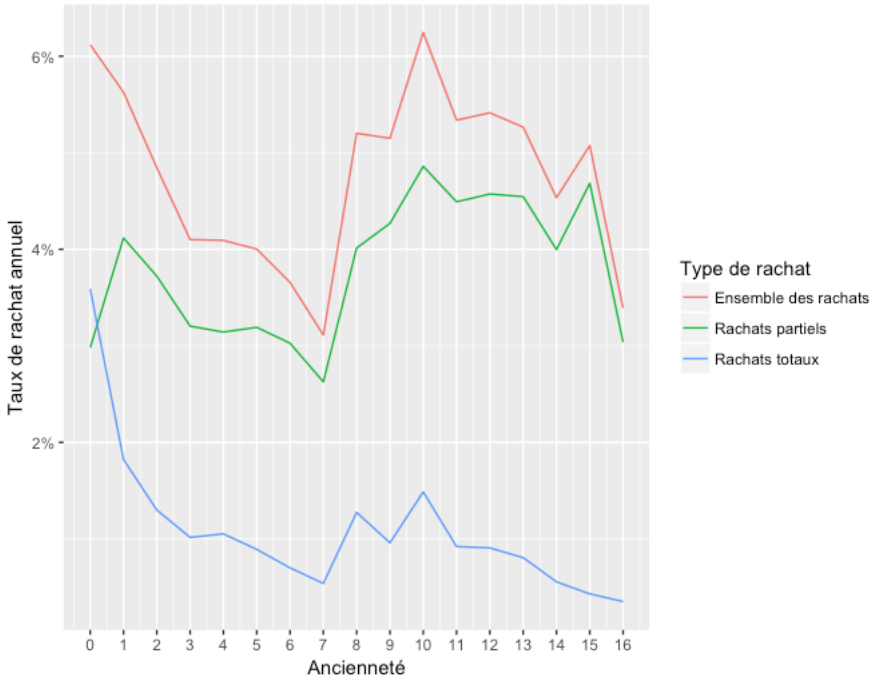

FIGURE 8 – Evolution of the repurchase rate by maturity of the contract.

Adaptation of the law of structural redemptions according to the seniority

The first method is to take into account the evolution of tax rules in the law of structural redemptions according to seniority. We used a parametric function to adapt this law. This function represents the gain or loss for policyholders (on a buy-back) on the products generated by the introduction of the Flat Tax fg. For example, in the case of a contract whose amount of premiums paid and not repurchased is 200 000 euros and which generated 10 000 euros of products (net of social security contributions), the results are as follows :

Contract maturity	Net proceeds (Taxation before)	Net proceeds (Taxation after)	Profit / Loss
0 - 4 years	6 500 €	8 720 €	34,15%
4 - 8 years	8 500 €	8 720 €	2,59%
>= 8 years	9 595 €	9 462,5 €	-1,38%

From the maturity and the characteristics of the contracts, the parametric function makes it possible to determine the profit or loss induced by the new taxation. In our redemption situation, these values are detailed in the last column of the table above. We propose to adjust the law of structural redemptions according to seniority in proportion to the profit / loss observed for each contract case. The different cases are as follows :

— Contracts subscribed after 27 September 2017 with premiums paid and not repurchased less than 150,000 euros.

— Contracts subscribed after September 27, 2017 with an amount of premiums paid and not repurchased higher than 150 000 euros.

— Contracts subscribed before September 27, 2017 with an amount of premiums paid

and not repurchased less than 150,000 euros without making any payments after that date.

— Contracts subscribed before 27 September 2017 with an amount of premiums paid and not repurchased less than 150,000 euros with installments after this date.

— Contracts subscribed before September 27, 2017 with an amount of premiums paid and not repurchased higher than 150 000 euros without making payments after that same date.

— Contracts subscribed before September 27, 2017 with an amount of premiums paid and not repurchased greater than 150,000 with installments after this date.

The law of adapted structural redemptions is defined as follows :

$$Loi_rachat_adaptée(x_i, y_i) = (1 + Fonction_paramétrique(x_i, y_i)) \times Loi_rachat(x_i, y_i)$$

With x_i and y_i respectively the maturity of the contract and the characteristics of the contract of an individual i.

The two portfolio redemption laws are represented below :

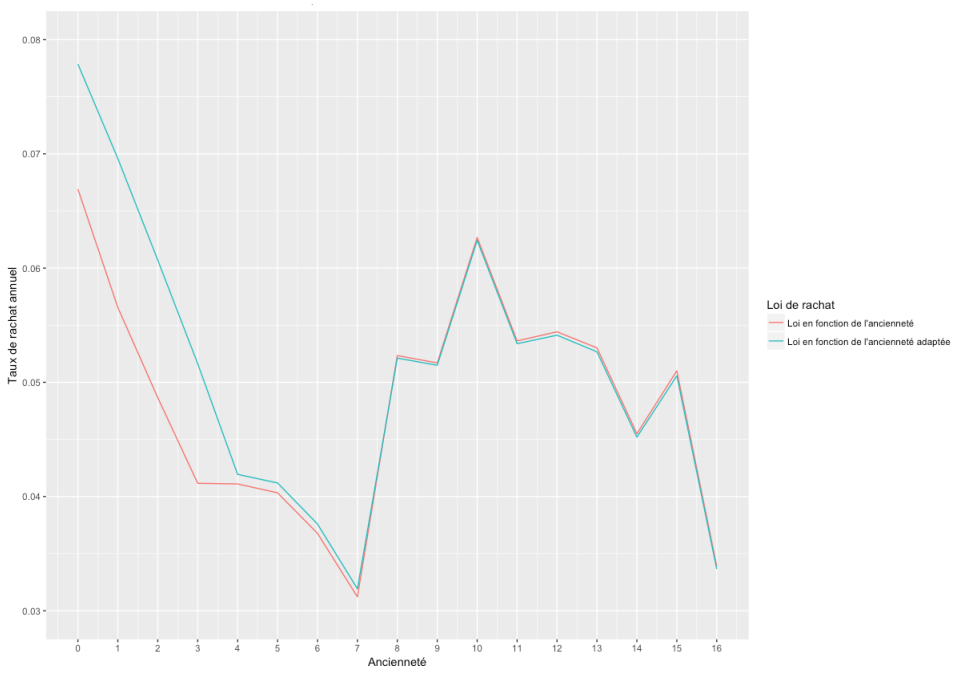

FIGURE 9 – Comparison of the two buy-back laws.

This first method is based on the strong assumption of the existence of a linear relationship between the repurchase and the evolution of the gain or loss of the products generated by the « Flat Tax ».

Considering that the buyout depends exclusively on the seniority of the contract, it can be assumed that the insured will be more inclined to buy between 0 and 8 years and less for maturity beyond 8 years (compared to before).

The laws *machine learning*

As a second method, we propose to use the *machine learning* techniques to predict the redemption rate per individual according to several explanatory variables. The redemption basis available to us contains the sex, age, maturity of the contract, the amount of mathematical provision and the presence or absence of payment and arbitration made on the contract.

We have selected two categories of models :

— The set-theoretic methods that make it possible to transform a simple model such as decision trees into a more efficient model capable of modeling the complex links existing between the variable to be explained and the explanatory variables. We will use 3 models based on this general principle : *Random Forest*, *Gradient Boosting Machine* and *Xgboost* ;

— Neural networks (*Deeplearning*) that can model strong nonlinear interactions between the explanatory variables and the variable to be explained.

The approach used to predict redemption rates is as follows :

— The processing of the database ;

— The sampling phase ;

— The learning and calibration phase of the parameters ;

— The prediction phase.

We conducted the learning between 2010 and 2016. The test is performed in 2017. The average redemption rate observed is 4.6%. The results are as follows :

Model	RMSE	Average redemption rate
Random Forest	0,1843	4,53%
Gradient Boosting	0,18416	4,56%
Xgboost	0,18428	4,53%
Deeplearning	0,18431	4,49%

The maturity buy-back act proposes an average redemption rate of 4.84% over the 2017 year. The predicted average redemption rates (based on the six explanatory variables) are all lower than the redemption rate actual average redemption but are closer to the actual rate observed than the average rate proposed by the buy-back law according to maturity. By taking maturity only as an explanatory variable, the redemption rate is overestimated. RMSE (*root mean square error*) is the indicator that allows to measure the quality of models (to be minimized). Of the models tested, the *Gradient Boosting Machine* has the best performance in terms of squared error.

The set methods are very efficient in determining the explanatory power of the variables. The *Gradient Boosting* provides as result :

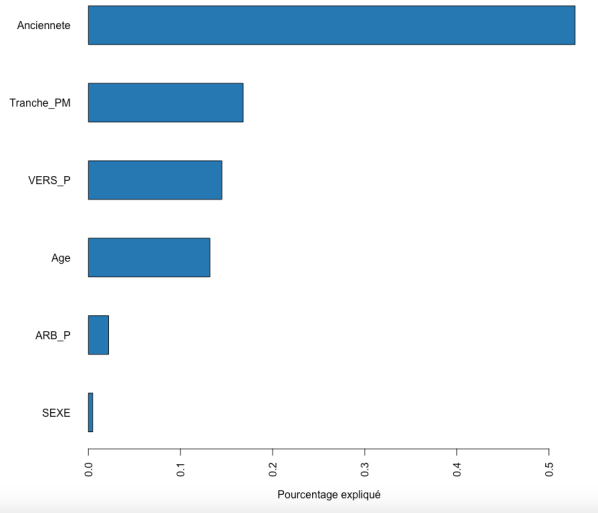

FIGURE 10 – Importance of the variables.

The maturity variable explains about 50% of the prediction of the redemption rate per individual. The amount of mathematical provision, the presence or absence of payment and the age are also important in the prediction of the rate. This analysis confirms the idea of taking into account several explanatory variables in the estimation of the structural redemption rate. In addition, *machine learning* models provided a better estimate of redemption rates than the law based on maturity in 2017.

The impact of changing buy-back laws on economic indicators

We then compared the values of the economic indicators resulting from the projection model by modifying the law of structural redemptions. The laws built are compared to a law of reference which is the law of redemptions depending on the maturity of the insurer (without taking into account tax developments). By construction of the laws, once the age of the *model points* reaches 16 years, the redemption rates are equal to 3.4 %.

The results from the projection model using the reference law and the adapted buy-back law are as follows :

Indicators	Reference situation	Appropriate buy-back law
Best Estimate	18 059 millions €	18 060 millions €
PVFP	757 millions €	755 millions €
Solvency ratio	121,2 %	121,0 %

By using the appropriate buy-back law, the solvency and profitability of the insurer decrease slightly. The associated redemption benefits are higher at the beginning of the projection. Indeed, the presence of contracts with less than 8 years of tax maturity induces this distinction. Actually, discounting impacts benefits less at the beginning than at the end of the period.

We then compared the value of the indicators with the laws built by the *machine learning* techniques. The results obtained are given below :

Indicators (M €)	Reference	*RF*	*GBM*	*XGB*	*DL*
Best Estimate	18 059	18 185	18 197	18 186	18 191
PVFP	757	631	618	629	624
Solvency ratio	121,2 %	114,3 %	113,7 %	114,2 %	114,0 %

All the indicators resulting from the projection with the laws *machine learning* are degraded compared to the indicators built with the law of redemptions according to the maturity. Solvency and profitability decrease significantly. To deepen our analysis, we propose to note the evolution of the average redemption rate predicted at each time step during the projection :

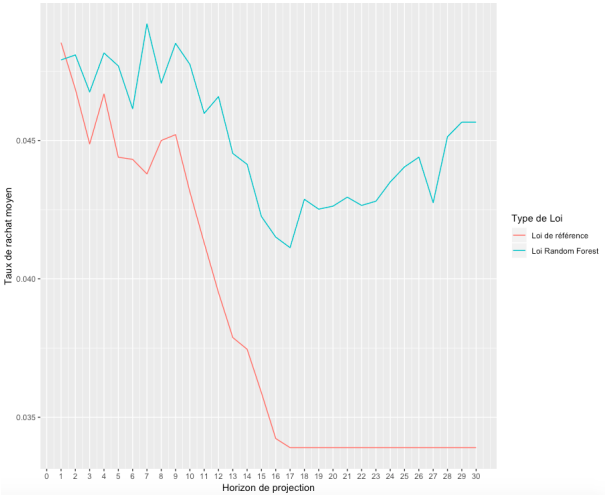

FIGURE 11 – Comparison of average redemption rates *Random Forest*.

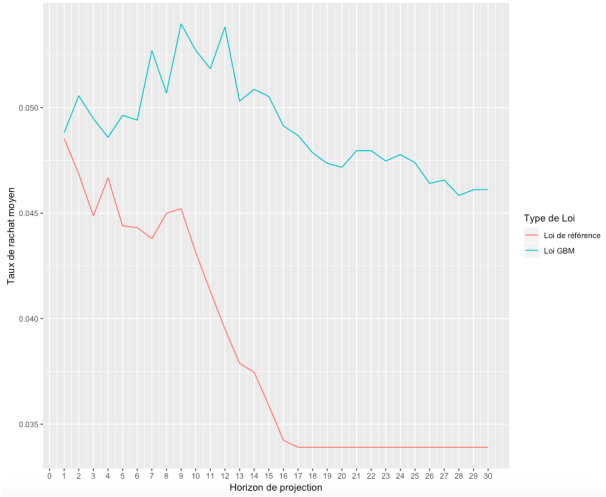

FIGURE 12 – Comparison of average redemption rates *GBM*.

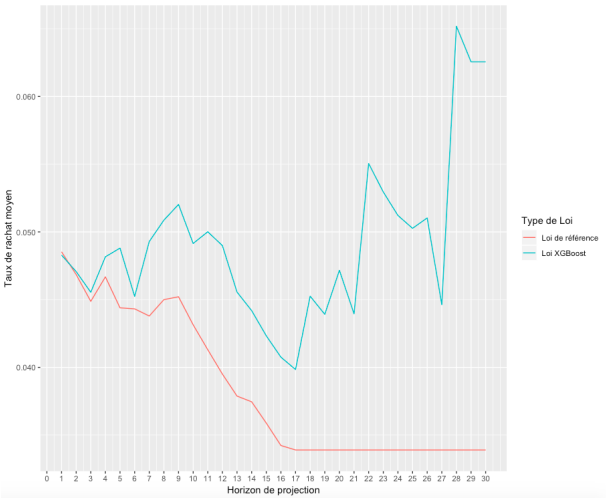

FIGURE 13 – Comparison of average redemption rates *Xgboost*.

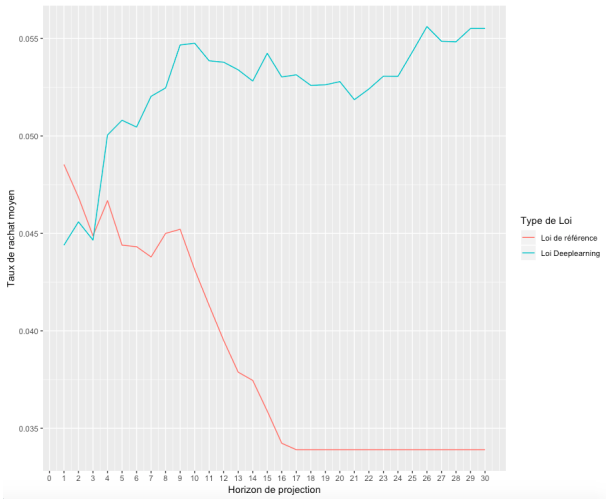

FIGURE 14 – Comparison of average redemption rates *Deeplearning*.

The evolution of the average redemption rates under the structural buyback explains the decline in profitability. Indeed, apart from the *Deeplearning* model over the first three years of projection, we can see that the predicted redemption rates are higher than the redemption rates provided by the law according to maturity. Our portfolio is sensitive to the increase of redemptions so profitability decreases.

The advantage of using the *machine learning* in the prediction of the redemption rates is that at each time step, the rates are estimated by taking into account the evolution of other variables that maturity as the age or amount of mathematical provision. We found on the graphs representing the evolution of the average redemption rate the contribution of *machine learning*. From the moment all portfolio contract years have reached 16, the rate becomes constant until the end of the projection. The methods of *machine learning* make it possible to give weight to each variable. Once all the contracts in the portfolio have maturity greater than 16 years, instead of giving the same redemption rate until the end of the projection, the models will also use the other variables to predict the redemption rate and so get different rates at each time step.

Conclusion

In this paper, we have proposed two alternative methods to the maturity buy-back law for estimating the structural redemption rate. The interest of the first method is to take into account the evolution of taxation in the law of structural redemptions according to the maturity. The second makes it possible to take into account other explanatory variables than maturity in the estimation of redemption rates. The laws *machine learning* provided a better estimate of the redemption rates compared to the law of structural redemptions according to maturity over the year 2017. The approach *machine learning* seems to be a good alternative for the construction of the law of structural redemptions. Finally, we found that a change in the law of structural redemptions could have a significant impact on the economic indicators of an insurer. Indeed, the use of laws *machine learning* has led to a decline in the profitability and solvency of our insurer.

A next step to complete our work would be to carry out a study on the estimation of the overall surrender rates (structural and cyclical) taking into account, in addition to the structural factors, variables describing the economic conditions of the market (rate served, interest rate growth, etc.).

Keywords : Flat Tax, Life Insurance, Taxation, Coverage Ratio, Present Value of Future Profits, Structural Redemptions, Machine Learning, Set-up Methods, Neural Networks, projection model.

Table des matières

Remerciements **i**

Résumé **ii**

Abstract **iii**

Synthèse **iv**

Summary **xi**

Introduction **1**

I Le contexte assurantiel de la « Flat Tax » en assurance vie **3**

1 Les différents types de contrats d'assurance vie **5**
1.1 Le contrat monosupport 5
1.2 Le contrat multisupport 5
1.3 L'arbitrage entre supports 6

2 L'option de rachat : garantie offerte à l'assuré **7**
2.1 Le risque de rachats 7
2.2 Le rachat structurel 8
2.3 Le rachat conjoncturel 8
2.4 La modélisation traditionnelle des deux lois de rachats 9

3 Les indicateurs de solvabilité et de richesse d'un assureur vie **11**
3.1 La directive Solvabilité II 11
3.1.1 Le bilan prudentiel 12
3.1.2 Le *Solvency Capital Requirement* 14
3.1.3 Le ratio de couverture comme indicateur de solvabilité 19
3.2 La Present Value of Future Profits comme indicateur de richesse 20

4 La loi de finances 2018 : "Flat Tax" **21**
4.1 La fiscalité de l'assurance vie avant le 27 septembre 2017 21
4.1.1 Les prélèvements sociaux 21
4.1.2 Le taux d'imposition 21
4.2 La mise en place de la "Flat Tax" 22
4.2.1 Les modalités de la "Flat Tax" en assurance vie 22
4.2.2 La mise en place de la « Flat tax » dans une simulation de rachat . 24

II Les méthodes de modélisation du comportement de rachat des assurés suite à l'entrée en vigueur de la « Flat Tax » 28

5 L'adaptation de la loi de rachats structurels en fonction de l'ancienneté 30
5.1 La construction de la loi de rachats structurels en fonction de l'ancienneté . 30
5.1.1 Les caractéristiques de la base de rachat 30
5.1.2 Le retraitement des données 30
5.1.3 La détermination des taux de rachat 31
5.2 La prise en compte de l'évolution de la fiscalité dans la loi construite 36

6 L'utilisation des techniques de *machine learning* 39
6.1 Présentation du *machine learning* 39
6.1.1 Une méthode d'apprentissage des données 39
6.1.2 Le type d'apprentissage 40
6.1.3 L'évaluation d'un modèle 40
6.1.4 Le choix des méthodes d'apprentissage 44
6.2 Présentation des modèles de prédiction 45
6.2.1 Les arbres de décision 45
6.2.2 Les *Random Forest* 47
6.2.3 Le *Gradient Boosting Machine* 48
6.2.4 *XGBoost* 49
6.2.5 Le *Deeplearning* 51

III L'évaluation des indicateurs économiques d'un assureur vie 57

7 La présentation du modèle de projection du portefeuille de contrats 59
7.1 Les caractéristiques du portefeuille de contrats 59
7.1.1 La composition de l'actif de l'assureur 59
7.1.2 La revalorisation des contrats 60
7.1.3 Les prestations 61
7.1.4 Les frais, chargements et commissions 63
7.1.5 L'estimation de la provision *Best Estimate* et de la *Present Value of Future Profits* 63
7.2 Le Générateur de Scénarios Économiques 65
7.2.1 La génération des taux courts nominaux 65
7.2.2 La génération de l'indice actions 68
7.2.3 La corrélation entre les actifs 71

8 Le résultat de la projection de référence en date initiale 72
8.1 La provision *Best Estimate* et la *Present Value of Future Profits* 72
8.1.1 La provision *Best Estimate* 72
8.1.2 La *Present Value of Future Profits* 73
8.1.3 La fuite de modèle 74
8.2 Le SCR et le ratio de couverture 75
8.3 Le bilan prudentiel 76

IV L'application des méthodes à un contrat d'assurance vie en euros 77

9 L'utilisation des techniques de *machine learning* pour construire les nouvelles lois de rachats structurels 78
9.1 La phase d'apprentissage et le choix des paramètres 78

9.1.1 La phase d'apprentissage 78
9.1.2 Le choix des paramètres 80
9.2 Les résultats des modèles d'apprentissage 82
9.2.1 Les *Random Forest* 82
9.2.2 Le *Gradient Boosting Machine* 82
9.2.3 *XGBoost* 83
9.2.4 Le *Deeplearning* 85
9.2.5 La comparaison des modèles 85

10 La sensibilité des indicateurs de solvabilité et de richesse au changement de la loi de rachats structurels 87
10.1 L'effet de la loi de rachats adaptée sur les indicateurs économiques 87
10.2 L'effet des lois *machine learning* sur les indicateurs économiques 90

Conclusion 93

Liste des abréviations, des sigles et des symboles 95

Annexe A 96

Annexe B 96

Annexe C 98

Bibliographie 100

Table des figures

1 Évolution du taux de rachat par ancienneté du contrat. v
2 Comparaison des deux lois de rachats. vi
3 Importance des variables. viii
4 Comparaison des taux de rachat moyens *Random Forest* avec la loi de référence. ix
5 Comparaison des taux de rachat moyens *GBM* avec la loi de référence. . . ix
6 Comparaison des taux de rachat moyens *XGBoost* avec la loi de référence. ix
7 Comparaison des taux de rachat moyens *Deeplearning* avec la loi de référence. ix
8 Evolution of the repurchase rate by maturity of the contract. xii
9 Comparison of the two buy-back laws. xiii
10 Importance of the variables. xv
11 Comparison of average redemption rates *Random Forest*. xvi
12 Comparison of average redemption rates *GBM*. xvi
13 Comparison of average redemption rates *Xgboost*. xvi
14 Comparison of average redemption rates *Deeplearning*. xvi
15 Évolution des versements par support. 4

2.1 Lois de rachats en fonction de l'ancienneté. 9
2.2 Lois de rachats conjoncturels. 10
2.3 Représentation graphique des lois de rachats conjoncturels. 10

3.1 Les 3 piliers de Solvabilité II. 11
3.2 Le bilan prudentiel simplifié. 12
3.3 La décomposition du SCR selon la formule standard pour une entreprise d'assurance vie. 15
3.4 Courbe des taux centrale et choquée au 31/12/2017. 17

4.1 Tableau d'imposition avant la réforme. 22
4.2 Schéma explicatif de la nouvelle fiscalité. 23

5.1 Évolution du taux de rachat annuel sur la période 2010 - 2017. 32
5.2 Évolution du taux de rachat par sexe. 32
5.3 Évolution du taux de rachat annuel par tranche d'encours. 33
5.4 Évolution du taux de rachat par âge de l'assuré. 34
5.5 Évolution du taux de rachat par ancienneté du contrat. 34
5.6 Comparaison des deux lois de rachats. 37

6.1 La notion de surapprentissage : résultat de la simulation. 42
6.2 Exemple d'arbre de décision. 46
6.3 Perceptron simple. 52

6.4 Perceptron multicouches. 53

7.1 Représentation graphique de loi de rachats conjoncturels. 62
7.2 Caractéristiques des trajectoires stochastiques du taux court nominal au 31/12/2017. 67
7.3 Test de martingalité pour les taux courts. 68
7.4 Caractéristiques des trajectoires stochastiques des performances actions au 31/12/2017. 69
7.5 Test de martingalité pour les actions. 70

8.1 Convergence du BE. 73
8.2 Convergence du PVFP. 74
8.3 Bilan prudentiel au 31/12/2017. 76

9.1 Variables contenues dans la base initiale. 79
9.2 Importance des variables Random Forest. 80
9.3 Importance des variables XGBoost. 80
9.4 Importance des variables RF. 82
9.5 RMSE en fonction du nombre d'arbres. 82
9.6 Importance des variables GBM. 83
9.7 RMSE en fonction du nombre d'arbres. 83
9.8 Importance des variables XGBoost. 84
9.9 RMSE en fonction du nombre d'arbres. 84
9.10 Importance des variables Deeplearning. 85
9.11 RMSE en fonction du nombre d'arbres. 85

10.1 Comparaison des taux de rachat moyens issus de la projection. 88
10.2 Chronique des flux de prestation du rachat. 89
10.3 Comparaison des taux de rachat moyens *Random Forest* avec la loi de référence. 90
10.4 Comparaison des taux de rachat moyens *GBM* avec la loi de référence. . . 90
10.5 Comparaison des taux de rachat moyens *Xgboost* avec la loi de référence. . 91
10.6 Comparaison des taux de rachat moyens *Deeplearning* avec la loi de référence. 91
10.7 Cas 1 : Nouveaux contrats 96
10.8 Cas 2 : Nouveaux contrats 96
10.9 Cas 3 : Anciens contrats 96
10.10Cas 4 : Anciens contrats 96
10.11Cas 5 : Anciens contrats 96
10.12Cas 6 : Anciens contrats 96
10.13Chonique des prestations RF 99
10.14Chronique des prestations GBM 99
10.15Chronique des prestations XGBoost 99
10.16Chronique des prestations Deeplearning 99

Liste des tableaux

3.1 La matrice de corrélation des chocs sur les indices actions 18

7.1 Caractéristiques des *model points* 59

8.1 SCR des sous-modules de risque marché. 75
8.2 SCR des sous-modules de risque vie. 75
8.3 SCR des modules de risque. 75

10.1 Comparaison des indicateurs avec la loi de rachats adaptée. 88
10.2 Comparaison des indicateurs avec les lois *machine learning*. 90
10.3 Matrice de corrélation des modules de risque composant le SCR 97
10.4 Matrice de corrélation des sous-modules de risque composant le SCR marché 97
10.5 Matrice de corrélation des sous-modules de risque composant le SCR souscription vie 98

Introduction

Dans sa loi de finances 2018, le gouvernement a instauré un prélèvement forfaitaire unique d'imposition (PFU, également appelé « Flat tax ») des revenus mobiliers de 30% se décomposant en un taux d'impôt sur le revenu de 12,8% auquel s'ajouteront les prélèvements sociaux au taux de 17,2%. Les contribuables, dont le niveau d'imposition résultant du barème serait plus favorable, auront toujours la possibilité de déclarer les plus-values mobilières au barème de l'impôt sur le revenu.

Les produits des contrats d'assurance vie afférents à de nouveaux versements à compter du 27 septembre 2017 sont concernés par l'application du prélèvement forfaitaire unique. Lorsque le montant total des encours, net des produits, détenu par l'assuré ne dépasse pas le seuil de 150 000 € (sur l'ensemble des contrats d'assurance vie), le prélèvement forfaitaire dérogatoire de 7,5% continue de s'appliquer aux produits contenus dans un rachat sur un contrat d'une durée supérieure à 8 ans. Les produits issus de nouveaux versements perçus par un assuré qui, au 31 décembre de l'année précédant le rachat, détient un encours supérieur au seuil de 150 000 € seront imposés à 12,8% au prorata de l'encours dépassant le seuil.

Jusqu'à présent, la fiscalité des contrats d'assurance vie dépendait de l'ancienneté du contrat souscrit. Le comportement de rachat des assurés pouvait dépendre de l'ancienneté de leur contrat. C'est ainsi que les assureurs modélisaient les lois de rachats structurels (rachats observés indépendamment des conditions de marché) pour alimenter leurs modèles de projection en vue du calcul de leur ratio de couverture (indicateur de solvabilité de l'assureur) et de leur *Present Value of Future Profit* (valeur actuelle des profits futurs générés par le portefeuille de contrats). Le prélèvement forfaitaire unique modifie les compartiments fiscaux. Le compartiment fiscal à 4 ans est supprimé et la différence entre les taux fiscaux des compartiments avant et après 8 ans d'ancienneté est réduite (voire supprimée à partir de 150 000 € versés sur le contrat).

L'entrée en vigueur de la « Flat tax » devrait modifier le comportement des détenteurs de contrats en euros. Les taux de rachat, la politique commerciale et la plupart des grands postes des bilans comptables et prudentiels devraient en être affectés.

La mise en place de ce prélèvement forfaitaire unique remet en cause l'utilisation de l'ancienneté du contrat comme seul facteur explicatif pour la construction de la loi de rachats structurels des assureurs. L'objectif de ce mémoire est de proposer une adaptation de la modélisation des rachats structurels et d'en estimer les effets sur les indicateurs de solvabilité et de richesse d'un assureur vie. Pour se faire, nous utiliserons un portefeuille de contrats d'épargne en euros d'un assureur anonyme.

Dans ce mémoire, nous présenterons les principes de base de l'assurance vie en définissant le vocabulaire utilisé, en introduisant la notion de rachat et en présentant la nouvelle fiscalité de l'assurance vie. Nous proposerons deux méthodes pour modéliser la loi de rachats structurels de l'assureur. La première est d'adapter la loi de rachats en fonction de l'ancienneté. La seconde est d'utiliser une approche *machine learning* pour prendre en compte d'autres facteurs explicatifs que l'ancienneté du contrat dans le comportement de rachat des assurés. Nous présenterons également l'évaluation du ratio de couverture sous Solvabilité II et de la *Present Value of Future Profits*. Nous appliquerons ces méthodes d'évaluation à un portefeuille d'un assureur vie dans la situation de référence, c'est à dire en utilisant la loi de rachats structurels de l'assureur dépendant de l'ancienneté, puis avec les nouvelles lois de rachats construites afin de comparer et d'analyser l'évolution des indicateurs de solvabilité et de richesse dans les deux situations.

Première partie

Le contexte assurantiel de la « Flat Tax » en assurance vie

Dans cette partie, les différentes notions utilisées lors de l'étude seront présentées.

Le contrat d'assurance vie est un placement financier doté d'avantages en matière de transmission du capital et de fiscalité. L'investissement se fait par des versements qui peuvent être programmés ou libres. La durée d'un contrat peut être viagère ou fixe.
C'est le produit d'assurance vie qui est le plus répandu en France (1 676 milliards d'euros en 2017).

Selon la Fédération Française de l'Assurance, le montant des cotisations nettes collectées par les sociétés d'assurances en 2017 s'élève à environ 133,9 milliards d'euros et a diminué de 1,79% par rapport à 2016.
Les versements sur les supports en unités de compte représentent 36,7 milliards d'euros soit 28% des cotisations (20% en 2016).

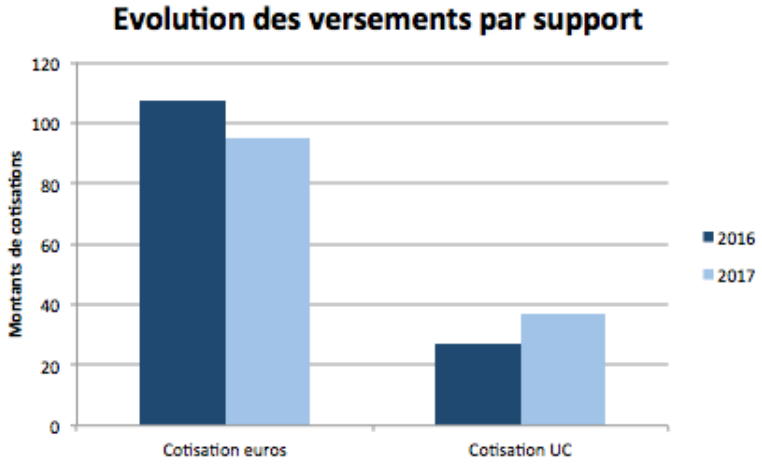

FIGURE 15 – Évolution des versements par support.

Sur l'année 2017, les prestations (décès, rachats,...) versées par les sociétés d'assurances ont augmenté de 6% par rapport à 2016. L'annonce de la loi de finance 2018 est certainement une cause de cette hausse. En effet, les détenteurs des contrats en euros pouvaient s'attendre à une baisse des avantages fiscaux et ont racheté une partie de leur épargne.

Toutefois, malgré le fait que l'épargne française soit majoritairement investie en assurance vie du point de vue des montants, le nombre de contrats ouverts est inférieur au nombre de contrats d'épargne bancaire classique (80% des français détiennent un livret A contre 40% qui détiennent une assurance vie).

Chapitre 1

Les différents types de contrats d'assurance vie

Les contrats d'assurance vie se divisent en deux grandes familles : les contrats monosupport et multisupport.

1.1 Le contrat monosupport

Le contrat monosupport propose un seul support d'investissement : le fonds en euros ou fonds général. Ce support correspond à un profil sécuritaire. Le capital placé ne peut être perdu et l'assureur garanti un taux minimum de rémunération (TMG) supérieur ou égal à 0. L'assuré porte peu de risque car celui-ci ne prend que le risque de défaut de l'assureur. De son côté, celui-ci investit les fonds principalement en obligations donc le rendement du portefeuille dépend majoritairement des taux obligataires. Ce type de contrat bénéficie d'une double garantie :

— Le taux minimum garanti auquel s'ajoute une participation aux bénéfices ;
— Un effet cliquet. En fin de chaque année, le capital du souscripteur est revalorisé et cette revalorisation est acquise.

1.2 Le contrat multisupport

Le contrat multisupport contient au moins un fonds en euros et un fonds en unités de compte (UC). Cela permet à l'assuré de constituer une épargne diversifiée et de choisir selon son aversion au risque parmi les différents supports proposés. Les assureurs proposent en particulier ce type de contrat avec des fonds « profilés » qui ont pour objectif d'investir sur des supports en unités de compte en fonction de la stratégie d'allocation d'actif choisie (prudente, équilibrée, dynamique).

Contrairement au fonds en euros, il n'existe pas de garantie de taux minimum de rendement, ni d'effet cliquet sur les supports en UC. Le contrat multisupport permet d'investir et de répartir son épargne entre différents supports tout en bénéficiant de possibilités d'arbitrage et de diversification. Étant donné que le choix d'investissement est plus large, à long terme, l'espérance de rendement est supérieure à celle du contrat monosupport car l'assuré aura la possibilité de profiter des opportunités de marché. Les contrats multisupports permettent d'investir son épargne sur le marché boursier et immobilier tout en

profitant de la fiscalité de l'assurance vie.

L'assureur n'est engagé que sur le nombre de parts d'UC et l'assuré supporte l'intégralité des risques financiers. C'est donc sur les fonds en euros que l'assureur porte le risque. Seul ce type de support sera utilisé dans notre étude.

1.3 L'arbitrage entre supports

Les supports en unités de compte (UC) sont exposés aux mouvements des marchés financiers, un risque de perte est présent. Ni le capital ni les intérêts ne sont garantis. Arbitrer son capital signifie modifier sa répartition entre les différents supports. Il est possible d'arbitrer son capital depuis le fonds en euros vers UC, depuis UC vers le fonds euros ou depuis UC vers d'autres UC. Cette option permet pour un assuré de transférer son capital vers d'autres supports pour sécuriser les gains réalisés sur les UC ou dynamiser ses revenus.

Lors d'une mauvaise période sur les marchés financiers, plutôt que de voir un rachat massif se faire, on peut observer un arbitrage des contrats des assurés.

Chapitre 2

L'option de rachat : garantie offerte à l'assuré

L'option de rachat est la possibilité offerte aux assurés d'effectuer un rachat partiel ou total de leurs contrats d'assurance vie. C'est à dire que l'assuré peut à tout moment récupérer tout ou une partie de l'épargne acquise sur son contrat, appelée valeur de rachat. En cas de rachat partiel, le contrat se poursuit dans les mêmes conditions, tandis que le rachat total met fin au contrat. Selon l'ancienneté et la nature du contrat, le rachat peut avoir des conséquences fiscales pour l'assuré.

Les rachats observés sur les contrats d'assurance vie peuvent être décomposés en deux catégories :

— Les rachats structurels ou statiques ;
— Les rachats conjoncturels ou dynamiques.

2.1 Le risque de rachats

Afin de garantir les intérêts promis à l'assuré, l'assureur investit les primes sur le marché. Lorsque l'option de rachat est déclenchée, l'assureur paye les prestations avec les primes perçues ou avec sa trésorerie. Néanmoins, si le montant de prestations est trop élevé, il sera amené à vendre ses actifs pour récupérer la valeur du rachat. Cela peut présenter deux situations défavorables possibles. En effet, si un actif peu liquide doit être vendu rapidement, sa valeur de marché peut baisser et si les titres sont en situation de moins-value latente, la vente va faire réaliser cette moins-value à l'assureur.

Ces deux phénomènes permettent d'illustrer le risque de rachat. Il est important pour l'assureur de maîtriser ce risque en le modélisant. Pour se faire, il est primordial de comprendre au mieux le comportement des assurés et de visualiser quels facteurs motivent ces derniers à racheter leurs contrats.

2.2 Le rachat structurel

Les rachats structurels correspondent au comportement moyen des assurés du portefeuille, indépendamment des conditions de marché. La décision de rachat peut être motivée par plusieurs facteurs :

— Un besoin de liquidité de l'assuré ;
— Des dispositions fiscales plus ou moins favorables : le volume de rachats partiels augmente après les 8 premières années d'ancienneté du contrat en raison de la fiscalité plus avantageuse ;
— L'âge de l'assuré : un assuré qui atteint un âge avancé entre dans une logique de transmission de patrimoine et souhaite bénéficier d'une fiscalité avantageuse sur les droits de succession ;
— Le changement de situation de l'assuré tels que la capacité à payer les primes, le chômage, le divorce peut influencer le comportement de rachat.

Les contrats d'assurance vie sont souvent souscrits dans le cadre d'un projet à moyen long terme. C'est à dire que lorsqu'il y a un rachat, le projet du détenteur du contrat est réalisé (retraite, étude des enfants ...). Généralement, ces projets ont un horizon de réalisation supérieur à 8 ans.

Les rachats structurels sont les rachats observés dans un contexte économique « normal ».

2.3 Le rachat conjoncturel

Les rachats conjoncturels correspondent à l'augmentation ou à la diminution des rachats dans un contexte concurrentiel lorsque l'assuré arbitre son contrat d'assurance vers d'autres supports financiers (produits assuranciels, bancaires ou immobiliers). Cet arbitrage peut être motivé par :

— Les fluctuations sur le marché des taux d'intérêt : si les taux montent, les assurés auront tendance à racheter leur contrat pour accéder à des taux plus compétitifs ;
— D'autres produits plus rentables, exonérés de taxes (PEA [1]) ;
— La solvabilité de l'assureur : la détérioration de l'état de solvabilité d'un assureur peut inciter un assuré à racheter son contrat.

Les rachats conjoncturels dépendent du comportement des assurés en réponse à l'écart constaté entre le taux de revalorisation servi par l'assureur et un taux de marché. Lorsque le taux de marché est supérieur au taux servi par l'assureur, l'assuré peut décider de racheter son contrat. Pour un écart entre le taux servi par l'assureur et le taux de marché de l'année, la totalité des assurés n'agira pas de la même façon.

Le rachat est un risque non réplicable car il dépend exclusivement du comportement des assurés. Puisque la gestion des rachats dynamiques est harmonisée dans le QIS 5 [2], c'est en particulier le risque de rachats structurels qui doit être évalué de manière robuste par les assureurs.

1. Plan Épargne Action
2. Les spécifications techniques du QIS 5 (5e étude quantitative d'impact) fournissent des formules fermées pour modéliser les rachats dépendant des taux servis et des conditions des marchés financiers.

2.4 La modélisation traditionnelle des deux lois de rachats

Afin d'évaluer le coût de l'option de rachat, les assureurs doivent modéliser le comportement des assurés. La Directive Solvabilité II impose que la projection des flux futurs tienne compte de la probabilité de rachat.

Loi de rachats structurels

Les rachats structurels, observés dans un contexte économique « normal », sont modélisés grâce à l'expérience passée de l'assureur. Les paramètres dont les hypothèses de rachat peuvent dépendre sont nombreux : l'âge de l'assuré, l'ancienneté du contrat, la gamme de produits, la provision mathématique initiale etc.

Néanmoins, les assureurs modélisent en général le rachat structurel en fonction d'un seul facteur explicatif qui est l'ancienneté du contrat. Avec l'entrée en vigueur de la « Flat Tax », la fiscalité ne dépend pratiquement plus de l'ancienneté du contrat amenant les assureurs à adapter la modélisation de la loi de rachats structurels.

La figure ci-dessous, publiée par la Fédération Française de l'Assurance (FFA), représente la loi de rachats en fonction de l'ancienneté.

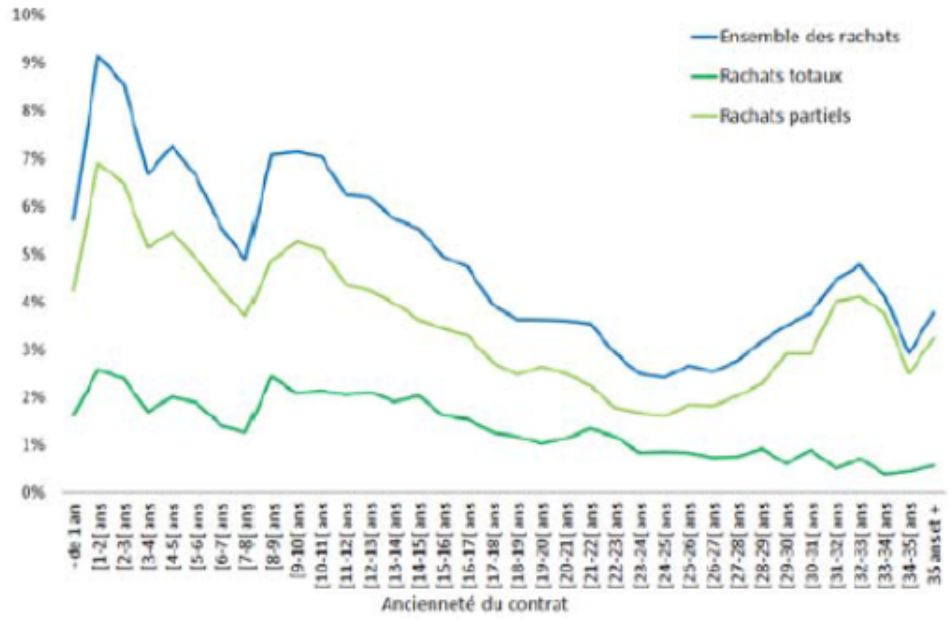

FIGURE 2.1 – Lois de rachats en fonction de l'ancienneté.

Loi de rachats conjoncturels

Les rachats conjoncturels ne peuvent être estimés à partir de l'historique des rachats. En effet, la simulation du taux de rachat conjoncturel dans le futur nécessite de projeter le taux intervenant sur le marché, le taux servi par l'assureur et de définir une loi de comportement qui, en fonction du taux servi (TS) et du taux attendu, associe un taux de rachat conjoncturel. Si le taux servi est inférieur au taux attendu, les assurés auront tendance à racheter davantage que dans une situation normale. Les Orientations Nationales

Complémentaires aux spécifications Techniques fournies par l'ACPR proposent de définir le taux de rachat conjoncturel (RC) comme suit :

$$RC(R) = \begin{cases} RC_{max} & si\ R - TA < \alpha \\ RC_{max} \cdot \dfrac{R - TA - \beta}{\alpha - \beta} & si\ \alpha < R - TA < \beta \\ 0 & si\ \beta < R - TA < \gamma \\ RC_{min} \cdot \dfrac{R - TA - \gamma}{\delta - \gamma} & si\ \gamma < R - TA < \delta \\ RC_{min} & si\ R - TA > \delta \end{cases}$$

Paramètre	Maximum	Minimum
α	-4%	-6%
β	0%	-2%
γ	1%	1%
δ	4%	2%
RC_{min}	-4%	-6%
RC_{max}	40%	20%

FIGURE 2.2 – Lois de rachats conjoncturels.

Graphiquement les lois sont représentées ainsi :

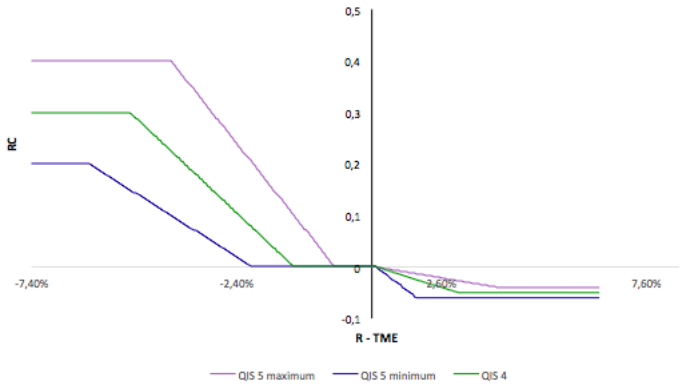

FIGURE 2.3 – Représentation graphique des lois de rachats conjoncturels.

Le rachat conjoncturel défini par la loi proposé par QIS4 (courbe verte) est nul quand l'écart entre le taux servi et le taux attendu est compris entre - 1% et 0,5%. Il augmente dès lors que ce même écart est plus petit que - 1% et diminue quand l'écart est plus grand que 0,5%.

Le taux de rachat total est alors obtenu en additionnant le taux de rachat conjoncturel avec le taux de rachat structurel.

Chapitre 3

Les indicateurs de solvabilité et de richesse d'un assureur vie

3.1 La directive Solvabilité II

La solvabilité est la capacité de l'assureur à faire face à ses engagements envers ses assurés. La directive Solvabilité II a pour objectif principal la meilleure adaptation des fonds propres exigés des compagnies d'assurances face aux risques auxquels elles sont exposées dans leur activité. Elle repose sur sa structure en 3 piliers :

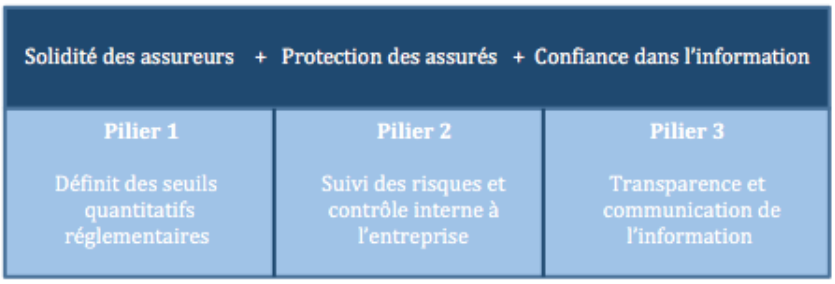

FIGURE 3.1 – Les 3 piliers de Solvabilité II.

— Le pilier 1 donne le choix entre formule standard et modèle interne pour le calcul des deux seuils réglementaires à respecter pour le montant des fonds propres : le *Solvency Capital Requirement* (SCR) et le *Minimum Capital Requirement* (MCR). En notant FP le montant de fonds propres de la compagnie d'assurance :
  - Si $SCR \leq FP$, l'assureur peut continuer à exercer son activité.
  - Si $MCR \leq FP <$ SCR, des mesures de redressement sont prises par l'Autorité de Contrôle des assurances en France (ACPR), qui peut imposer une stratégie à l'entreprise.
  - Si FP $<$ MCR, le droit d'exercer son activité d'assurance est retiré par l'ACPR.
— Le pilier 2 fixe des normes qualitatives de suivi des risques en interne (ORSA [1]) et d'exigence de gouvernance de l'entreprise.

1. *Own Risk and Solvency Assessment*, processus d'évaluation interne des risques et de la solvabilité de l'entreprise. L'ORSA demande une évaluation globale de la solvabilité compte tenu des risques auxquels elle est exposée.

— Le pilier 3 définit le détail des informations qualitatives et quantitatives à fournir aux investisseurs, assurés et autorités.

3.1.1 Le bilan prudentiel

Le référentiel Solvabilité II propose une vision économique du bilan de l'assureur : le bilan prudentiel. La directive européenne accompagnée par la commission européenne et l'EIOPA ont fixé les règles de l'établissement du bilan prudentiel. Par exemple :

— Les provisions techniques sont calculées de manière prudente. Elles sont constituées de la meilleure estimation des engagements de l'assureur, appelé *Best Estimate*, valorisée sans marge de prudence et d'une marge pour risque (*Risk Margin*).

— La valorisation du bilan doit se faire de manière « Market-Consistenty », c'est à dire de manière cohérente avec le marché.

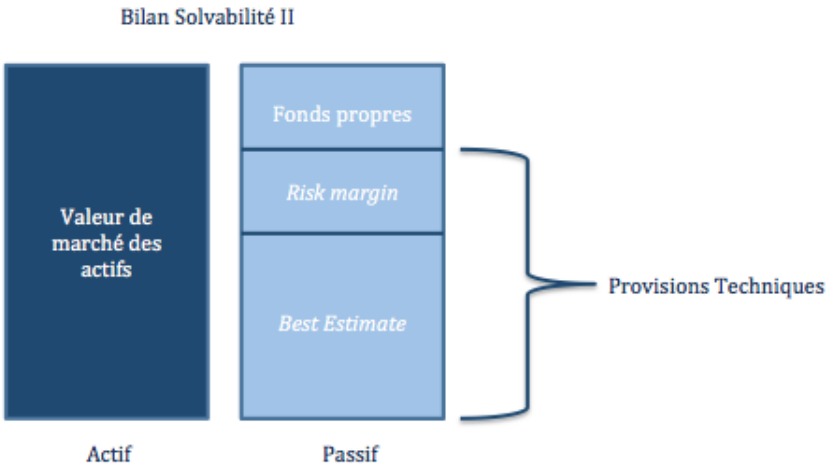

FIGURE 3.2 – Le bilan prudentiel simplifié.

• Valorisation de l'actif :

Sous Solvabilité II, les actifs de placements sont valorisés à leur Juste Valeur (valeur de marché), c'est à dire comme étant le montant pour lequel ils peuvent être échangés sur le marché.

• Valorisation du passif :

Le passif est composé des Fonds Propres et des provisions techniques (*Best Estimate* et *Risk Margin*).

◦ *Best Estimate* :

Dans le cadre de Solvabilité II, le *Best Estimate* est défini comme étant « la moyenne pondérée par leurs probabilités des flux de trésorerie futurs, compte tenu de la valeur de l'argent, estimée sur la base de la courbe des taux sans risque ». On peut exprimer le *Best Estimate* de façon mathématique :

$$BE = \mathbb{E}^{P*Q}(\mid \sum_{t=1}^{T} \Delta_t \cdot L_t \mid) \tag{3.1}$$

Avec :

— Q la probabilité risque neutre. Sous cette probabilité, la valeur actualisée du processus de flux est une martingale ;

— P la probabilité réelle qui correspond à la probabilité historique des variations des actifs observés sur le marché ;

— Δ_t le facteur d'actualisation (au taux sans risque r_h) $= e^{-\int_0^t r_h dh}$;

— L_t représente les flux correspondants aux engagements de l'assureur envers l'assuré ;

— T l'horizon de projection choisi.

Dans la projection des flux, il faut prendre en compte la revalorisation de la provision mathématique pour les contrats en euros (intérêts techniques et participation aux bénéfices).

En pratique, le *Best Estimate* est déterminé par la méthode de Monte-Carlo. Nous avons à disposition cet estimateur du *Best Estimate* :

$$BE^k = \frac{1}{k} \sum_{i=1}^{k} \sum_{t=1}^{T} \Delta_{t,i} \cdot L_{t,i} \tag{3.2}$$

Avec k le nombre de scénarios économiques et $\Delta_{t,i} = e^{-\int_0^t r_h^i dh}$ le facteur d'actualisation de la *ième* réalisation du *Best Estimate*. Ces scénarios seront générés par un Générateur de Scénarios Economiques. Les méthodes seront expliquées dans la troisième partie.

◦ La *Risk Margin* :

Dans le cadre de Solvabilité II, la *Risk Margin* est une composante des provisions techniques qui garantit que le montant total de provisions est équivalent au montant que devrait avoir une compagnie d'assurance pour assurer ses engagements envers ses assurés.

Elle consiste à mesurer le coût en fonds propres pour un investisseur qui reprendrait l'activité d'assurance en *run-off* (plus de contrats entrant dans le portefeuille).

La *Risk Margin* est calculée de manière mathématique :

$$RiskMargin = CoC \cdot \sum_{t \geq 0} \mathbb{E}^{Q*P}\left(\frac{SCR(t)}{(1 + r_{t+1})^{t+1}}\right) \tag{3.3}$$

Avec :

— CoC : Taux du coût du capital fixé à 6% ;

— Q la probabilité risque neutre ;

— P la probabilité réelle ;

— SCR(t) le SCR l'année t.

Les anciennes versions des spécifications du LTGA (*Long-Term-Guarantees-Assessment*) proposaient plusieurs simplifications pour le calcul de la *Risk Margin*. Ce n'est pas l'objet du mémoire donc nous la négligerons. La valeur des provisions techniques sera la valeur de la provision *Best Estimate*.

◦ Fonds Propres :

Les Fond Propres sont obtenus par différence en soustrayant de la valeur de l'actif les provisions techniques (formule simplifiée). C'est à dire :

$$FondsPropres = Actifvaleurmarché - Provisionstechniques \tag{3.4}$$

3.1.2 Le *Solvency Capital Requirement*

Le *Solvency Capital Requirement* (SCR) est le montant de fonds propres économiques qu'une compagnie d'assurance doit détenir pour faire face à une ruine économique à horizon 1 an dans 99,5% des cas.

Le capital économique s'écrit de la manière suivante :

$$SCR = FP_0 - P(0,1) \cdot q_{0,5\%}(FP_1) \tag{3.5}$$

— Le premier terme représente les fonds propres économiques initiaux et le second représente le surplus (algébrique) de capital à ajouter en t = 0.

— P(0,1) est le prix d'un prix zéro coupon 1 an utilisé comme facteur d'actualisation.

— $q_{0,5\%}$ est le quantile à 0,5% des fonds propres économiques en t=1.

Le SCR peut être estimé selon deux méthodes :

— Formule standard proposé par l'EIOPA

— Modèles internes totaux ou partiels

Dans ce mémoire, la formule standard sera utilisée pour déterminer le SCR. Nous présentons donc seulement cette méthode.

• Le SCR selon la formule standard

Selon la formule standard, le SCR est la somme de 3 éléments :

$$SCR = BSCR + SCR_{operationel} + Ajustement \tag{3.6}$$

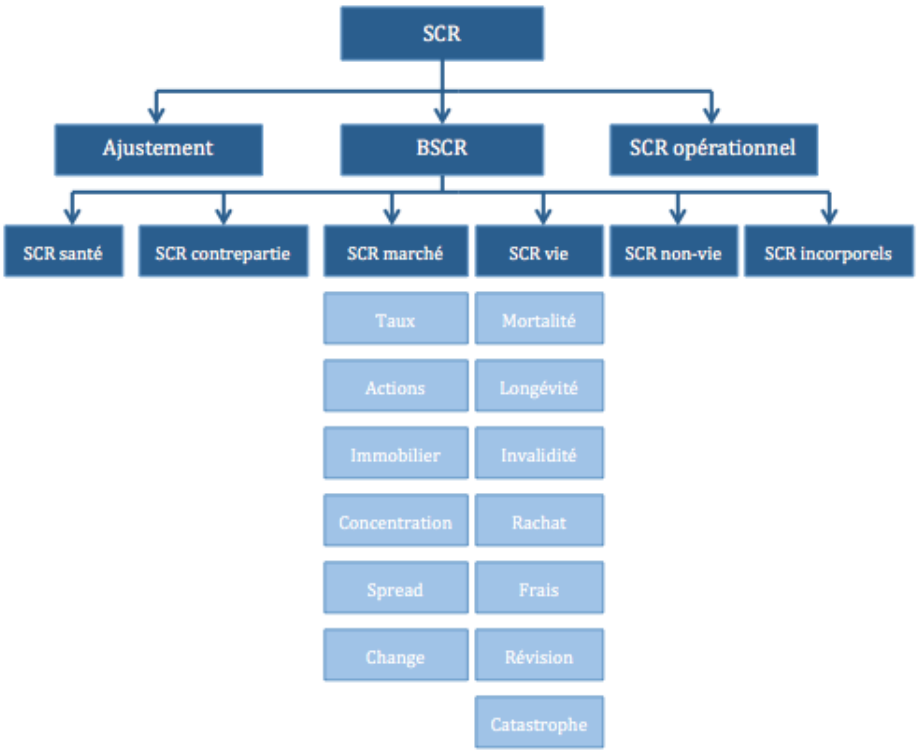

FIGURE 3.3 – La décomposition du SCR selon la formule standard pour une entreprise d'assurance vie.

• Le BSCR (*Basic Solvency Capital Requirement*) est l'agrégation de 6 modules de risque :

- Risque de marché
- Risque de souscription santé
- Risque de contrepartie
- Risque de souscription vie
- Risque de souscription non vie
- Risque lié aux immobilisations incorporelles

• Le SCR opérationnel représente les risques spécifiques aux processus internes de l'organisme et les erreurs du personnel.

• L'ajustement tient compte de la capacité d'absorption des pertes des provisions techniques et des impôts différés.

La méthode de calcul du SCR par la formule standard repose sur 3 étapes fondamentales :

— Détermination d'un capital économique pour chaque risque élémentaire (action, mortalité, taux . . .) ;
— Agrégation des capitaux au sein de chaque module de risques (marché, vie, non-vie. . .), ce qu'on appelle l'agrégation intra-modulaire ;

— Agrégation des capitaux des différents modules (agrégation inter-modulaire).

Étape 1 :

Pour déterminer le capital économique de chaque risque élémentaire (marché et souscription vie), on réalise la différence en t = 0 entre les fonds propres centraux et les fonds propres choqués (chocs définis par l'EIOPA pour chaque risque). C'est à dire, en appelant C_i le capital associé à un risque élémentaire i, on a comme formule :

$$\begin{aligned} C_i &= \Delta FP_i \\ &= \Delta Valeurmarche_Actif - \Delta BE \\ &= (Valeurmarche_Actif_{central} - Valeurmarche_Actif_{choque}) - (BE_{central} - BE_{choque}) \end{aligned}$$

Etape 2 et 3 :

On agrège ensuite les capitaux sous risques puis ces capitaux entre eux pour déterminer le BSCR.

$$SCR_k = \sqrt{\sum_{(i,j)\in\mathbb{R}_k^2} \rho_{i,j}^{\mathbb{R}_k} \cdot C_i \cdot C_j} \tag{3.7}$$

Avec SCR_k le capital associé au module k et C_i le capital associé au risque élémentaire i.

Puis à partir de tous les « SCR modules » obtenus et à l'aide de la matrice de corrélation des risques :

$$BSCR = \sqrt{\sum_{(i,j)\in K} \rho_{i,j}^{\mathbb{R}_k} \cdot SCR_i \cdot SCR_j} \tag{3.8}$$

Le BSCR correspond donc au capital global (avant SCR opérationnel et l'ajustement).

Le SCR marché

Le risque de marché résulte de l'incertitude de la variation des taux d'intérêts, du cours des actions, du prix de l'immobilier, du niveau de concentration et des taux de *spread* et de change. Le SCR marché est déterminé par agrégation des SCR sous modulaires (action, taux, etc) à l'aide de la matrice de corrélation des risques (la matrice est donnée dans l'annexe B).
Dans l'application réalisée, on considère un actif simplifié d'un assureur qui serait constitué de x% d'actions et (100-x)% d'obligations gouvernementales européennes (pas de risque de spread). Ces actifs sont répartis entre de nombreuses contreparties (pas de risque de concentration) et exprimés en € (pas de risque de change). Cet assureur n'est donc soumis qu'au risque actions et au risque de taux.

• Le risque de taux d'intérêt

Le risque de taux d'intérêt existe pour tous les actifs et passifs dont la valeur est sensible aux changements de taux d'intérêt. On retrouve les produits dérivés de taux et les obligations comme actifs concernés. Concernant le passif, les provisions techniques actualisées au taux zéro coupon sont également sensibles au risque de taux.

Pour matérialiser le choc de taux, nous recalculons la valeur des éléments de l'actif et du passif concernés en appliquant un choc de taux à la hausse et à la baisse. L'EIOPA propose d'appliquer un pourcentage de taux, pour chaque maturité, à la courbe des taux de référence. Dans la réglementation actuelle, le scénario à la hausse conduit à une variation minimale de 1% et les taux négatifs ne sont pas abaissés dans le scénario à la baisse.

Au 31/12/2017, les courbes à la hausse et à la baisse sont les suivantes :

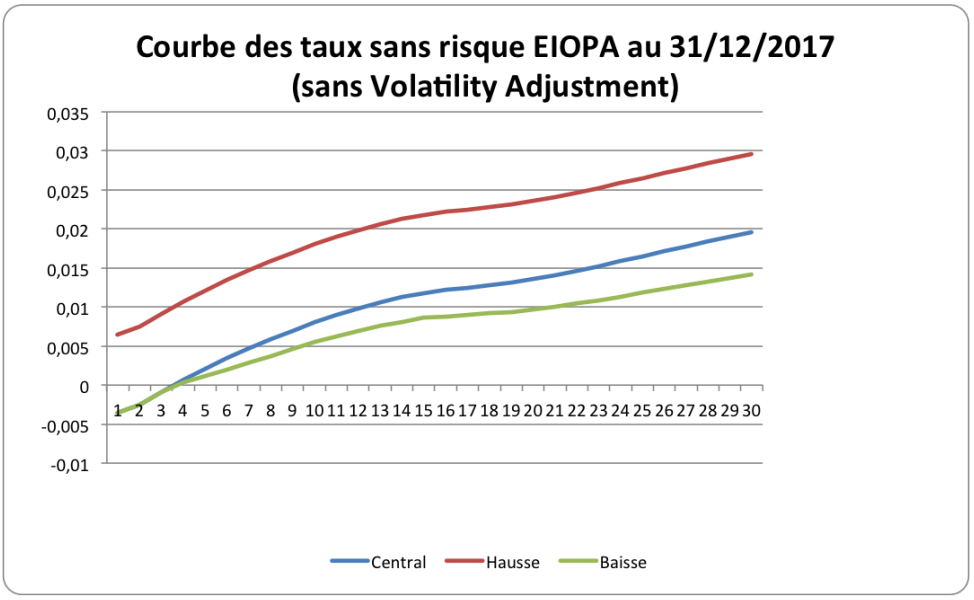

FIGURE 3.4 – Courbe des taux centrale et choquée au 31/12/2017.

Les SCR hausse et baisse des taux sont égaux à la variation de fonds propres suites à l'application des chocs. L'assureur retient le SCR hausse ou baisse des taux qui donne un SCR marché le plus important (après avoir agrégé les autres risques).

• Le risque action

Le risque sur les actions résulte de la volatilité des prix de marché des actions. L'approche standard se base sur deux indices :

— Les actions de type 1 : les actions échangées sur les marchés des pays de l'OCDE (Organisation de Coopération et de Développement Economique) et/ou de l'EEE (Espace Economique Européen).
— Les actions de type 2 : les action de pays émergents, les actions non cotées, les produits de base et autres investissements alternatifs.

Le choc est différent en fonction du type d'action. Pour les actions de type 1, il est de 39% et 49% pour les actions de type 2, modulé par un effet « d'ajustement symétrique » fonction du niveau actuel d'un indice et de la moyenne pondérée de celui-ci calculée sur 36 mois. C'est l'EIOPA qui communique cet ajustement et au 31/12/2017 il s'élevait à 1,9%. Pour chaque type, on peut déterminer le SCR. Ensuite, il suffit de les agréger avec la matrice de corrélation :

Indice	Type 1	Type 2
Type 1	1	0,75
Type 2	0,75	1

TABLE 3.1 – La matrice de corrélation des chocs sur les indices actions

Le SCR souscription vie

Le risque de souscription Vie reflète le risque découlant des engagements d'assurance vie, compte tenu des périls couverts et des procédés appliqués dans l'exercice de cette activité. Il est calculé comme résultant de la combinaison des exigences de capital applicables aux sous-modules suivants :

— Risque de mortalité
— Risque de longévité
— Risque d'invalidité
— Risque de rachat
— Risque de dépenses en vie
— Risque de révision
— Risque de catastrophe en vie

Le SCR souscription vie est obtenu par agrégation des SCR des sous modules et de la matrice de corrélation des différents sous modules du SCR souscription vie (cf annexe B).

Dans l'application réalisée, on considère un produit d'assurance vie en euros sans possibilité de conversion en rente donc l'assureur n'est pas soumis aux risques de révision et d'invalidité.

• Le risque de mortalité

Ce module est calculé pour les contrats garantissant des prestations en cas de décès, c'est à dire pour les contrats pour lesquels une augmentation des taux de mortalité entraîne une augmentation de la valeur des engagements d'assurance.
Le choc de mortalité est un choc supposant une augmentation annuelle (permanente) des taux de mortalité de 15% pour chaque âge (article 137 du règlement délégué).

• Le risque de longévité

Ce module est calculé pour les contrats garantissant des prestations en cas de vie, c'est à dire pour les contrats pour lesquels une diminution des taux de mortalité entraîne une augmentation de la valeur des engagements d'assurance.
L'exigence de capital pour risque de longévité est définie dans l'Article 138 du règlement délégué. Elle correspond à la perte de fonds propres de base de l'entreprise résultant d'une

baisse soudaine et permanente de 20 % des taux de mortalité utilisés pour le calcul des provisions techniques.

• Le risque de rachat

Le risque de rachat désigne le risque de constater un taux imprévu de rachat des polices, de cessation de paiement de primes, d'échéance ou de renouvellement. Ce taux imprévu peut être un taux plus élevé ou plus bas que le taux anticipé.
L'approche vise donc à évaluer l'impact sur l'ensemble des contrats selon trois hypothèses distinctes :

— Une baisse des taux de rachats annuels de 50% ;
— Une hausse des taux de rachats annuels de 50% ;
— Un rachat de masse de 40% la première année puis un retour aux taux de rachats central

Le SCR retenu pour le risque de cessation est le SCR le plus important entre ces trois chocs.

• Le risque de dépenses

Le risque de dépenses en vie est un risque de perte résultant de la variation des dépenses encourues pour la gestion des contrats d'assurance et de réassurance.
L'Article 140 du règlement délégué précise que le choc à appliquer correspond à la conjugaison des modifications soudaines permanentes suivantes :

— une augmentation de 10% du montant des dépenses prises en considération dans le calcul des provisions techniques ;
— une augmentation de 1% du taux d'inflation des dépenses utilisées pour le calcul des provisions techniques.

• Le risque de catastrophe

Le risque de catastrophe en vie correspond au risque de perte résultant d'une hausse absolue de la mortalité des adhérents de 0,15% durant l'année à venir.

3.1.3 Le ratio de couverture comme indicateur de solvabilité

Le ratio de couverture, également appelé ratio de solvabilité, est le rapport entre les fonds propres et le SCR. Si celui est plus petit que 100%, cela signifie que les fonds propres ne couvrent pas le SCR et que l'assureur devra fournir un plan de redressement à l'ACPR.

Si le ratio de couverture est élevé, l'assureur peut plus facilement manœuvrer son allocation d'actif et prendre plus de risques en investissant sur des actions plus risquées (l'espérance de rendement d'un actif risqué est plus élevée que celle d'un actif moins risqué).

Dans ce mémoire, le ratio de couverture est l'indicateur de solvabilité utilisé.

3.2 La Present Value of Future Profits comme indicateur de richesse

La Présent Value of Future Profits (PVFP) est définie comme « la valeur actuelle des profits futurs générés par le portefeuille de contrats calculée dans un scénario déterministe sans prime de risque ». En d'autres termes, la PVFP correspond à la valeur actuelle des résultats futurs. En effet, le résultat se calcule ainsi :

$$\text{Résultat} = \text{Produits financiers} - \text{Revalorisation} + \text{Chargements} - \text{Frais} - \text{Commission} \tag{3.9}$$

La PVFP est une des composantes de la *Market Consistent Embedded Value*.

• La *Market Consistent Embedded Value*

L'*Embedded Value* correspond à la richesse de la compagnie. Ce concept de valeur de société est utilisé depuis les années 1990. C'est le CFO Forum [2] qui a harmonisé les méthodes de calcul de l'*Embedded Value*.

En juin 2008, les principes de la MCEV (cohérence avec les données du marché) ont été publiés. La MCEV est déterminée par la formule :

$$MCEV = ANR + PVFP - TVFOG - CFCR - CRNHR \tag{3.10}$$

Avec :

— ANR : actif net réévalué

— TVFOG (Time value of financial options and guarantees) : valeur temps des options et garanties financières

— CFCR (Frictional cost of required capital) : coût de blocage du capital requis

— CRNHR (Cost of residual no hedgeable risk) : coût des risques résiduels non couvrables.

Dans ce mémoire, la PVFP est utilisée comme indicateur de richesse de l'assureur.

Dans le prochain chapitre, nous énoncerons la nouvelle fiscalité mise en place : la "Flat Tax".

2. Association européenne formée par les directeurs financiers des grandes compagnies d'assurance côtées sur le marché financier. Son objectif est d'influencer sur le développement de l'information financière pour les entreprises d'assurance qui représentent une part importante du secteur européen de l'assurance.

Chapitre 4

La loi de finances 2018 : "Flat Tax"

Afin de simplifier et d'harmoniser la fiscalité de l'épargne, le gouvernement a instauré un prélèvement forfaitaire unique : la « Flat Tax ».
Dans ce chapitre, la fiscalité avant la réforme sera rappelée puis les modalités de la « Flat Tax » seront expliquées.

4.1 La fiscalité de l'assurance vie avant le 27 septembre 2017

En cas de rachat d'un contrat assurance vie, les revenus générés par ce dernier sont soumis aux prélèvements sociaux et à l'impôt.

4.1.1 Les prélèvements sociaux

Depuis 1997, des prélèvements sociaux existaient sur les contrats monosupport euros. C'est seulement depuis 2011 que les fonds en euros des contrats multisupports sont soumis aux prélèvements sociaux.

A la fin de chaque année, lors du versement des intérêts sur le fonds euros, des cotisations sociales sont prélevées. Au moment du rachat, un complément de prélèvements sociaux est appliqué uniquement à la part de gain.

Sur le fonds en unité de compte, ce sont les plus-values qui sont soumises aux cotisations sociales.

4.1.2 Le taux d'imposition

La part de plus-value rachetée est soumise à l'impôt (sauf cas d'exonération). L'assuré a le choix entre le prélèvement forfaitaire et une imposition selon le barème progressif de l'impôt sur le revenu. Pour les rachats qui ont lieu après 8 ans, un abattement s'applique à hauteur de 4 600 euros pour un assuré seul et 9 200 euros pour un couple sur les produits générés.

Lors de chaque rachat, l'assureur calcule la part de plus-value au titre de chaque compartiment fiscal. Le taux d'imposition dépend de la date de souscription du contrat et de l'ancienneté du contrat au moment de la sortie. On peut voir ci-dessous l'évolution des compartiments et des taux d'imposition entre 1983 et le 27 septembre 2017.

<table>
<tr><th>Ancienneté au moment de la sortie</th><th>[0 ; 4 ans [</th><th>[4 ; 8 ans [</th><th>Après 8 ans</th></tr>
<tr><td>Contrats souscrits avant 1983</td><td colspan="3">Contrats exonérés d'impôt</td></tr>
<tr><td>Contrats souscrits de 1983 à 1989</td><td colspan="2">PFL 45%/ 25% / 15% en fonction de la durée moyenne pondérée ou IRPP</td><td>Versements avant 1998 : exonérés d'impot
Versements à partir de 1998 : PFL 7,5% ou IRPP après abattement (4 600 € /9 200 € pour un couple)</td></tr>
<tr><td>Contrats souscrits entre 1990 et le 26/09/1997</td><td rowspan="2">PFL 35% ou IRPP</td><td rowspan="2">PFL 15% ou IRPP</td><td>Versements avant 1998 : exonérés d'impot
Versements à partir de 1998 : PFL 7,5% ou IRPP après abattement (4 600 € /9 200 € pour un couple)</td></tr>
<tr><td>Contrats souscrits depuis le 27/09/1997</td><td>PFL 7,5% ou IRPP après abattement (4 600 € / 9 200 € pour un couple)</td></tr>
</table>

FIGURE 4.1 – Tableau d'imposition avant la réforme.

Les versements ayant lieu entre le 27/09/1997 et le 31/12/1997 sont partiellement exonérés après 8 ans.

4.2 La mise en place de la "Flat Tax"

Applicable depuis le 1er janvier 2018, le prélèvement forfaitaire unique de 30% se décompose d'un taux forfaitaire d'impôt sur revenu de 12,8% et de prélèvements sociaux au taux de 17,2%. Elle impacte les revenus mobiliers (intérêts, dividendes), les plus-values de cessions de valeurs mobilières et les gains réalisés sur les contrats d'assurance vie. Les épargnants sont par principe soumis à la « Flat tax » mais peuvent opter pour le barème progressif de l'impôt sur le revenu. Cette option concerne l'ensemble des revenus et les plus values de l'année. En assurance vie, la part d'imposition s'applique aux gains des versements réalisés depuis le 27 septembre 2017. La part de prélèvements sociaux ne s'applique qu'à partir de janvier 2018.

Dans la suite de la présentation, nous ne parlerons que de la part d'imposition étant donné que les prélèvements sociaux seront applicables à tous les contrats (avant et pendant la réforme).

4.2.1 Les modalités de la "Flat Tax" en assurance vie

Le prélèvement forfaitaire unique concerne les contrats ayant effectués des versements après le 27 septembre 2017. Pour les contrats souscrits avant la réforme, le PFU s'applique seulement aux produits issus des versements ayant lieu après le 27 septembre. Pour les autres correspondant aux versements ayant lieu avant le 27 septembre 2017, c'est l'ancien taux d'imposition qui s'applique.

Le taux d'impositon du PFU est le suivant :

— Pour les contrats disposant au 31 décembre de l'année précédente d'un ou plusieurs contrats (tous les assureurs confondus) dont le montant total des primes (non rachetées) dépasse 150 000 euros, les produits sont imposés à 12,8% quelque soit l'ancienneté du contrat.

— Pour les autres contrats, c'est à dire ceux qui n'ont pas un niveau de primes non rachetées supérieur au seuil de 150 000 euros, leurs produits sont imposés, lors d'un rachat, à hauteur de 12,8% si le contrat a une ancienneté de moins de 8 ans sinon

ils sont imposés à 7,5%.

L'abattement sur les produits de 4 600 euros pour un assuré et 9 200 euros pour un couple s'applique toujours pour les rachats qui ont lieu après 8 ans d'ancienneté. Pour un rachat effectué après le 27 septembre 2017, l'abattement s'applique en priorité :

— Aux produits correspondant aux primes versées avant le 27 septembre 2017.

— Puis aux produits correspondant aux primes versées à partir du 27 septembre qui sont imposés au taux de 7,5%.

— Enfin, aux produits correspondant aux primes versées à partir du 27 septembre qui sont imposés au taux de 12,8%.

Pour illustrer les modalités de la "Flat Tax", on peut voir ci-dessous un schéma explicatif :

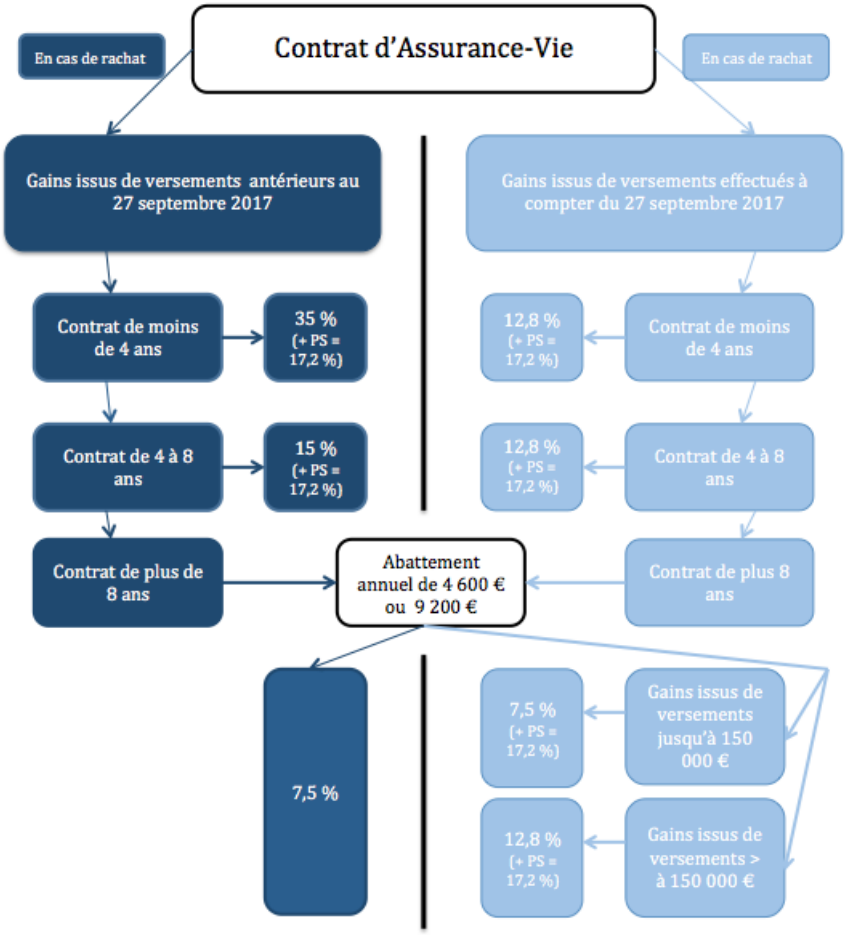

FIGURE 4.2 – Schéma explicatif de la nouvelle fiscalité.

4.2.2 La mise en place de la « Flat tax » dans une simulation de rachat

Dans cette partie nous proposons de créer « une calculatrice » sur Excel qui estime les produits nets d'impôts lors d'un rachat d'un contrat sur les fonds en euros. L'objectif étant de comparer dans les différentes situations, les gains ou pertes sur les produits générés que procure la nouvelle taxe par rapport à l'ancienne. Nous ne retirons pas les prélèvements sociaux, l'étude porte principalement sur le changement de taux d'imposition.

Les situations possibles sont :

— Les contrats souscrits après le 27 septembre 2017 avec un montant de primes versées et non rachetées inférieur à 150 000 euros.

— Les contrats souscrits après le 27 septembre 2017 avec un montant de primes versées et non rachetées supérieur à 150 000 euros.

— Les contrats souscrits avant le 27 septembre 2017 avec un montant de primes versées et non rachetées inférieur à 150 000 euros sans effectuer de versements après cette même date.

— Les contrats souscrits avant le 27 septembre 2017 avec un montant de primes versées et non rachetées inférieur à 150 000 euros avec versements après cette même date.

— Les contrats souscrits avant le 27 septembre 2017 avec un montant de primes versées et non rachetées supérieur à 150 000 euros sans effectuer de versements après cette même date.

— Les contrats souscrits avant le 27 septembre 2017 avec un montant de primes versées et non rachetées supérieur à 150 000 avec versements après cette même date.

Dans toutes les situations, nous effectuons une simulation de rachat dans les 3 compartiments fiscaux, c'est à dire pour des rachats entre 0 et 4 ans, 4 et 8 ans et 8 ans et plus. Dans la deuxième colonne se trouveront les produits nets d'impôts dans la situation où il n'y aurait pas eu la loi de finances 2018 et dans la troisième les produits net d'impôts dans le cas de la réforme. Nous rappelons le tableau permettant de calculer les produits nets de prélèvements sociaux et d'impôts :

Ancienneté contrat	Fiscalité avant 27/09	Fiscalité après 27/09
0 - 4 ans	35%	12,80%
4 - 8 ans	15%	12,80%
>= 8 ans	7,50%	7,50%

• Cas des nouveaux contrats souscrits après le 27 septembre 2017

◦ Premier cas :

Pour un encours inférieur à 150 000 euros, nous supposons que les primes ont généré 10 000 euros de produits. On obtient les résultats suivants :

Ancienneté contrat	Produit net (Fiscalité avant)	Produit net (Fiscalité après)	Gain / Perte
0 - 4 ans	6 500 €	8 720 €	34,15%
4 - 8 ans	8 500 €	8 720 €	2,59%
>= 8 ans	9 595 €	9 595 €	0%

Sur le compartiment 0 - 4 ans d'ancienneté, la nouvelle taxe génère un gain de 34,15% sur les produits. Entre 4 - 8 ans, on observe un gain faible. Pour les rachats supérieurs à 8 ans, le taux d'imposition est le même donc il n'y a pas de changement.

◦ Deuxième cas :

Pour un encours supérieur à 150 000 euros, nous supposons que 200 000 euros de primes ont généré également 10 000 euros de produits. Les résultats sont les suivants :

Ancienneté contrat	Produit net (Fiscalité avant)	Produit net (Fiscalité après)	Gain / Perte
0 - 4 ans	6 500 €	8 720 €	34,15%
4 - 8 ans	8 500 €	8 720 €	2,59%
>= 8 ans	9 595 €	9 462,5 €	-1,38%

Entre 0 et 8 ans, le pourcentage de gain est le même que pour le premier cas comme le taux d'imposition ne varie pas. Au delà de 8 ans, on remarque que les contrats avec un encours supérieur à 150 000 euros enregistrent une perte. Si l'on augmente le montant d'encours ou les plus-values, on observe une perte légèrement plus importante et inversement si l'on diminue l'encours ou les plus-values, la perte diminue.

• Cas des anciens contrats souscrits avant le 27 septembre 2017

◦ Troisième cas :

Pour un encours inférieur à 150 000 euros, nous supposons que 10 000 euros de produits ont été générés et qu'aucun versement n'est effectué après le 27 septembre. Dans ce cas présent, le taux d'imposition de la « Flat tax » n'aura pas d'impact sur les produits. La différence est seulement sur le taux de prélèvements sociaux qui sera de 17,2% pour les produits à partir de 2018.

◦ Quatrième cas :

Pour un encours inférieur à 150 000 euros, nous supposons que 10 000 euros de produits ont été générés et que des versements ont été effectués (sans dépasser le seuil de 150 000 euros toutes primes confondues) après le 27 septembre et qui ont généré 2 000 euros de produits.

Ancienneté contrat	Produit net (Fiscalité avant)	Produit net (Fiscalité après)	Gain / Perte
0 - 4 ans	7 800 €	8 244 €	5,69%
4 - 8 ans	10 200 €	10 244 €	0,43%
>= 8 ans	11 445 €	11 445 €	0%

Avec ces hypothèses, le gain pour le premier compartiment baisse considérablement par rapport aux autres cas. On observe également pour le deuxième compartiment une légère diminution des bénéfices. En faisant varier les produits issus des primes antérieures au 27 septembre, on observe une diminution du gain quand on augmente le montant de produits et inversement quand on le baisse. En augmentant les produits issus des versements ayant lieu après le 27 septembre, on observe une augmentation du gain pour le premier compartiment et une forte diminution pour le deuxième. Le taux d'imposition est le même pour le troisième compartiment donc il n'y a pas de changement.

◦ Cinquième cas :

Pour un encours supérieur à 150 000 euros, nous supposons que le montant de primes versées avant le 27 septembre est supérieur à 150 000 euros et ont généré 10 000 euros de produits et qu'aucun versement n'a lieu ensuite. Nous sommes dans la même situation que le troisième cas où la « Flat Tax » n'a pas d'impact sur les produits.

◦ Sixième cas :

Pour un encours supérieur à 150 000 euros, supposons que 100 000 euros de primes ont été versées avant le 27 septembre 2017 et ont généré 10 000 euros de produits et que 100 000 euros de primes versés après ont généré également 10 000 euros de produits. Dans cette situation, il faut prendre en compte la part des primes versées après le 27 septembre dans 150 000 euros pour bien appliquer l'abattement lors des rachats au delà de 8 ans. Cette part est ici de 50 %.

Ancienneté contrat	Produit net (Fiscalité avant)	Produit net (Fiscalité après)	Gain / Perte
0 - 4 ans	13 000 €	17 440 €	34,15%
4 - 8 ans	17 000 €	17 440 €	2,59%
>= 8 ans	18 845 €	18 580 €	-1,41%

Entre 0 et 8 ans, on retrouve la même situation que dans les trois premiers cas donc le pourcentage de gain est le même. Pour un rachat au delà de 8 ans, avec nos hypothèses de départ, on observe une perte de 1,41% des produits avec la nouvelle réforme. Néanmoins, ce pourcentage varie selon que l'on modifie les produits et le niveau d'encours. En effet, si on augmente le montant d'encours le pourcentage de gain diminue et inversement si on diminue ce montant, le gain augmente. De plus, si on augmente les produits issus des versements ayant lieu avant le 27 septembre, on observe une augmentation du gain et inversement pour les produits issus des versements après le 27 septembre.

• Analyse générale [1] :

Dans toutes ces situations on soulève 3 faits remarquables sur l'impact de la nouvelle fiscalité sur les produits générés par les contrats :

— Entre 0 et 4 ans d'ancienneté, la « Flat Tax » avantage clairement le profit des assurés ;
— Entre 4 et 8 ans, un léger gain est observé ;
— Pour une ancienneté au delà de 8 ans, il n'y a pas de changement pour les assurés qui ont un montant de primes non rachetées inférieur à 150 000 euros. Pour les assurés qui ont plus de 150 000 euros de primes versées non rachetées, la « Flat tax » génère une légère perte.

En première estimation, si l'on considère que le rachat dépend exclusivement de l'ancienneté, on peut supposer que les assurés seront plus enclins à racheter entre 0 et 8 ans et moins pour une ancienneté au delà de 8 ans. Néanmoins ce raisonnement est très simplifié et servira aux assureurs qui conservent l'ancienneté du contrat comme seul facteur explicatif dans l'attente d'avoir un historique de rachat plus long depuis la mise en place de la « Flat Tax ».

Dans la partie suivante, les différentes méthodes que nous proposons à un assureur-vie pour modéliser la loi de rachats structurels seront exposées.

1. Les six cas sont illustrés dans l'annexe 1 sous forme de graphique

Deuxième partie

Les méthodes de modélisation du comportement de rachat des assurés suite à l'entrée en vigueur de la « Flat Tax »

Dans ce mémoire, nous nous intéressons exclusivement à la modélisation du rachat structurel. La mise en place de la « Flat Tax » pourrait remettre en cause l'utilisation de la variable ancienneté comme seul facteur explicatif pour la construction de la loi de rachats structurels. Cette partie a pour vocation de proposer aux assureurs deux méthodes de modélisation de cette loi. La première est d'adapter leur loi de rachats structurels en fonction de l'ancienneté et la seconde est de modéliser le rachat structurel en utilisant une approche *machine learning* qui permet de prendre en compte d'autres facteurs explicatifs.

Chapitre 5

L'adaptation de la loi de rachats structurels en fonction de l'ancienneté

5.1 La construction de la loi de rachats structurels en fonction de l'ancienneté

Cette section vise à présenter la base de données de rachats de l'assureur anonyme que nous avons à disposition, les traitements que nous avons dû effectuer et le calcul du taux de rachat. L'objectif est de déterminer les taux de rachat par individu et de construire la loi de rachats structurels en fonction de l'ancienneté.

5.1.1 Les caractéristiques de la base de rachat

Pour notre étude, nous avons à disposition un portefeuille de contrats multi-support d'un assureur anonyme couvrant la période 2010 à 2017. Nous nous intéressons seulement aux fonds en euros. Dans la suite, nous simplifierons en parlant de contrats en euros. Le portefeuille comporte 400 535 contrats d'un même produit avec des variables caractérisant le contrat (encours de début d'année, arbitrage, rachat partiel, rachat total, versement, ancienneté, date d'entrée, date de sortie, statut du contrat) et l'assuré (date de naissance, date de décès, âge, sexe).

• Statistiques descriptives du portefeuille :

Notre portefeuille comporte 8,7% de contrats rachetés totalement et 25% rachetés partiellement au moins une fois sur la période. Les femmes sont plus nombreuses (55%) que les hommes. L'âge moyen de nos assurés est d'environ 62 ans avec un peu plus de la moitié qui ont plus de 60 ans. Pour ces derniers l'encours moyen est de 64 000 euros contre 52 000 euros pour les assurés ayant moins de 60 ans. On observe un encours maximal s'élevant à 11 millions d'euros. L'ancienneté moyenne de notre portefeuille est d'environ 4 ans.

5.1.2 Le retraitement des données

Dans cette partie, les retraitements de données que nous avons dû effectuer sur notre base initiale sont expliqués.

• Valeurs aberrantes :

— Les variables « encours », « rachat partiel » et « rachat total » contiennent des valeurs négatives. Ces valeurs pourraient servir à corriger des erreurs commises l'année précédente mais après analyse de chaque contrat concerné, on constate finalement que la présence de valeurs négatives n'est pas justifiée. On décide donc de toutes les annuler.

— Les contrats ayant des dates d'entrée supérieures aux dates de sortie sont également supprimés de la base.

• Contrats conjugaux :

Plusieurs contrats ayant les mêmes encours, les mêmes rachats partiels et totaux sont observés. Il s'agit des contrats conjugaux. Nous choisissons de garder le conjoint le moins âgé, celui qui a le plus de chance de survivre le plus longtemps donc qui a plus de temps pour racheter le contrat.

5.1.3 La détermination des taux de rachat

Le taux de rachat est l'indicateur d'intensité de la survenance d'un rachat. On détermine le taux de rachat annuel par individu ainsi :

$$\text{Taux de rachat} = \frac{\text{Montant du rachat}}{\text{Assiette du rachat}} \tag{5.1}$$

N'ayant pas d'information sur la date de rachat, nous faisons l'hypothèse que les rachats ont lieu en moyenne en milieu d'année. L'assiette de rachat dépend du type de rachat et de l'existence ou non d'arbitrage ou de versements dans l'année. Ainsi, on peut définir :

$$\text{Assiette du rachat} = \begin{cases} \text{Montant du rachat} & \text{si rachat total} \\ \dfrac{\text{Encours}_{\text{ouverture}} + \text{Encours}_{\text{cloture}} + \text{Montant du rachat}}{2} & \text{si rachat partiel} \end{cases}$$

Une fois que le taux de rachat annuel par individu est obtenu, le rachat global du portefeuille en fonction de différentes variables peut être observé. Les taux de rachat globaux, partiels et totaux annuels sur toute la période 2010-2017 sont présentés à la page suivante.

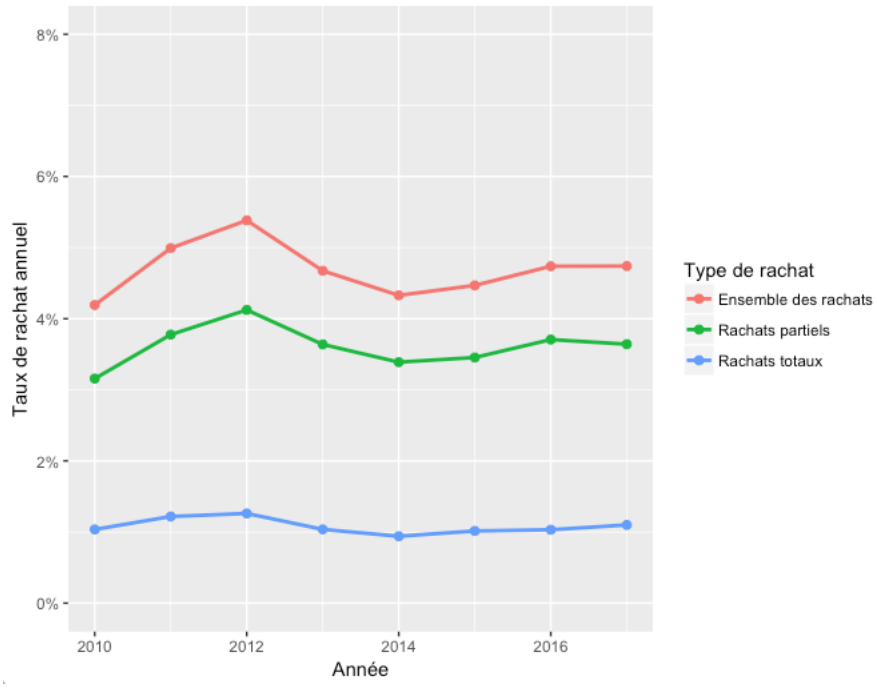

FIGURE 5.1 – Évolution du taux de rachat annuel sur la période 2010 - 2017.

Sur la période 2010-2017, le taux de rachat a connu un pic dans les années 2011 et 2012. Cette hausse des rachats peut s'expliquer par la période de crise en Grèce (il ne s'agit donc pas de rachats structurels uniquement). De plus, on observe que le taux de rachat partiel (3,1% - 4,1%) est plus élevé que le taux de rachat total (0,9% - 1,3%) et que ces deux taux évoluent de la même manière tout au long de la période.

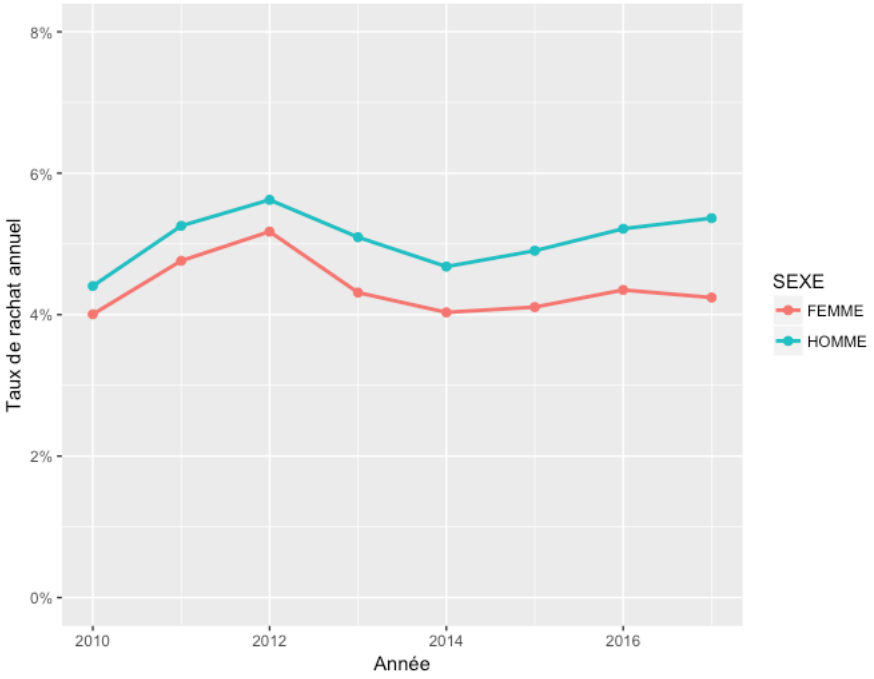

FIGURE 5.2 – Évolution du taux de rachat par sexe.

Le taux de rachat des hommes (4,4% - 5,6%) est plus élevé que celui des femmes (4% - 5,1%). De plus ces deux taux ont la même tendance sur la période hormis en 2017 où le taux de rachat chez l'homme augmente alors que celui de la femme diminue.

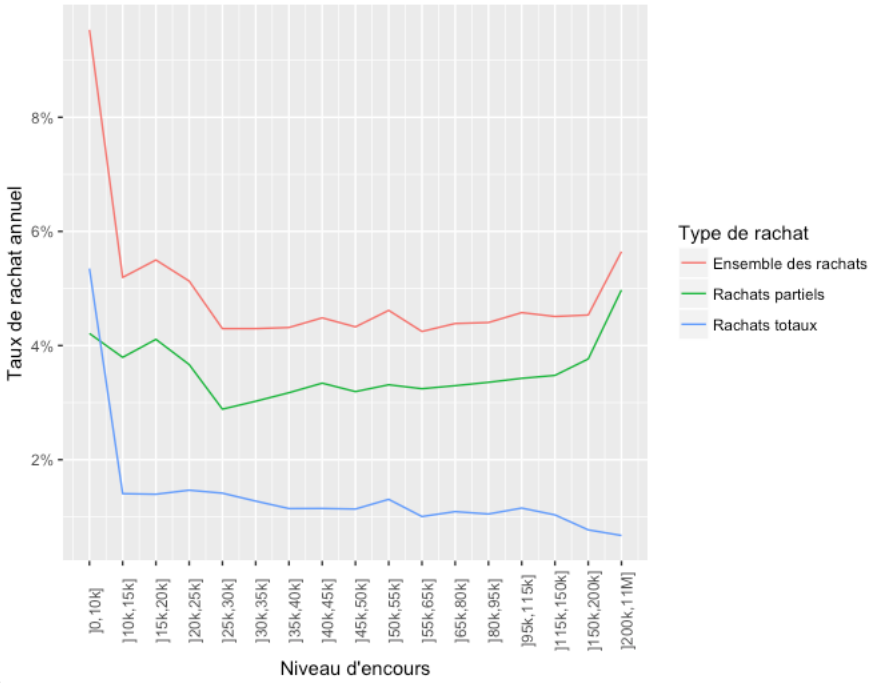

FIGURE 5.3 – Évolution du taux de rachat annuel par tranche d'encours.

Dans la construction du taux de rachat par niveaux d'encours, la variable est classée par tranche. Les niveaux d'encours sont construits de façon à avoir des classes homogènes en nombre de contrats. Sur le graphique, un point remarquable est visible sur la première classe entre 0 et 10 000 euros d'encours où le taux de rachat total est nettement supérieur aux taux de rachat total des autres classes d'encours (5,3% contre 1,1% en moyenne sur l'ensemble des classes). Cela s'explique par le fait que les assurés les moins riches rachètent plus souvent pour couvrir leurs besoins de liquidité. De plus, les personnes les plus riches effectuent plus de rachats partiels. En effet, ce type d'assuré s'intéresse davantage aux opportunités d'investissements plus rentables pour leur épargne.

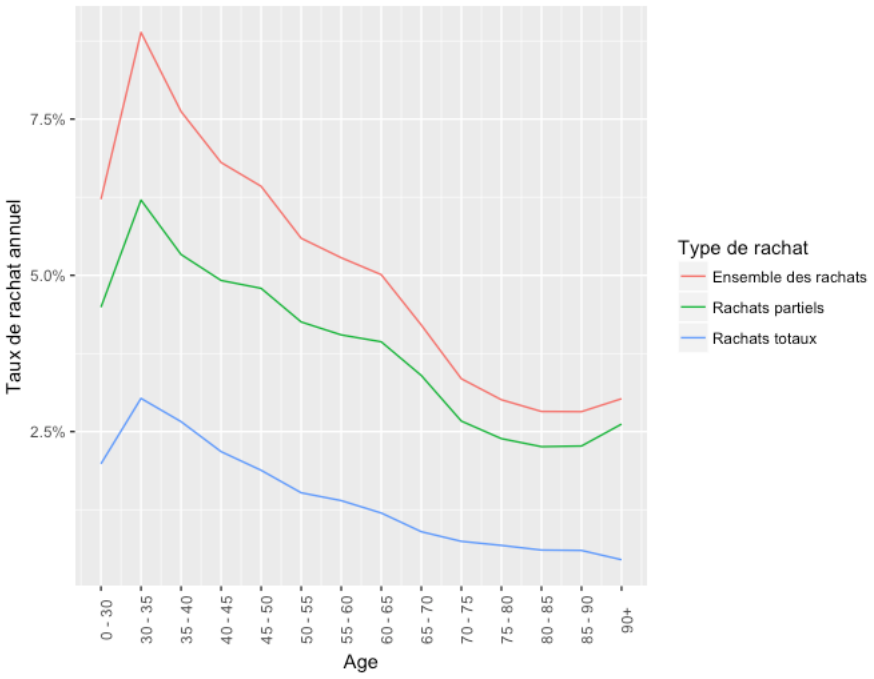

FIGURE 5.4 – Évolution du taux de rachat par âge de l'assuré.

Dans la construction du taux de rachat par âge, la variable est classée par tranche. Une première entre 0 et 30 ans et ensuite une toute les 5 ans. Le taux le plus élevé observé correspond à la classe des âges entre 30 et 35 ans. Les projets d'investissements mobiliers ou d'enfants peuvent expliquer ce pic. Une décroissance du taux de rachat est constaté à partir de 35 ans.

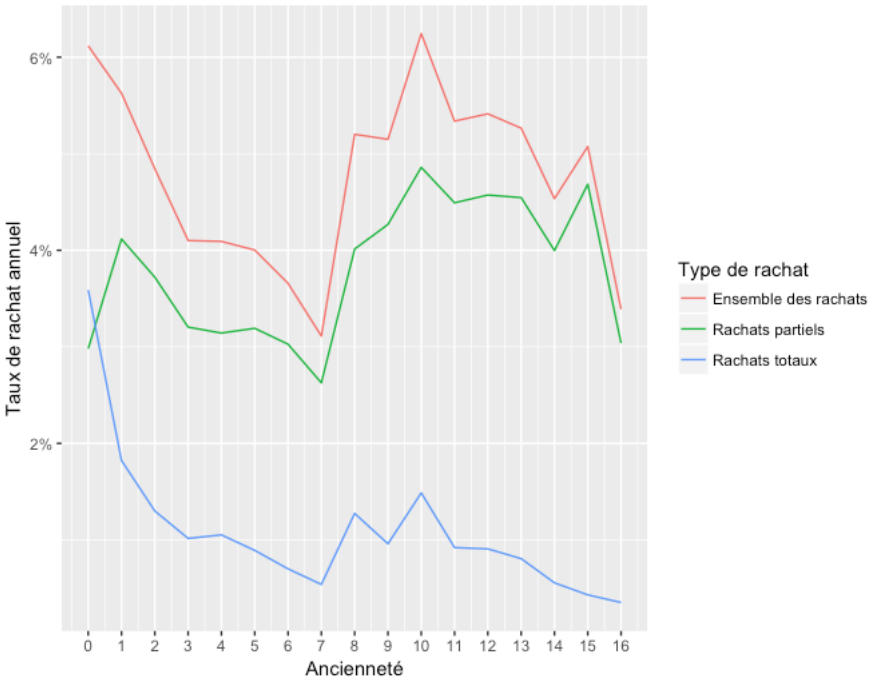

FIGURE 5.5 – Évolution du taux de rachat par ancienneté du contrat.

On observe deux points significatifs dans l'évolution du taux de rachat par ancienneté. La première année, le taux est très élevé notamment pour le rachat total. Cela peut s'expliquer par le fait que les assurés souscrivent un contrat d'assurance vie mais se rendent compte peu de temps après la souscription qu'ils ont finalement des opportunités de placements financiers plus intéressantes ailleurs (par des incitations commerciales par exemple). Le deuxième point remarquable est le pic à 8 ans. Nous pouvions l'anticiper du fait de l'important changement de fiscalité qui a lieu à 8 ans d'ancienneté du contrat (dans l'ancienne fiscalité). On peut ajouter qu'entre 3 et 4 ans, le taux ne baisse pas car les assurés profitent également du changement de fiscalité à 4 ans d'ancienneté (mais qui est moins important que celui à 8 ans).

L'évolution du taux de rachat en fonction de l'ancienneté est très intéressante pour modéliser le rachat structurel. La fonction traduisant le taux de rachat n'est pas monotone et présente des points remarquables. Cela est due à l'existence des trois compartiments fiscaux.

La mise en place de la « Flat Tax » supprime le compartiment fiscal à 4 ans et réduit (-2,2%) la différence du taux d'imposition entre le deuxième et troisième compartiment fiscal (seulement pour les contrats ayant un montant de primes versées et non rachetées inférieur à 150 000€ sinon le compartiment est supprimé pour les encours supérieur à 150 000€). Nous pouvons donc penser que la stabilisation du taux à 4 ans et le pic à 8 ans ne seront plus aussi constatable. L'utilisation de cette loi pour modéliser le rachat structurel est remis en question par le changement de fiscalité.

Dans la prochaine section, nous proposons une première méthode pour modéliser le comportement de rachat des assurés. Dans l'attente d'avoir un historique de rachats conséquent (suite à la mise en place de la « Flat Tax »), cette méthode pourrait être utiliser comme méthode transitoire.

5.2 La prise en compte de l'évolution de la fiscalité dans la loi construite

Dans cette section une adaptation de la loi « classique » du rachat structurel est proposée. Cette méthode peut être une transition en attendant d'avoir un historique assez long des rachats suite à la nouvelle réforme. Nous proposons de prendre en compte l'évolution des règles fiscales dans la loi de rachats structurels en fonction de l'ancienneté. Pour cela, nous avons recours à une fonction paramétrique présentée dans la suite.

• Détermination de la fonction paramétrique

Dans la première partie « une calculatrice » Excel a été présentée permettant de calculer les produits nets d'impôts lors de rachats dans différentes situations. Celle-ci a permis d'obtenir les gains et pertes sur les produits qu'engendrait la « Flat Tax ».

Afin d'avoir une loi de rachats structurels dépendant de l'ancienneté plus cohérente avec la nouvelle réforme, un ajustement de cette loi est proposé à l'aide de cette fonction « gain/perte ». En effet, dans les compartiments 0 - 4 ans, 4 - 8 ans et plus de 8 ans d'ancienneté, le gain ou la perte sur les produits ont été déterminés. L'ajustement de la loi dans les 3 compartiments peut être fait proportionnellement aux gains/pertes ainsi obtenus pour chaque rachat.

L'utilisation de la fonction paramétrique implique de prendre comme hypothèse forte que le rachat évolue linéairement par rapport à la variation des gains ou des pertes sur les produits. C'est à dire que si sur le compartiment 0 - 4 ans, la « Flat Tax » génère 2 fois plus de gains, il y aura alors 2 fois plus de rachats.

Par exemple, pour les nouveaux contrats qui sont souscrits après la réforme, le gain généré par la « Flat Tax » pour un rachat entre 0 et 4 ans d'ancienneté est de 34,15%. Nous pouvons considérer qu'il y aura 1,3415 fois plus de rachats aujourd'hui (dans l'hypothèse que le rachat ne dépend que de l'ancienneté fiscale du contrat).

Finalement, la fonction paramétrique dépend de la situation de rachat, c'est à dire si le contrat est souscrit avant ou après le 27 septembre 2017, du niveau d'encours et de produits générés par le contrat et le compartiment fiscal associé lors du rachat. Nous reprenons l'ordre des différentes situations de rachats de la première partie. Par exemple, la fonction paramétrique associée au deuxième cas est définie ainsi :

$$Fonction_param_cas2 = \begin{cases} 34,15\% & \text{si Ancienneté} \in [0,4[\\ 2,59\% & \text{si Ancienneté} \in [4,8[\\ -1,38\% & \text{si Ancienneté} \in [8,+\inf[\end{cases}$$

Pour chaque situation de rachats, une fonction paramétrique est obtenue. En pratique, la base ne permet pas d'appliquer correctement cette méthode car celle-ci ne renseigne pas les primes versées par les assurés. Nous disposons seulement des provisions mathématiques (primes + produits). Néanmoins la valeur du gain/perte ne varie pas beaucoup selon les différentes situations de rachats.

Pour notre étude, les fonctions paramétriques utilisées sont celles associés aux deux premiers cas.

La loi de rachats structurels dépendant de l'ancienneté adaptée est définie ainsi :

$$Loi_adaptée_ancienneté = (1 + Fonctionparam) \times Loi_ancienneté \tag{5.2}$$

Les deux lois de rachats du portefeuille sont représentées ci-dessous :

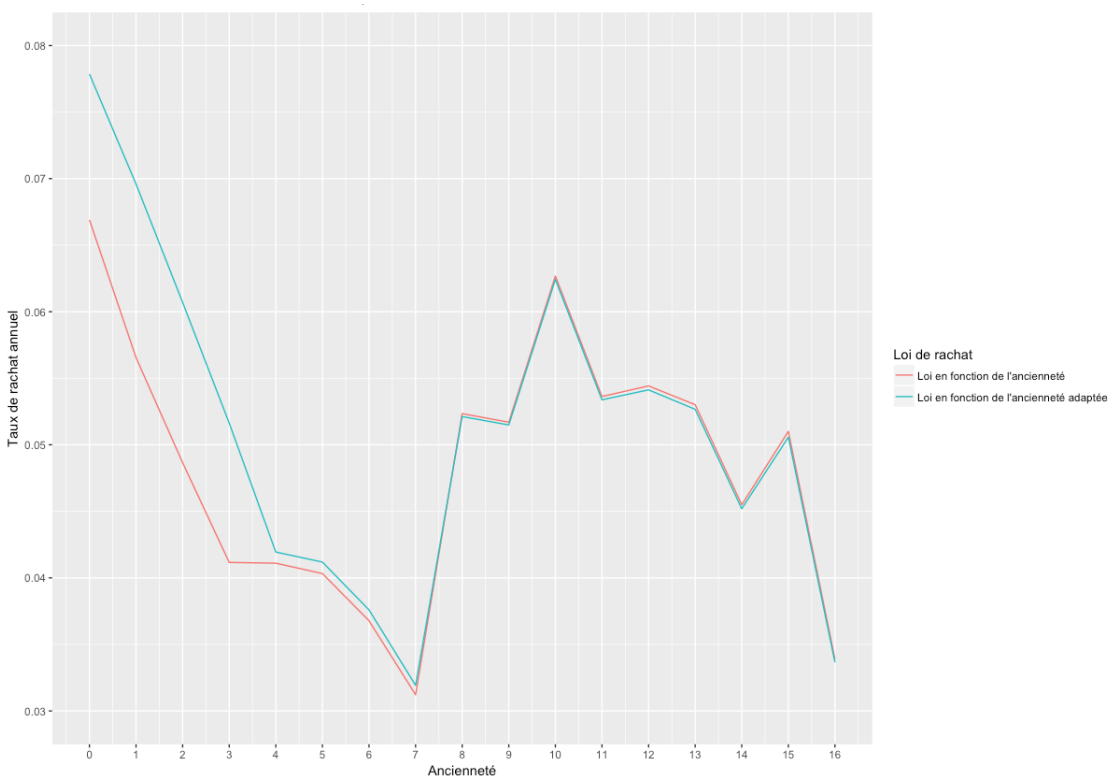

FIGURE 5.6 – Comparaison des deux lois de rachats.

Il faut bien prendre en considération que l'hypothèse d'existence d'une relation linéaire entre le rachat et l'évolution du gain ou perte des produits n'a pas été vérifiée expérimentalement (statistiquement) car nous n'avons pas encore d'historique pour le réaliser. L'objectif est de proposer une méthode avec les hypothèses connues (gain ou perte que génère la « Flat Tax » sur les produits) afin d'adapter la loi en fonction de l'ancienneté. Un assureur disposant de la décomposition de la provision mathématique (primes versées et produits) peut appliquer correctement la fonction paramétrique dans toutes les situations de rachats.

• Proposition de deux autres méthodes simples

L'adaptation de la loi de rachats en fonction de l'ancienneté peut être réalisée de différentes façons. En effet, un premier choix très simple serait de considérer un taux constant pour la loi adaptée, le taux de rachat moyen de la loi de rachats structurels en fonction de l'ancienneté. Le taux de rachat moyen de la base de l'assureur anonyme est 4,75%.

En deuxième choix, en considérant que l'ancienneté fiscale ne sera plus un facteur déterminant dans la décision de rachat des assurés, le taux de rachat moyen correspondant aux rachats ayant lieu après 8 ans d'ancienneté est utilisé quelle que soit l'ancienneté. Le taux d'imposition après 8 ans ne varie plus donc l'ancienneté fiscale n'est plus un facteur explicatif du rachat. Le taux de rachat moyen obtenu est 5,44%.

Les avantages de ces deux dernières méthodes sont qu'elles sont facilement compréhensibles et simples à utiliser. La méthode qui consiste à prendre en compte la fonction paramétrique est plus fine mais repose sur l'hypothèse que le rachat évolue linéairement avec le gain ou la perte sur les produits.

Afin de ne plus prendre en compte qu'une seule variable dans la construction de la loi de rachats structurels, nous proposons dans ce mémoire d'utiliser les techniques de *machine learning* afin de construire une loi prenant en compte d'autres facteurs explicatifs.

Chapitre 6

L'utilisation des techniques de *machine learning*

Dans ce chapitre, les principes du *machine learning* seront énoncés puis nous présenterons les modèles choisis pour construire la loi de rachats structurels.

6.1 Présentation du *machine learning*

6.1.1 Une méthode d'apprentissage des données

Arthur Samuel, l'un des fondateurs de l'intelligence artificielle, définit en 1959 le *machine learning* comme « le champ d'étude visant à donner la possibilité à une machine d'apprendre sans être formellement programmée ».

• Cas simple d'application du *machine learning* : le filtre antispam d'une messagerie

Le service de messagerie, représentant la *machine*, va examiner la manière dont l'utilisateur va classer ses mails entrant en spam ou non. Après cette période d'*apprentissage*, la machine retient des critères de classification. Si le mail contient par exemple les mots « argent » ou « rencontre facile », la probabilité que la machine classe ce mail en spam va augmenter. Au contraire, si par exemple l'expéditeur est connu et que les mots du mail sont plus courants, la probabilité va baisser.

L'accès à des données est donc primordial dans l'application du *machine learning*. Celles-ci se divisent en variables explicatives et une variable à expliquer. Les modèles vont utiliser les variables explicatives pour apprendre la variable à expliquer.

• Les types de données :

On peut caractériser deux types de données :

— Les données quantitatives qui sont des valeurs décrivant une quantité mesurable (sous la forme de nombres). Elles peuvent être continues ou discrètes ;

— Les données qualitatives qui sont des valeurs décrivant des caractéristiques, sous la forme de modalités. Elles peuvent être catégorielles ou ordinales.

La qualité des modèles de *machine learning* dépend de la qualité des données. Il est important de bien préparer la base de données avant d'utiliser les modèles. Il faut notamment éviter d'avoir des valeurs aberrantes et des variables explicatives qui ne sont pas pertinentes à l'apprentissage de la variable à expliquer.

6.1.2 Le type d'apprentissage

Les algorithmes n'ont pas tous la même fonction d'usage. Ils sont classés selon deux critères :

— Les algorithmes supervisés ou non supervisés (mode d'apprentissage) ;
— Les algorithmes de régression ou de classification (type de problème).

• Le mode d'apprentissage :

Les algorithmes supervisés prélèvent de la connaissance à partir d'une base de données contenant des couples d'entrée / sortie. Ce type d'algorithme est performant mais présente un coût dans la recherche de bases contenant des couples.
Au contraire, les algorithmes non supervisés ne contiennent pas le concept d'entrée / sortie. Ce type d'algorithme est en général moins performant.

Dans ce mémoire, l'objectif est de prédire le taux de rachat par individu en fonction de plusieurs variables explicatives. Nous nous intéresserons exclusivement aux algorithmes supervisés.

Les algorithmes supervisés cherchent à définir une représentation des associations entrée / sortie par l'intermédiaire d'une fonction de prédiction. Si l'on considère une base de données décrite par deux variables X_1 et X_2, il faudra leur associer une variable de sortie Y. L'algorithme proposera alors une fonction de prédiction f de la forme :

$$Y = f(X_1, X_2)$$

• Le type de problème à traiter :

Les algorithmes de régression et de classification se caractérisent par le type de valeurs de la variable à expliquer : donnée quantitative ou qualitative.

La variable à expliquer dans ce mémoire est le taux de rachat par individu. Nous utiliserons donc les algorithmes de régression.

6.1.3 L'évaluation d'un modèle

Avant de se lancer directement dans la partie prédiction, il est important de faire attention au phénomène récurrent dans l'application du *machine learning* : l'*overfitting* (le surapprentissage).

6.1.3.1 Le surapprentissage

La notion d'*overfitting* ou surapprentissage pour un modèle supervisé désigne le fait que le modèle prédictif produit par celui-ci s'adapte « trop bien » à la base d'apprentissage. Le modèle pourra donner de très bonnes prédictions sur les données d'apprentissage, mais il prédira mal sur des données qu'il n'a pas encore vues lors de sa phase d'apprentissage. On dit que la fonction prédictive se généralise mal et que le modèle « souffre » d'*overfitting*.

• Cas de surapprentissage dans un modèle de régression polynomiale :

Dans leur ouvrage « Data Science : fondamentaux et études de cas », E. Biernat et M. Lutz propose un exemple simple du risque de surapprentissage.

Le risque de surapprentissage dans un modèle de régression polynomiale est très fréquent lorsque l'on choisit un degré de polynôme trop grand. La fonction construite passera par un nombre important de points des données d'apprentissage et aura une forme oscillante. Une fonction oscillante ne permettra pas de généraliser suffisamment le modèle et donc pourra entraîner une mauvaise prédiction si les données testées varient des données de la base d'apprentissage.

Pour illustrer, une simulation sur le logiciel R est réalisée sur un jeu de données trivial de 10 points (y en fonction de x).
Un modèle est ajusté pour expliquer y en fonction de x selon 3 formes de modèles différentes :

— Une régression linéaire simple ;

— Une régression polynomiale d'ordre 3 ;

— Une régression polynomiale d'ordre 9.

Les 3 modèles sont ensuite utilisés pour réaliser des prédictions de y à partir d'un vecteur contenant 250 points répartis à intervalles réguliers entre 1 et 10.
Les résultats sont représentés sur la figure ci-dessous :

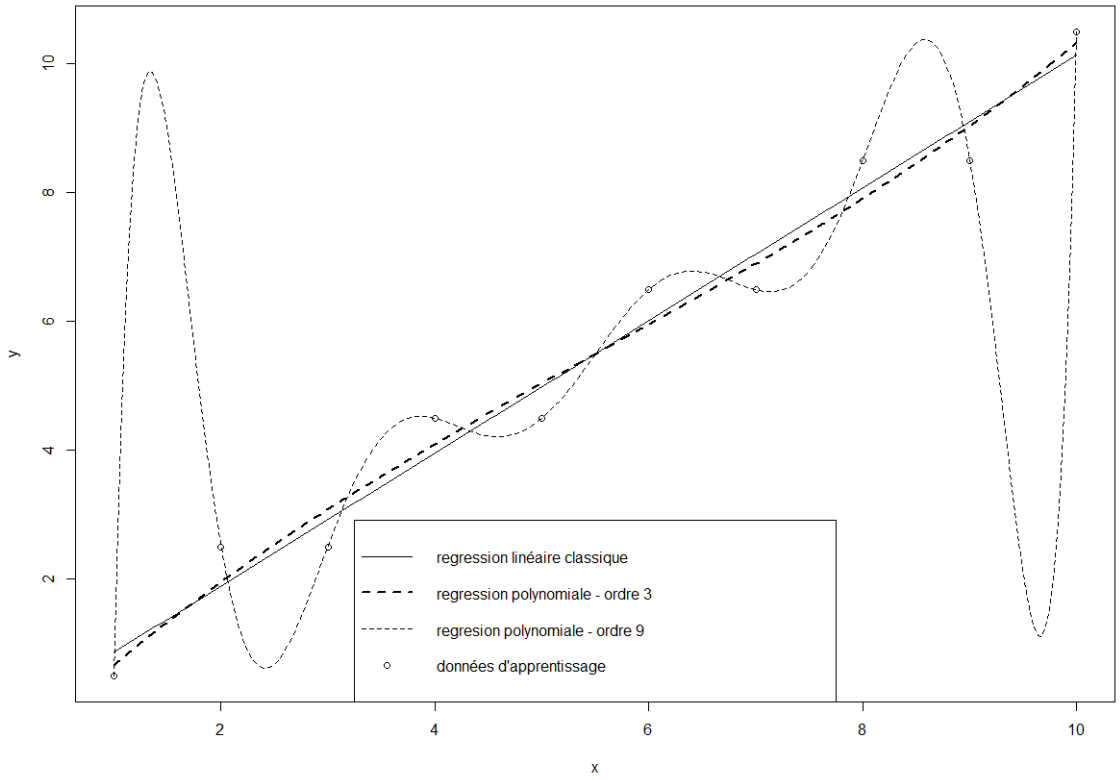

FIGURE 6.1 – La notion de surapprentissage : résultat de la simulation.

La régression polynomiale d'ordre 9 est celle qui ajuste le mieux les données d'apprentissage. Toutefois, nous pouvons remarquer que si les valeurs de x changent, il y aura de grands écarts. Au contraire, les régressions linéaires et polynomiales d'ordre 3 ne passent pas par les données d'apprentissage mais fournissent une représentation moyenne plutôt fiable. Le modèle de régression polynomiale d'ordre 9 n'est pas convaincant car si les données d'apprentissage varient cela va entraîner des ajustements très différents dans la régression. Ce modèle caractérise très bien le phénomène de surapprentissage.

En pratique, il peut être compliqué de distinguer un bon modèle d'un modèle surappris. Le surapprentissage est caractérisé par un modèle qui a une erreur de prédiction sur la base d'apprentissage très faible (biais) mais se généralise mal (variance). Généralement quand on diminue le biais, la variance augmente et inversement lorsqu'on diminue la variance on augmente le biais. Il faut chercher un compromis biais - variance.

6.1.3.2 Le compromis biais - variance

Lorsqu'un modèle cherche à prédire une variable, il commet deux types d'erreurs qui sont le biais et la variance. Le biais caractérise l'erreur de prédiction (complexité du modèle) et la variance est liée à la capacité du modèle à se généraliser. Le compromis va permettre de trouver un équilibre entre ces deux erreurs.
Dans l'exemple présenté ci-dessus, nous étions dans le cas d'un modèle ayant un biais très faible mais une forte variance.

Une stratégie classique peut être adoptée en divisant la base de données en deux. Une

partie sera utilisée pour effectuer l'apprentissage et la prédiction sera effectuée sur la seconde. Généralement 70% de la base de données est réparti dans la base d'apprentissage et 30% dans la base de test (nous détaillerons la stratégie de division de données dans le chapitre 9).
Nous pouvons mesurer deux types d'erreurs : l'erreur de modélisation et l'erreur de prévision.

L'erreur de modélisation s'évalue sur la base d'apprentissage alors que l'erreur de prévision se détermine sur la base de test. En reprenant l'exemple du modèle de régression polynomiale, nous pouvons calculer ces erreurs avec différents degrés de complexité, en augmentant les degrés du polynômes. Comme nous l'avons vu ci-avant, l'erreur de modélisation est plus faible pour une régression polynomiale d'ordre 9 que d'ordre 3. L'erreur de prévision se comporte autrement, elle va diminuer en augmentant la complexité jusqu'à un certain seuil de complexité. Au delà de ce seuil, l'erreur de prévision augmente à nouveau.

La problématique de surapprentissage est résolue lorsque le seuil optimal qui réduit au maximum ces deux erreurs est estimé. Un choix de la métrique de performance doit être fait pour évaluer la qualité du modèle.

6.1.3.3 Le choix de la métrique de performance

Les mesures à disposition pour évaluer la qualité d'un modèle de régression sont nombreuses. Elles se basent sur trois grandeurs :

— La valeur observée y_i ;
— La valeur prédite pour y_i appelée $\widehat{y_i}$;
— La moyenne de la valeur observée $\bar{y}$.

• L'erreur moyenne absolue (MAE, *Mean Absolute Error*) :

$$\frac{1}{n}\sum_{i=1}^{n}|y_i - \widehat{y_i}|$$

• La racine carrée de la moyenne du carré des erreurs (RMSE, *Root Mean Squared Error* :

$$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y_i})^2}$$

La RMSE permet, par rapport à la MAE, de donner un poids relativement élevé aux erreurs importantes. Ces deux indicateurs ont l'avantage d'être dans l'unité de la variable à expliquer ce qui permet d'être facilement interprétable.

Dans ce mémoire, la métrique RMSE sera utilisée pour l'évaluation de la qualité des modèles de prédiction.

6.1.3.4 Préparation à la programmation

L'application du *machine learning* s'effectuera à l'aide du logiciel R. De nombreuses fonctions sont présentes dans ses librairies ce qui permet de mettre en oeuvre tout le processus de la préparation des données à la comparaison des modèles. En pratique, nous utiliserons la librairie H2O.

• La librairie H2O :

H2O est une librairie *open source* distribuée par *H2O.ai* (« programme informatique dont le code source est distribué sous une licence permettant à quiconque de lire, modifier ou redistribuer ce logiciel »). Le logiciel R peut être rapidement limité pour l'étude de grandes bases de données. H2O a la capacité de paralléliser les opérations et d'utiliser ses serveurs pour garder en mémoire les modèles de *machine learning* construits. La librairie offre une puissance de calcul nettement supérieure aux librairies existantes sur R.

Les fonctions comprises dans le *package* H2O donnent un intérêt supplémentaire à son utilisation. Il propose un large choix de modèles de prédiction permettant de ne pas se restreindre sur le choix des modèles à utiliser.

Pour éviter le surapprentissage des modèles, H2O dispose de la fonction *h2o.grid* qui permet de gérer le problème d'erreur de prédiction et de généralisation des modèles. Elle sera explicitée plus en détails dans la partie 9.1.2.

6.1.4 Le choix des méthodes d'apprentissage

Nous proposons dans cette deuxième méthode d'apprendre le taux de rachat par individu en fonction de plusieurs variables caractérisant le contrat et l'assuré. Le lien existant entre ces facteurs peut s'avérer complexe. Pour cela, nous proposons deux approches.

• La première approche est d'utiliser les méthodes ensemblistes. Le principe est de combiner des modèles avec des performances faibles pour obtenir un modèle plus efficace. Parmi cette grande famille, il existe deux sous-familles :

— Les méthodes parallèles ;

— Les méthodes séquentielles.

Les méthodes parallèles visent à entraîner des modèles de manière indépendante pour ensuite les regrouper afin de prendre une décision, ce qui permet de réduire la variance. Le modèle *Random Forest* repose sur ce principe.

Les méthodes séquentielles visent à entraîner des modèles en série. Chaque modèle pourra connaître l'erreur du modèle précédent. Ainsi, ayant connaissance de l'erreur passée, il pourra se concentrer sur la correction des erreurs connues. Yoav Freund et Robert Schapire réalisèrent les premiers travaux en 1990 reposant sur cette idée. Les méthodes séquentielles sont également appelées les méthodes de *boosting*. Le modèle *Gradient Boosting Machine* repose sur ce principe. Une variante de ce modèle existe et a souvent été utilisée dans les compétitions *Kaggle*[1] : *XGBoost*.

1. Kaggle est une plateforme web organisant des compétitions en science des données.

Les modèles proposés seront un modèle de *Random Forest* et deux modèles de *Boosting* : le *Gradient Boosting Machine* et *XGBoost*. Ces trois modèles agrègent des arbres de décision.

• La seconde approche est d'utiliser les méthodes connexionnistes : les réseaux de neurones. Ce modèle complexe permet de modéliser de fortes interactions non linéaires entre la variable à expliquer et les variables explicatives. Depuis peu, une méthode d'apprentissage basée sur les réseaux de neurones est apparue : l'apprentissage profond (le *Deeplearning*).

Après avoir expliqué le choix des méthodes d'apprentissage qui seront utilisées, nous présentons plus en détail les modèles.

6.2 Présentation des modèles de prédiction

Dans cette section, les 3 méthodes ensemblistes seront présentées puis le modèle connexionniste sera décrit. Nous introduirons tout d'abord les grands principes des arbres de décision qui sont un prérequis à la compréhension des *Random Forest*.

6.2.1 Les arbres de décision

Un arbre de décision a pour objectif d'expliquer une valeur à partir d'une série de variables discrètes ou continues. On parle d'arbre de régression lorsque la variable à expliquer est quantitative et d'arbre de classification pour une variable qualitative.
Le principe de ces méthodes est de grouper les individus de manière homogène du point de vue de la variable à prédire en prenant en compte la qualité de prédiction des variables explicatives. La hiérarchie formée permet de représenter les résultats dans un arbre et de définir des règles explicatives.

Les règles de construction d'un arbre sont très simples. Le modèle effectue plusieurs itérations et à chacune d'elle, divise les individus en k classes (généralement $k = 2$) pour expliquer la variable que nous voulons prédire. La première division est obtenue en choisissant la variable explicative qui effectuera la meilleure séparation des individus. Les sous-populations créées sont représentées par les noeuds de l'arbre. Une mesure de proportion est associée à chaque noeud. Elle permet d'expliquer l'appartenance à une classe d'une variable de sortie. L'opération de division est effectuée pour chaque sous-population jusqu'à ce que plus aucune séparation ne soit possible. Lorsque le partitionnement est fini, les derniers noeuds deviennent une feuille à laquelle est attribuée une valeur quantitative (pour une variable quantitative comme variable à expliquer).

Exemple : On peut représenter sous forme d'arbre de décision le fait d'accepter ou non un prêt.

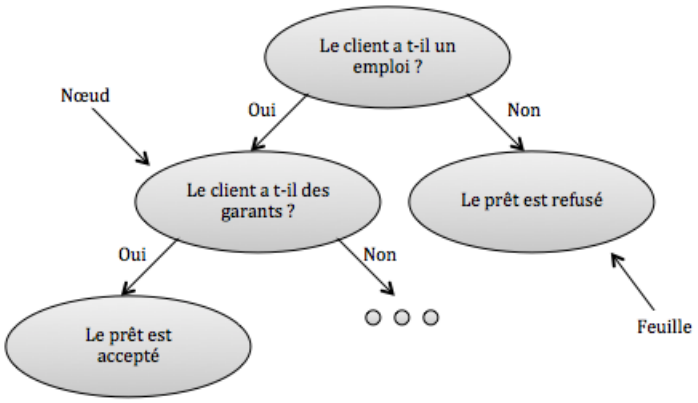

FIGURE 6.2 – Exemple d'arbre de décision.

L'avantage de cette méthode est sa facilité de compréhension et d'interprétation des résultats. L'inconvénient est qu'un arbre de décision peut être sujet aux problématiques de surapprentissage. Pour éviter ce problème, il est possible d'élaguer l'arbre (choisir un nombre de feuilles maximal) ou de choisir un critère d'arrêt.

L'algorithme de la construction d'un arbre de décision implique d'avoir :

— Un critère de division ;
— Une règle d'arrêt ;
— Une affectation des feuilles à une valeur de la variable à expliquer.

• Le critère de division

Une division d'individus est considérée comme admissible si les « noeuds fils » obtenus ne sont pas vides. Le critère de division repose sur la définition d'une fonction d'hétérogénéité. Pour un arbre de régression, l'hétérogénéité du noeud k peut être définie par la variance :

$$V_k = \frac{1}{\#k} \cdot \sum_{i \in k} (Y_i - \bar{Y_k})^2$$

avec #k le cardinal du noeud k, Y la variable à expliquer et $\bar{Y_k} = \frac{1}{\#k} \cdot \sum_{i \in k} Y_i$.
L'objectif est de minimiser cette variance intra-classe de la variable à expliquer.

• La règle d'arrêt :

L'objectif de définir une règle d'arrêt est de choisir un niveau de noeuds optimal qui permet d'avoir le juste équilibre entre un arbre trivial et un arbre surappris (compromis biais/variance).

• L'affectation :

A chaque feuille est affectée une valeur de la variable à expliquer. Dans le cadre d'un arbre de régression, la valeur associée est la moyenne des observations associées à la feuille.

Nous allons désormais présenter les trois modèles ensemblistes qui agrègent des arbres et qui permettent de combler l'aptitude au surapprentissage des arbres de décision.

6.2.2 Les *Random Forest*

Les *Random Forest* ont été proposés par Leo Breiman en 2001. L'idée principale de cette méthode est d'utiliser plusieurs arbres de décision construits en parallèle pour pallier aux défauts d'un arbre seul. L'algorithme des forêts aléatoires repose sur les techniques de *bagging* et l'ajout d'une composante aléatoire. L'assemblage d'arbres de décision construits sur la base d'un tirage aléatoire des individus se nomme le *tree bagging*. A cette technique est ajouté un échantillonage sur les variables explicatives : le *feature sampling*.
On peut écrire :

$$Randomforest = Tree_bagging + Feature_sampling$$

• Le *Tree bagging* :

Le *Bagging* ou le *boostrap aggregation* est une méthode d'agrégation de modèles qui permet de réduire la variance. Il a été introduit par Leo Breiman en 1996. L'algorithme du *tree bagging* pour construire n arbres de décisions est le suivant :

— Pour un couple d'entrée/sortie (X,Y), on tire aléatoirement avec remise n échantillons de (X,Y) que l'on note (X_n, Y_n) ;
— Les arbres de décision sont entraînés ensuite sur chaque couple (X_n, Y_n).

On effectue ensuite une prédiction en effectuant la moyenne des valeurs prédites par les n arbres.

• Le *Feature sampling* :

Le *Feature sampling* propose en plus du tirage aléatoire sur les lignes effectué par les méthodes de *bagging*, un tirage aléatoire sur les variables explicatives. L'objectif est de réduire la variance de l'ensemble créé. La variance de l'ensemble se calcule ainsi :

$$V_{modele} = \rho \cdot \sigma^2 + \frac{1-\rho}{n} \cdot \sigma^2$$

Avec :

— ρ le coefficient de corrélation des paires de variables ;
— σ^2 la variance des variables explicatives.

Si on augmente le nombre d'arbres, le deuxième terme s'efface. Le *Feature sampling* aura comme rôle d'effacer le premier terme en baissant le coefficient de corrélation entre les variables.

Le *Tree bagging* et le *Feature sampling* visent à diminuer la variance du modèle.

• Les paramètres du *Random forest* en pratique :

Le modèle *Random Forest* sera implémenté sous R avec la librairie H2O. La fonction utilisée est *h2o.randomForest*. Nous présenterons les trois paramètres primordiaux dans le cadre d'une régression :

— ntrees : le nombre d'arbres différents à entraîner ;

— max_depth : la profondeur maximale des arbres qui dépend du niveau d'interaction entre les variables ;

— mtries : le nombre maximal de variables qui est tiré aléatoirement, le critère par défaut est p / 3 (p le nombre de variables).

Nous choisissons la métrique RMSE comme critère de séparation afin de couper les feuilles de chaque arbre en cours de construction.

Le modèle *Random Forest* étant présenté, nous passons désormais à la présentation des méthodes de *Boosting*.

6.2.3 Le *Gradient Boosting Machine*

Le *Gradient Boosting Machine* a été proposé par Jérome Freidman en 2002. Ce modèle repose sur les mêmes principes que les méthodes de *boosting* : « construire une séquence de modèles de sorte que chaque étape, chaque modèle ajouté à la combinaison, apparaisse comme un pas vers une meilleure solution ». La nouveauté dans le *Gradient Boosting Machine* est que le pas est franchi vers le gradient de la fonction de perte (pour mesurer l'ajustement).
Le *Gradient Boosting Machine*, comme le *Random Forest*, agrège des arbres de décision.

On peut écrire que :

$$GradientBoosting = Descente_de_Gradient + Boosting$$

Dans le cas d'une régression, l'algorithme du modèle peut se schématiser ainsi :

❑ La fonction de coût est choisie afin d'estimer l'erreur entre les valeurs prédites par le modèle et la variable à expliquer. On peut la définir comme suit :

$$G(x, y) = \frac{1}{2}(x - y)^2$$

❑ Le modèle est initialisé par :

$$\gamma_0 = \underset{\gamma \in \mathbb{R}}{argmin}(\sum_{i=1}^{n} G(y_i, \gamma))$$

Avec y_i la i^e composante de la variable à expliquer et n le nombre d'observations.

❑ Pour b allant de 1 à B (nombre d'arbres) :

— On note f_b le modèle obtenu à l'étape b ;

— On calcule le résidu y - $f_{b-1}(x)$ qui s'apparente au gradient négatif de la fonction de coût :

$$y_i - f_b(x_i) = -(\frac{dG(y_i, f_{b-1}(x_i))}{df_{b-1}(x_i)})$$

— Un nouvel arbre M_b est créé afin d'estimer y_i - $f_{b-1}(x_i)$;

— On résout le problème d'optimisation :

$$\gamma_m = \underset{\gamma \in \mathbb{R}}{argmin}(\sum_{i=1}^{n} G(y_i, f_{b-1}(x_i) + \gamma \cdot M_b(x_i)))$$

Le nouveau modèle f_b s'écrit :

$$f_b(x) = f_{b-1}(x) + \gamma \cdot M_b(x)$$

❑ Finalement, l'algorithme renvoie le modèle :

$$F(x) = \gamma_0 + \gamma \cdot M(x) + ... + M_B(x)$$

• Les paramètres du *Gradient Boosting Machine* en pratique :

La fonction utilisée est *h2o.gbm*. Lors de la phase de calibrage, nous optimiserons ces paramètres :

— ntrees : le nombre d'arbres différents à entraîner ;

— max_depth : la profondeur maximale des arbres qui dépend du niveau d'interaction entre les variables ;

— learn_rate : définit le pas à chaque descente de gradient.

La fonction de coût qui sera utilisée est la fonction des moindres carrés.

Nous passons désormais à la présentation d'une variante du *Gradient Boosting Machine* : le modèle *XGBoost*.

6.2.4 *XGBoost*

En 2016, Chen et Guestrin ont proposé une variante du modèle *Gradient Boosting Machine*. Le modèle *XGBoost* est la version « extrême » du *Gradient Boosting Machine*. Par son implémentation en parallèle, le temps d'entraînement du modèle *XGBoost* est beaucoup moins long que le *Gradient Boosting Machine*. De plus, il peut utiliser d'autres algorithmes sous-jacents que les arbres de décision.

Nous considérons dans cette présentation que ce modèle utilise les arbres de décision et nous notons sa fonction de perte *l*. Comme pour le *Gradient Boosting Machine*, *XGBoost* construit des arbres de décision en série et cherche à réduire le biais et la variance.

Pour rappel le modèle de *tree boosting* s'écrit :

$$\widehat{y_i} = \sum_{k=1}^{K} f_k(x_i), \forall i \in [1, n]\,,\ f_k \in F \text{ et (X,Y) le couple d'entrée.}$$

Avec F l'ensemble des arbres de décision possibles et K le nombre d'arbres. Afin de trouver les meilleurs paramètres, nous avons besoin de définir une fonction objectif.

• La fonction objectif

La fonction objectif permet de mesurer la performance du modèle en fonction d'un ensemble de paramètres. Elle est définie comme suit :

$$Obj(\theta) = \sum_{i=1}^{n} l(\widehat{y}_i, y_i) + \sum_{k=1}^{K} \beta(f_k)$$

Avec $\beta(f_k)$ la régularisation associée à chaque arbre. La fonction de perte mesure l'erreur entre la prédiction du modèle et la valeur de référence. Généralement, on utilise comme fonction de perte la fonction d'erreur quadratique moyenne. Le terme de régularisation a pour rôle de contrôler la complexité du modèle et éviter le surapprentissage.

Cette fonction objectif utilise des fonctions comme paramètres. Afin de les optimiser, nous utilisons la méthode *Additive training*. Le principe de cette approche est d'apprendre à chaque itération l'erreur de l'itération précédente. Le modèle apprend l'erreur commise à l'itération précédente et ajoute un arbre qui permet de réduire l'erreur de prédiction. Ainsi, la valeur prédite au pas t (t $\in$ [0,K]) est :

$$\widehat{y}_i^{(t)} = \sum_{k=1}^{t} f_k(x_i) = \widehat{y}_i^{(t-1)} + f_t(x_i)$$

Le but est de minimiser la fonction objectif. A chaque étape, l'arbre qui minimise la fonction objectif est choisi :

$$\begin{aligned} Obj(t) &= \sum_{i=1}^{n} l(\widehat{y}_i^{(t)}, y_i) + \sum_{k=1}^{K} \beta(f_k) \\ &= \sum_{i=1}^{n} l(\widehat{y}_i^{(t-1)}, y_i) + f_t(x_i) + \beta(f_t) \end{aligned}$$

Pour l'optimisation, nous devons d'abord utiliser une extension de la formule de Taylor. Le résultat est :

$$Obj(t) \approx \sum_{i=1}^{n} (g_i \cdot f_t(x_i) + \frac{1}{2} h_i \cdot f_t^2(x_i)) + \beta(f_t)$$

Avec :

$$g_i = \partial_{\widehat{y}_i^{(t-1)}} \cdot l(y_i, \widehat{y}_i^{(t-1)})$$

$$h_i = \partial^2_{\widehat{y}_i^{(t-1)}} \cdot l(y_i, \widehat{y}_i^{(t-1)})$$

g_i et h_i représentent respectivement les dérivées partielles d'ordre 1 et 2 de la fonction de perte.

Nous devons désormais trouver le facteur de régularisation qui « pénalisera » les modèles trop complexes. On pose $f_t(x_i) = c_{q(x)}$ avec q : $\mathbb{R} \to$ 1,2,...,T. Dans cet algorithme, le nombre de feuilles est pénalisé :

$$\beta(f_t) = \gamma \cdot T + \frac{1}{2}\lambda \sum_{j=1}^{T} c_j^2$$

Les paramètres λ et γ permettent d'ajuster le terme de régularisation. A chaque étape, la grandeur de l'ajustement diminue. Le facteur de retrait « pénalise » l'ajout d'un autre modèle pour ralentir l'apprentissage et éviter le surapprentissage.

• Les paramètres du modèle *XGBoost* en pratique :

La fonction utilisée est *h2o.xgboost*. Lors de la phase de calibrage, nous optimiserons ces paramètres :

— ntrees : le nombre d'arbres différents à entraîner ;

— max_depth : la profondeur maximale des arbres qui dépend du niveau d'interaction entre les variables ;

— learn_rate : définit le pas à chaque descente de gradient ;

— col_sample_rate : le taux d'échantillonnage des colonnes pour chaque division de chaque niveau ;

— sample_rate : le ratio d'échantillonnage des observations.

Les trois modèles ensemblistes ont été présentés, nous passons désormais à la présentation d'un autre modèle de prédiction : le *Deeplearning*.

6.2.5 Le *Deeplearning*

Le *Deeplearning*, ou apprentissage profond est une forme d'intelligence artificielle. Les algorithmes de *Deeplearning* sont fondés sur ceux des réseaux de neurones à plusieurs couches cachées.

Les premiers travaux de la modélisation neuronale datent des années 1940. Mac Culloch et Pitts (1943) ont réalisé les premières constructions par « neuromimétisme de l'unité de base du calcul neuronal ».

Le Perceptron de Rosenblatt est le premier neurone artificiel fonctionnel (1962). Le Perceptron désigne un ensemble de neurones formels connectés. Rosenblatt cherchait à modéliser la reconnaissance visuelle des images. Pour cela il s'est inspiré de la structure biologique de la vision. Tout d'abord, nous présentons le Perceptron simple.

6.2.5.1 Le Perceptron simple : un séparateur linéaire

Un Perceptron simple P_w peut être défini par :

— Un espace d'entrée X ;

— Un espace de sortie Y ;

— Un vecteur de poids w = $(w_1, ..., w_{n+1}) \in \mathbb{R}^{n+1}$;

— Une fonction de réponse ϕ(u) = $1_{u \geq 0}$.

$$P_w(x) = \phi(w \cdot x) = \phi(\sum_{i=1}^{n+1} w_i \cdot x_i) \; \forall x \in \mathbb{R}^n$$

Il existe plusieurs fonctions de réponses. Voici quelques exemples :

— $\phi(\mathrm{u}) = \mathrm{u}$ (neurone linéaire) ;

— $\phi_T(\mathrm{u}) = \dfrac{1}{1 + exp(-\dfrac{u}{T})}$ (sigmoïde) ;

— $\phi_T(\mathrm{u})$ comme variable de Bernoulli à valeurs dans [0,1] telle que :

$$P[\phi_T(u) = 1] = \frac{1}{1 + exp(-\frac{u}{T})}$$

Le Perceptron simple peut être représenté schématiquement de la manière suivante :

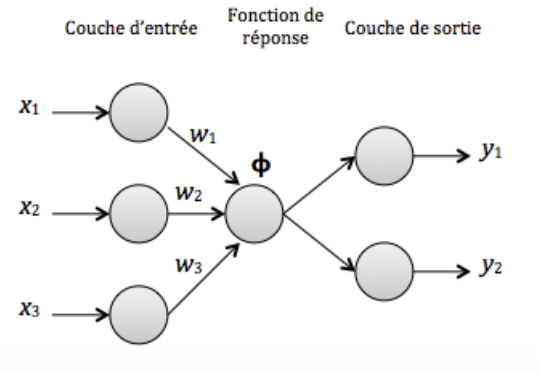

FIGURE 6.3 – Perceptron simple.

Le Perceptron défini par Rosenblatt est défini ainsi :

$$P_w(x) = \text{signe}(\sum_{i=1}^{n} w_i \cdot x_i)$$

Un perceptron peut s'interpréter comme une droite linéaire dans le cas d'une dimension d'ordre 2 et d'une sortie qui prend comme valeur 1 ou -1. Dans ce cas, résoudre un problème statistique revient à déterminer la droite linéaire qui permet de séparer l'ensemble des points en deux sous ensembles.

En analogie avec l'actuariat, la régression linéaire multiple (GLM) peut s'apparenter à un neurone formel où la fonction de réponse ϕ est la fonction inverse de la fonction lien g :

$$g(\mathbb{E}[Y|X]) = {}^t w \cdot X$$

Avec (X,Y) le couple d'entrée/sortie et w le vecteur des poids.

Une différence majeure existe entre le GLM et le modèle neuronal. Ce dernier n'introduit pas d'hypothèse de distribution sur Y|X. Le modèle linéaire se révèle souvent insuffisant pour modéliser des phénomènes complexes. Par exemple, il peut arriver que des espaces ne soient pas linéairement séparables. Dans ce cas, il est préférable d'utiliser des méthodes plus complexes : les réseaux de neurones multicouches.

6.2.5.2 Le Perceptron multicouches

Minsky et Papert ont constaté en 1969 qu'un réseau multicouches peut dépasser les limites du Perceptron simple. Pour que les méthodes connexionnistes soient utilisées, il faut attendre les travaux de Le Cun (1985), Mac Culland et Rumelhart (1986) avec l'algorithme d'apprentissage de *rétropropagation du gradient*. Cet algorithme est une généralisation d'un algorithme proposé par Widrow-Hoff (1960) et de la méthode du gradient.

A la différence du Perceptron simple, le Perceptron multicouches contient une ou plusieurs couches cachées. Le Perceptron multicouches peut être représenté schématiquement :

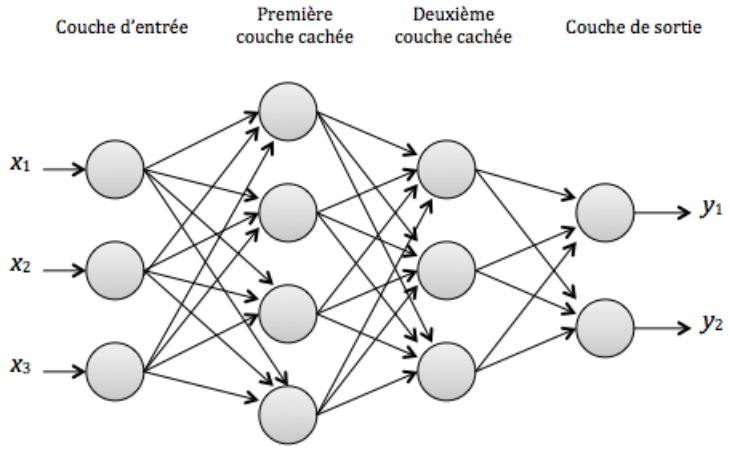

FIGURE 6.4 – Perceptron multicouches.

L'algorithme de *rétropropagation du gradient* sera justifié mathématiquement par les résultats de Hornik (1989, 1993), Cybenko (1989), Barron (1993) qui montrent par exemple que « toute fonction de $\mathbb{R}^l$ dans $\mathbb{R}^n$ ayant un nombre fini de discontinuités peut être approchée avec une précision arbitraire par certains réseaux à une couche cachée ».
Pour simplifier les notations, nous nous sommes limités par la suite à la définition d'un réseau à une couche cachée.

Théorème de Hornik(1993) :

Soit I = $[\theta^-, \theta^+] \subset \mathbb{R}$, $\theta^- \neq \theta^+$. Soit $\mathbb{U}$ = U un voisinage de 0 dans $\mathbb{R}^l$, et $\phi : \mathbb{R} \to \mathbb{R}$ borélienne bornée, non polynomiale sur I. Soit $P(\phi, I, U)$ l'ensemble des Perceptrons à une couche cachée de $\mathbb{R}^l$ dans $\mathbb{R}$, de la forme :

$$P_{\theta,W^1,W^2}(x) = \sum_{j=1}^{m} W_j^2 \cdot \phi(\sum_{i=1}^{l} W_{i,j}^1 x_i - \theta_j),$$

où m $\in N^*$, $W^2 = (W_j^2) \in \mathbb{R}^m$, $W^1 = (W_{i,j}^1) \in \mathbb{U}^{m \times l}$, $\Theta = (\theta_j) \in I^m$. Soit K un ensemble compact de $\mathbb{R}^l$. Alors :

i) Pour toute mesure finie μ sur K, $P(\phi, I, U)$ est dense dans tous les $L^p(\mu)$, p $\in$ [1,∞[: pour toute fonction f $\in L^p(\mu)$, il existe une suite (f_k) dans $P(\phi, I, U)$ telle que $\forall\ \epsilon > 0$, $\exists$ N $\in \mathbb{N}$, $\forall$ k $\leq$ N,

$$||f - f_k||_p = (\int_K |f - f_k|^p d\mu)^{1/p} < \epsilon$$

ii) Dès que ϕ est continue, $P(\phi, I, U)$ est dense dans (C(K),$||.||_\infty$).

Il s'agit d'un théorème d'existence. Ce résultat théorique est important mais ne donne pas d'information sur la taille de la couche cachée m qu'il faut choisir pour approximer une fonction donnée avec une précision voulue.

L'algorithme de rétropropagation du gradient

Nous allons présenter l'algorithme utilisé afin de déterminer les paramètres optimaux qui réduisent l'erreur entre les sorties fournies par le modèle et les sorties « désirées ». La rétropropagation vise à corriger cette erreur en modifiant les poids du réseau de neurones (initialement choisis de manière aléatoire). En appliquant cette étape plusieurs fois, le réseau offre une meilleure prédiction par la diminution de l'erreur.

On considère un ensemble de données qui vont servir à l'apprentissage. Chaque échantillon possède ses valeurs cibles qui sont celles que le réseau de neurones doit à terme prédire lorsqu'on lui présente le même échantillon.

Nous supposons connue une fonction $f : \mathbb{R}^l \to \mathbb{R}^n$ sur un ensemble fini $B_a = ((x_r, f(x^r)), r \in [\![1, p]\!])$ qui représente la base d'apprentissage. Un Perceptron est recherché à une couche cachée comportant l entrées, n unités en sortie, et un nombre m fixé d'unités dans la couche cachée approchant au mieux la fonction f. Une fois choisies les fonctions d'activation ϕ_c et ϕ_s, partout dérivables, un Perceptron est déterminé par les valeurs des poids (W^1_{ij}) et (W^2_{jk}). Définissons la fonction d'erreur quadratique sur la base d'apprentissage :

$$E((W^1_{ij}), (W^2_{jk})) = \sum_{r=1}^{p} ||f(x^r) - s^r||^2,$$

où s^r = P(W^1,W^2,ϕ_c,ϕ_s)(x^r) est la valeur calculée en x^r par le Perceptron, et || || est la norme Euclidienne usuelle sur $\mathbb{R}^n$. E est une fonction des (l + 1)m + (m + 1)n variables (W^1_{ij},W^2_{jk}), que l'on cherche à minimiser. On utilise pour cela une méthode itérative de minimisation : la méthode du gradient.

Notons $\nabla g(z) = (\frac{\partial g}{\partial z_1}(z),...,\frac{\partial g}{\partial z_n}(z))$ le vecteur gradient de g en z $\in \mathbb{R}^n$. L'algorithme est le suivant :

Algorithme (méthode du gradient)

1. Choisir z(0) dans $\mathbb{R}^n$, tel que $\nabla g(z(0)) \neq 0$;
2. Mettre t à 0 ;
3. Calculer z(t+1) := z(t) - $\eta(t) \cdot \nabla g(z(t))$, t := t+1 ;
4. Si | z(t) - z(t+1)| > ϵ aller en 3. Fin si.
5. Fin.

On impose ∇g(z(0)) $\neq$ 0, car z(0) pourrait être un maximum local. Si le paramètre $\eta(t)$ est choisi assez petit, on a bien g(z(t + 1)) < g(z(t)). En effet, le développement de Taylor de g en z(t + 1) s'écrit :

$$g(z(t+1)) = g(z(t)) - \eta(t)||\nabla g(z(t))||^2 + \eta(t) \cdot \epsilon(\eta(t)),$$

où ϵ(x) $\to$ 0, pour x $\to$ 0. On considérera dans la suite $\eta(t)$ constante ou décroissante vers 0 pour t $\to \infty$.

L'algorithme du gradient est appliqué au problème de minimisation de E. Il faut déterminer le gradient de E.

On pose $E_r = ||f(x^r) - s^r||^2$. Alors E = $\sum_{r=1}^{p} E_r$ et $\nabla E = \sum_{r=1}^{p} \nabla E_r$.

Calculons ∇E_r :

Par définition, $E_r = ||f(x^r) - s^r||^2 = \sum_{k=1}^{n} (y_k^r - s_k^r)^2$, où on a posé $f(x^r) = (y_1^r, ..., y_n^r)$. Posons $c^r = (c_j^r = \phi_c(W_{.,j}^1 \cdot x^r)$, j = 1,...,m) où $W_{.,j}^1 = (W_{1,j}^1$,...,$W_{n+1,j}^1$ et $c^r = (c_1^r$,...,c_m^r,c_{m+1}^r = - 1).

Pour j = 1,...,m + 1 et k = 1,...,n fixés, on a :

$$\frac{\partial E_r}{\partial W_{j,k}^2} = -2(y_k^r - s_k^r)\phi_s^{'}(W_{.,k}^2 \cdot c^r) \cdot c_j^r = \delta_j^r c_j^r,$$

où on a posé $\delta_k^r = -2(y_k^r - s_k^r) \cdot \phi_s^{'}(W_{.,k}^2 \cdot c^r)$.

Pour i = 1,...,l et j = 1,...,m + 1 fixés, on a :

$$E_r = \sum_{k=1}^{n} (y_k^r - \phi_s(W_{.,k}^2 \cdot c^r))^2,$$

Donc :

$$\begin{aligned}
\frac{\partial E_r}{\partial W_{i,j}^1} &= -2 \sum_{k=1}^{n} (y_k^r - \phi_s(W_{.,k}^2 \cdot c^r))\phi_s^{'}(W_{.,k}^2 \cdot c^r) \frac{\partial}{\partial W_{i,j}^1}(W_{.,k}^2 \cdot c^r) \\
&= -2 \sum_{k=1}^{n} (y_k^r - s_k^r)\phi_s^{'}(W_{.,k}^2 \cdot c^r) W_{j,k}^2 \phi_c^{'}(W_{.,j}^1 \cdot x^r) x_i^r \\
&= \gamma_j^r x_i^r,
\end{aligned}$$

où on a posé γ_j^r = -2 $\sum_{k=1}^{n}$ $((y_k^r$ - $s_k^r)$ $\phi_s^{'}(W_{.,k}^2 \cdot c^r) W_{j,k}^2 \phi_c^{'}(W_{.,j}^1 \cdot x^r)$.

Nous pouvons désormais écrire l'algorithme d'apprentissage d'un réseau à une couche cachée par la méthode de rétropropagation du gradient (RPG) :

Algorithme (RPG)

1. Soit B_a = (x^r, f(x^r)), r = 1,...,p, M le pas d'apprentissage et des constantes d'apprentissage (η (t) > 0, t = 1,...,M) ;
2. Initialiser les poids $(W_{i,j}^1)$ et $(W_{j,k}^2)$;
3. Pour t = 1 à M :

4. Calculer c^r, s^r, $(\delta_l^r)_l$ et $(\gamma_j^r)_j$, pour r = 1,...,p ;
5. Modifier les poids, pour i =1,...l + 1, j = 1,...,m + 1, k = 1,...,n :

$$W_{i,j}^1 := W_{i,j}^1 - \eta(t) \sum_{r=1}^{p} \gamma_j^r x_i^r,$$

$$W_{j,k}^2 := W_{ij,k}^2 - \eta(t) \sum_{r=1}^{p} \delta_k^r c_j^r;$$

6. Fin Pour ;
7. Fin.

Les modèles retenus pour la prédiction du taux de rachat par individu ont été présentés. Dans la prochaine partie, nous allons présenter l'outil qui permettra de valoriser le bilan prudentiel et ainsi évaluer les indicateurs de solvabilité et de rentabilité.

Troisième partie

L'évaluation des indicateurs économiques d'un assureur vie

Cette partie vise à présenter le modèle de projection utilisé pour déterminer la valeur des indicateurs économiques de l'assureur. Nous proposerons les premiers résultats de la projection au 31/12/2017 obtenus en utilisant la loi de rachats structurels dépendant de l'ancienneté.

Chapitre 7

La présentation du modèle de projection du portefeuille de contrats

Dans ce chapitre, nous présenterons les hypothèses de modélisation considérées.

7.1 Les caractéristiques du portefeuille de contrats

Le portefeuille étudié est l'ensemble des contrats fonds en euros (302 618) issus de la base de rachat (décrite dans la partie précédente) en stock au 31/12/2017. En vue de sa projection, nous effectuons un regroupement homogène des individus par âge, ancienneté et provisions mathématiques (encours) : 137 *model points*[1] sont créés. Les caractéristiques du sous-portefeuille sont données ci-dessous :

	Minimum	Quantile à 5%	Médiane	Moyenne	Quantile à 95%	Maximum
Ancienneté	0	3	6	6,27	10	16
Encours	100 €	24 909 €	40 063 €	55 684 €	69 890 €	11 749 668 €
Âge	18	31	47	47,73	64	100

TABLE 7.1 – Caractéristiques des *model points*

L'ensemble des *models points* sont composés de 47,44% d'hommes, 7,3% d'assurés qui ont effectué un arbitrage et 51,84% qui ont versé des primes sur leur contrat.
L'ensemble des contrats représente une provision mathématique de 16,3 milliards d'euros.

7.1.1 La composition de l'actif de l'assureur

L'actif de l'assureur en représentation de la provision mathématique et des fonds propres sociaux est respectivement de 16,3 milliards d'euros et de 1,6 milliards d'euros.

L'actif est composé de 85% d'obligations d'état et 15% d'actions de type 1.

On considère l'hypothèse que l'assureur a acquis une plus-value latente de 15% au cours de la détention de ses actifs. La valeur de marché de son actif au 31/12/2017 est de

1. Un *model points* est un sous-portefeuille homogène de contrats.

18,8 milliards d'euros pour l'actif en représentation de la provision mathématique et de 1,8 milliards d'euros pour la valeur de marché des fonds propres sociaux.

7.1.2 La revalorisation des contrats

Chaque année, l'encours de l'assuré est garanti par l'assureur (pour les contrats en euros). Le contrat est de plus revalorisé tout au long de sa durée de vie.

7.1.2.1 Le taux minimum garanti

L'assureur s'engage lors de la souscription d'un contrat en euros à revaloriser chaque année l'encours de l'assuré d'un taux minimum garanti (TMG). Ce taux est défini dans l'article A. 132-3 du Code des assurances. Il est plafonné en fonction de :

— La moyenne des taux moyens servis lors des deux exercices précédents (TMS) ;

— Le taux d'intérêt technique maximal (TTM).

Le TMG à l'année N peut se définir ainsi :

$$TMG_N = min(max(120\% \cdot TTM_N; 110\% \cdot \frac{TMS_N + TMS_{N-1}}{2}); 150\% \cdot TTM_N) \quad (7.1)$$

Nous considérons un TMG nul net de frais pour nos contrats afin d'être en cohérence avec le contexte des taux bas actuel.

La revalorisation annuelle est fonction du TMG mais aussi d'une participation aux bénéfices.

7.1.2.2 La participation aux bénéfices

L'Article A331-3 définit la participation aux bénéfices : elle doit être supérieure à 85% du résultat financier et 90% du résultat technique de l'assureur.
L'assureur doit verser cette participation aux bénéfices sur tous les contrats en euros. Il dispose de 8 ans pour les distribuer à ses assurés.

Le modèle utilisé engage l'assureur à verser 85% de ses produits financiers réalisés dans l'année. Si le résultat technique est négatif pour une année, ce résultat n'affectera pas le résultat financier.

Le taux de participation aux bénéfices (PB) est calculé ainsi :

$$taux_PB = 85\% \cdot rendement_portefeuille \quad (7.2)$$

Finalement, le taux servi à l'assuré est le maximum entre le TMG et le taux de PB :

$$taux_servi = max(TMG, taux_PB) \quad (7.3)$$

7.1.2.3 Le rendement du portefeuille

Afin de déterminer le taux servi aux assurés, il faut estimer les produits financiers réalisés dans l'année pour ensuite obtenir le rendement comptable du portefeuille. Le modèle propose une modélisation des actifs simplifiée où les produits financiers sont déterminés ainsi :

— Pour les actions, le rendement considéré est la moyenne sur les cinq dernières années du rendement des actions car elles rapportent le taux de marché depuis leur achat ;
— Pour les obligations, le rendement considéré est la moyenne sur les cinq dernières années du taux à 10 ans.

Le rendement comptable du portefeuille est donc la somme des deux rendements énoncés précédemment pondérés par leur part respective dans le portefeuille d'actifs :

$$rendement_portfeuille = part_{action} \cdot rdt_action + part_{obligation} \cdot rdt_obligation \quad (7.4)$$

Afin de constater si cette hypothèse considérée sur le rendement du portefeuille est cohérente et ne fausse pas la projection, nous pouvons constater si, avec ce choix, le modèle génère des fuites de modèle.

7.1.2.4 La fuite de modèle

La fuite de modèle, ou écart de convergence, est la différence entre la valeur de marché initiale des actifs en représentation des provisions mathématiques et la somme de la provision *Best Estimate* et de la *Present Value of Future Profits*.

$$fuite_modele = VM_0 - BE - PVFP \quad (7.5)$$

Si la fuite est positive, cela signifie que l'assureur n'a pas distribué toute sa richesse entre les assurés et les actionnaires. A contraire, si la fuite est négative l'assureur a distribué plus de richesse que disponible entre les assurés et les actionnaires.

Dans le cadre de Solvabilité II, l'ACPR préconise d'être prudent, c'est à dire que si la fuite est positive, elle appartient aux assurés.

La fuite de modèle obtenue est de -0,04%. On peut considérer que l'hypothèse du rendement comptable de l'actif est cohérent avec nos scénarios économiques générés (détail dans le prochain chapitre) et donc ne fausse pas la projection. L'hypothèse est validée.

7.1.3 Les prestations

L'assureur verse des prestations à l'assuré en cas de rachat ou de décès. Le rachat est modélisé par le rachat structurel et conjoncturel. On considère que les prestations ont lieu en milieu d'année.

7.1.3.1 La mortalité

Lors du décès d'un assuré, l'assureur verse au bénéficiaire le montant de provision mathématique présent sur le contrat. Nous utilisons la table réglementaire TH 002 et TF 002 afin d'avoir les probabilités de décès q_x pour chaque âge x et donc de diminuer chaque année le montant de provisions mathématiques des contrats.

La prestation mortalité se calcule ainsi :

$$prestation_mortalité = PM_{Ouverture} \cdot q_x \quad (7.6)$$

7.1.3.2 Le rachat structurel

Lors d'un rachat, l'assureur verse au bénéficiaire le montant de provisions mathématiques égal au montant demandé par l'assuré. Nous utiliserons les lois construites dans la deuxième partie, c'est à dire la loi en fonction de l'ancienneté, celle en fonction de l'ancienneté adaptée et les lois construites par les techniques de *machine learning*.

Nous considérons que le taux de rachat est constant au delà de 16 ans d'ancienneté du contrat.

L'utilisation du *machine learning* dans la projection

A chaque pas de temps, l'âge, l'ancienneté et la provision mathématique évoluent. Chaque assuré voit son profil changé et donc sa capacité à racheter son contrat également. Pour rappel, les modèles prédictifs utilisent aussi la variable « Sexe », « Versement » et « Arbitrage ». Dans le modèle de projection, nous ne projetons pas les primes et les assurés ne peuvent pas arbitrer car seuls les fonds en euros du contrat multi-support sont étudiés. Nous considérons que les variables « Versement » et « Arbitrage » caractérisent le profil de l'assuré et qu'il n'évolue pas lors de la projection.

Lors de la projection, le taux de rachat sera prédit à chaque pas de temps basé sur les six variables explicatives énoncées ci-avant (avec l'âge, l'ancienneté et la provision mathématique qui auront évolué).

Les prestations associées aux rachats structurels sont calculées ainsi :

$$prestation_rachat_struc = PM_ouverture \cdot taux_rachat_struc \tag{7.7}$$

7.1.3.3 Le rachat conjoncturel

Le rachat conjoncturel ou rachat dynamique dépend du contexte économique du marché. Pour estimer ce rachat, l'ACPR propose dans ses Orientations Nationales Complémentaires du LTGA deux lois qui modélisent un plafond et un plancher de rachat. Dans le modèle, nous choisissons la loi de rachats médiane entre les deux lois proposées.

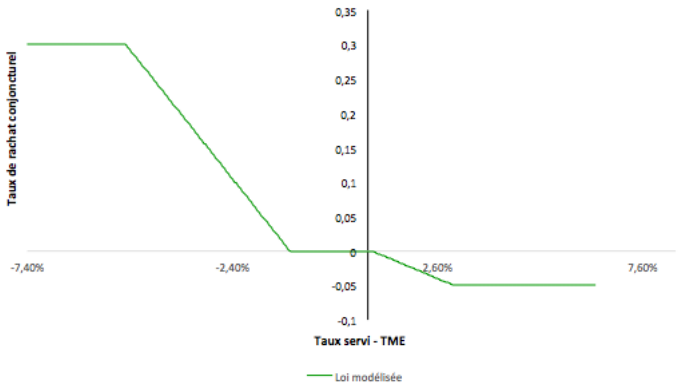

FIGURE 7.1 – Représentation graphique de loi de rachats conjoncturels.

Comme pour le rachat structurel, les prestations associées aux rachats conjoncturels se calculent ainsi :

$$prestation_rachat_conj = PM_ouverture \cdot taux_rachat_conj \tag{7.8}$$

Finalement, les prestations annuelles du portefeuille sont la somme des trois prestations exprimées auparavant :

$$prestation_totale = prestation_mortalité + prestation_rachat_struc + prestation_rachat_conj \tag{7.9}$$

7.1.4 Les frais, chargements et commissions

Pour exercer son activité, l'assureur paye des frais de gestion. Nous faisons l'hypothèse que les frais représentent 0,25% de l'encours total du portefeuille. Tous les ans, l'assureur détermine ses frais comme suit :

$$frais_gestion = taux_frais \cdot (PM_{Ouverture} - \frac{prestation_totale}{2}) \tag{7.10}$$

Nous considérons que le portefeuille est en *run-off*, c'est à dire qu'il n'y a pas d'entrée au cours de la projection. Les seuls chargements envisagés sont ceux associés à la gestion des contrats. Le taux de chargement à appliquer est supposé égal à 0,6 % des provisions mathématiques. Ils sont calculés ainsi :

$$chargement = taux_chargement \cdot (PM_{Ouverture} - \frac{prestation_totale}{2}) \tag{7.11}$$

Enfin, l'assureur rémunère son réseau de distribution à hauteur de 0,3% de l'encours total du portefeuille. Le taux de commission annuel est calculé comme suit :

$$commission = taux_commission \cdot (PM_{Ouverture} - \frac{prestation_totale}{2}) \tag{7.12}$$

7.1.5 L'estimation de la provision *Best Estimate* et de la *Present Value of Future Profits*

Afin de déterminer la provision *Best Estimate* et la *Present Value of Future Profits*, nous devons estimer tous les flux futurs, c'est à dire les prestations, les commissions, les frais, les chargements et le rendement comptable.

A chaque pas de temps i, la provision mathématique est déterminée ainsi :

$$PM[i] = PM[i-1] + revalorisation[i] - prestation[i] - chargement[i] \tag{7.13}$$

La provision *Best Estimate* est la somme des flux futurs actualisés. Nous utilisons les taux courts pour le calcul des facteurs d'actualisations.

La provision *Best Estimate* se détermine ainsi :

$$BE = \frac{1}{1000} \sum_{i=1}^{1000} (\sum_{k=1}^{30} [\sum_{j=1}^{137} (Flux_{i,k,j} \cdot exp(-\sum_{l=0}^{k-1} Taux_Actualisation_{k,l}))]) \tag{7.14}$$

Avec 30 l'horizon de projection, 1000 le nombre de scénarios économiques et 137 le nombre *model points*. Au terme de la projection, il faut liquider la provision mathématique, c'est à dire que dans le dernier flux nous intégrons la provision mathématique de fermeture.

La *Present Value of Future Profits* représente les résultats futurs acutalisés de l'assureur. Le résultat est la somme d'un résultat technique et d'un résultat financier.

La formule est la suivante :

$$PVFP = \frac{1}{1000} \sum_{i=1}^{1000} (\sum_{k=1}^{30} [\sum_{j=1}^{137} (Résultat_{i,k,j} \cdot exp(-\sum_{l=0}^{k-1} Taux_Actualisation_{k,l}))]) \tag{7.15}$$

La modélisation du passif a été décrite, nous allons maintenant présenter le Générateur de Scénarios Économiques.

7.2 Le Générateur de Scénarios Économiques

Afin d'évaluer la valeur actuelle probable des flux de trésorerie futurs, nous avons besoin de projeter différents scénarios risque neutre des variables financières suivante :

— Les taux d'intérêt nominaux ;

— Le rendement des actions.

Le Générateur de Scénario Economique (GSE) utilisé dans ce mémoire est le GSE de FIXAGE. 1 000 scénarios seront simulés sur un horizon de 30 ans.

7.2.1 La génération des taux courts nominaux

La génération des taux courts nominaux permet de calculer les prix des obligations zéro-coupon P(t,T) à la date t de maturité T. Le prix est déterminé comme suit :

$$P(t,T) = \mathbb{E}[e^{\int_t^T r_s ds}] \tag{7.16}$$

Avec r_s le taux court nominal de la date s. Un modèle de taux pourra nous fournir la dynamique de ce taux.

7.2.1.1 Le modèle de Vasicek à deux facteurs

Le GSE de FIXAGE utilise le modèle de Vasicek à deux facteurs pour décrire la dynamique des taux courts nominaux à partir des taux longs. L'équation de diffusion est la suivante :

$$\begin{cases} \mathrm{d}r_t = \alpha_1 (l_t - r_t)\,\mathrm{d}t + \sigma_1 \mathrm{d}W_t^1 \\ \mathrm{d}l_t = \alpha_1 (\mu - l_t)\,\mathrm{d}t + \sigma_2 \mathrm{d}W_t^2 \end{cases} \tag{7.17}$$

Avec :

— r_t le taux court en t ;

— α_1 la vitesse de convergence du taux court vers le taux long ;

— σ_1 la volatilité du taux court ;

— l_t le taux long terme en t ;

— μ la moyenne du taux long ;

— α_2 la vitesse de retour à la moyenne du taux long ;

— σ_1 la volatilité du taux long terme ;

— W_t^1 et W_t^2 deux mouvements browniens standards indépendants.

L'avantage de ce modèle est qu'il dispose d'une formule fermée permettant d'obtenir le prix des obligations zéro-coupon pour chaque date t et pour toute maturité T.
La formule permettant d'obtenir le prix des zéro-coupon de maturité T à la date t est :

$$P(t,T) = e^{A(T-t) - B_1(T-t)\cdot r_t - B_2(T-t)\cdot l_t} \tag{7.18}$$

Avec :

$$B_1(t) = \frac{1 - e^{-\alpha_1 \cdot t}}{\alpha_1}$$

$$B_2(t) = \frac{\alpha_1}{\alpha_1 - \alpha_2}(\frac{1 - e^{-\alpha_2 \cdot t}}{\alpha_2} - \frac{1 - e^{-\alpha_1 \cdot t}}{\alpha_1})$$

$$A(t) = (B_1(t) - t)(\mu - \frac{\sigma_1^2}{2\alpha_1^2}) + B_2(t) \cdot \mu - (\sigma_1^2 \cdot \frac{B_1(t)^2}{4\alpha_1}) +$$

$$\frac{\sigma_2^2}{2}(\frac{t}{\alpha_2^2} - \frac{2(B_1(t) + B_2(t))}{\alpha_2^2} + \frac{1 - e^{-2\alpha_1 t}}{2\alpha_1(\alpha_1 - \alpha_2)^2} - \frac{\alpha_1(1 - e^{-(\alpha_1 + \alpha_2)t})}{\alpha_2(\alpha_1 - \alpha_2)^2(\alpha_1 + \alpha_2)} + \frac{\alpha_1^2(1 - e^{2\alpha_2 t})}{2\alpha_2^3(\alpha_1 - \alpha_2)^2})$$

Pour chaque scénario, la courbe des taux est déterminée comme suit :

$$R(t,T) = -\frac{ln(P(t,T))}{T - t} \tag{7.19}$$

Le modèle de Vasicek à deux facteurs est discrétisé avec un pas de temps annuel selon un schéma d'Euler afin d'obtenir les trajectoires de taux pour tout t entier positif :

$$\begin{cases} r_{t+1} = r_t + \alpha_1 (l_t - r_t) + \sigma_1 \epsilon_{0,1}^1 \\ l_{t+1} = l_t + \alpha_2 (\mu - l_t) + \sigma_2 \epsilon_{0,1}^1 \end{cases} \tag{7.20}$$

Avec $\epsilon_{0,1}^1$ et $\epsilon_{0,1}^1$ deux lois normales centrées réduites indépendantes.

Le calibrage du modèle

Dans l'objectif de générer des trajectoires des taux courts nominaux cohérentes avec le marché, un calibrage est effectué en minimisant l'écart quadratique entre les taux zéro-coupon publiés par l'EIOPA au 31/12/2017 et les taux zéro-coupon simulés tout en respectant les tests de martingalité et de *market consistency*.
Les paramètres obtenus pour le modèle de Vasicek à deux facteurs sont :

r_0	α_1	σ_1	l_0	α_2	σ_2	μ
-0.00358	0.241	0.011	0.009	0.207	0.006	0.0251

Avec les paramètres calibrés, nous obtenons les caractéristiques des trajectoires générées :

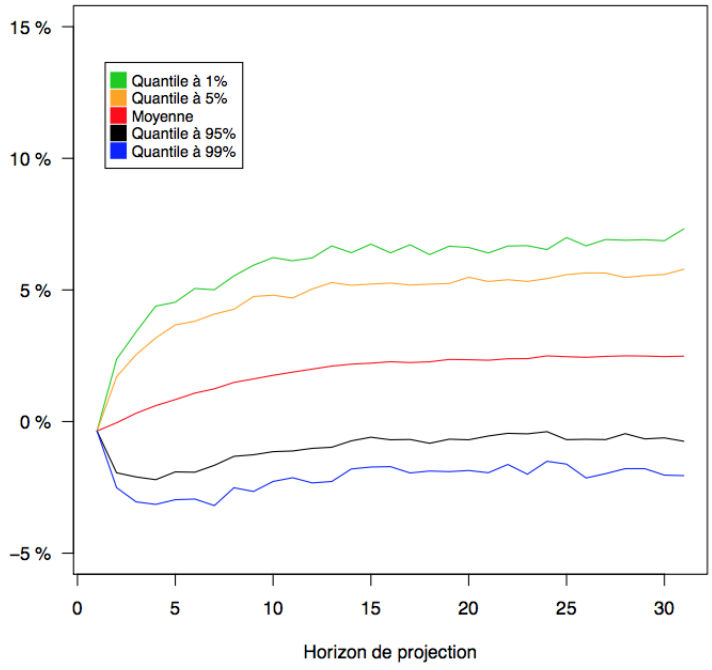

FIGURE 7.2 – Caractéristiques des trajectoires stochastiques du taux court nominal au 31/12/2017.

Le test martingale

Le test martingale pour les taux nominaux consiste à vérifier que la moyenne des prix actualisés des taux zéro-coupon à tout instant t de maturité T est égale à la valeur en t=0 du prix zéro coupon de maturité T donnée par la courbe de taux de référence :

$$P_{ref}(0,T) = \mathbb{E}[e^{-\int_0^T r_t dt}] \tag{7.21}$$

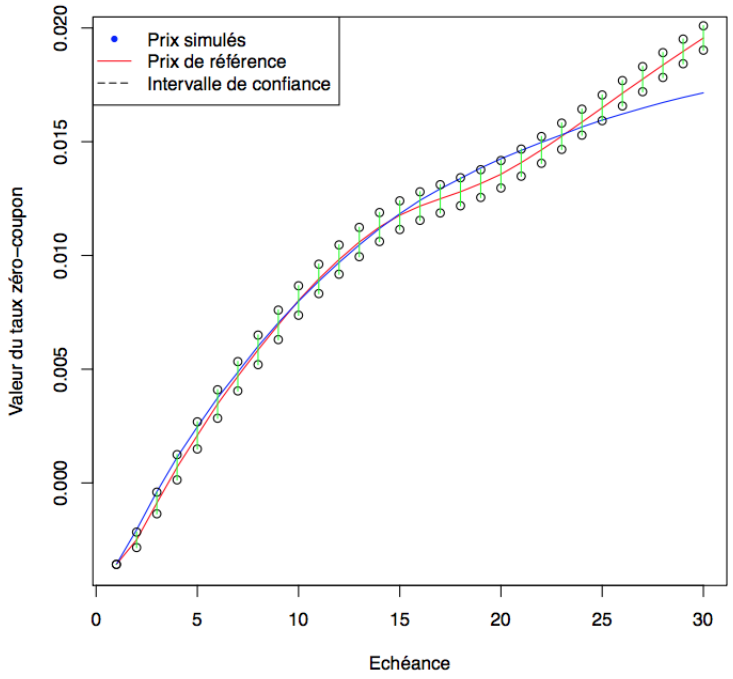

FIGURE 7.3 – Test de martingalité pour les taux courts.

Le graphe ci-dessus nous montre que les maturités inférieures à 16 ans sont bien comprises dans l'intervalle de confiance. L'écart existant au delà de 16 ans est dû à la mauvaise réplication des prix zéro-coupon à l'instant inital par le modèle de Vasicek à deux facteurs.

Le test de *market consistency*

Le test *market consistency* permet de vérifier que les trajectoires simulées permettent de retrouver le prix de marché d'un put de maturité 1 an sur une obligation d'État de maturité 10 ans de nominal égal à 100 émise au pair et qui verse un coupon fixe annuel. Au 31/12/2017, le prix de l'option est de 2,72.

L'intervalle de confiance à 95% est [2,53 ; 2,9]. Le test est validé.

7.2.2 La génération de l'indice actions

7.2.2.1 Le modèle de Black & Scholes

La performance annuelle des actions est modélisée par le modèle de Black & Scholes.

Le processus suivi par l'action est un mouvement brownien géométrique. Il est caractérisé par son équation de diffusion :

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{7.22}$$

Avec :

— S_t est le prix de l'actif sous-jacent au temps t

— μ la dérive

— σ la volatilité
— $W = (W_t)_{t\geq 0}$ est un mouvement Brownien

En appliquant la formule d'Itô à $\ln(S_t)$, la résolution de l'équation différentielle donne :

$$S_t = S_0 \cdot e^{\left[(\mu - \frac{\sigma^2}{2})t + \sigma W_t\right]} \tag{7.23}$$

où S_0 est la valeur de l'actif au temps $t = 0$.

Le calibrage du modèle

Comme pour le modèle de taux, les paramètres du modèle action sont calibrés. Ces derniers sont tout d'abord calibrés en monde réel puis les trajectoires de la performance de l'indice sont calibrées en risque neutre afin de centrer les trajectoires autour du taux court et ajuster la volatilité afin qu'elle soit égale à la volatilité implicite du CAC 40 au 31/12/2017 (12,41%).
Les paramètres obtenus pour le modèle de Black & Scholes sont :

μ	σ
0.143	0.154

Avec les paramètres calibrés, nous obtenons les trajectoires suivantes :

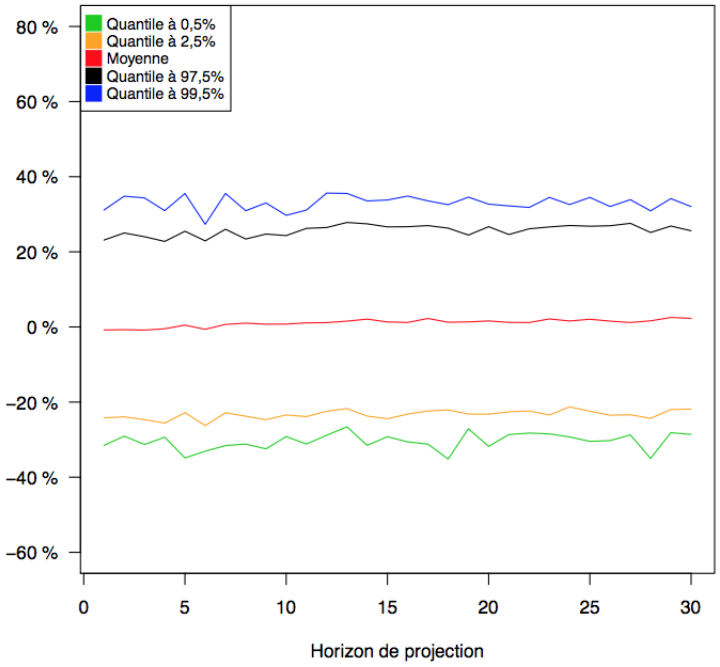

FIGURE 7.4 – Caractéristiques des trajectoires stochastiques des performances actions au 31/12/2017.

Le test martingale

Le test martingale pour les performances des actions consiste à vérifier que la moyenne des valeurs de marché actualisées des actions soit une martingale. C'est à dire que pour un prix de marché de 100 euros, les prix actualisés à chaque pas de temps doivent être égaux à 100.

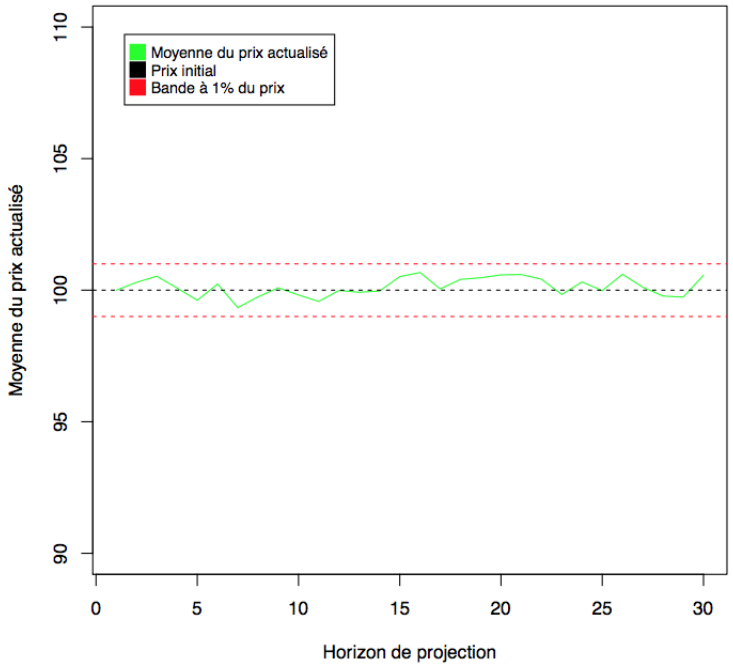

FIGURE 7.5 – Test de martingalité pour les actions.

Le graphe ci-dessus permet de constater que, dans un intervalle de confiance à 1%, le test est vérifié.

Test de *market consistency*

Le test de *market consistency* pour les performances actions consiste à choisir une volatilité implicite pour le modèle action afin d'obtenir des prix cohérents avec le marché. La volatilité choisie pour le modèle action est celle de l'indice du CAC 40 au 31/12/2017. Le test est vérifié.

Après avoir décrit la génération des taux des courts nominaux et de l'indice des actions, nous allons décrire la dépendance entre les actifs lors de la génération des scénarios économiques.

7.2.3 La corrélation entre les actifs

Le GSE de Fixage est un modèle qui génère les scénarios des différents actifs puis les agrègent entre eux.

En supposant que la dépendance entre les taux courts et les actions est linéaire, une matrice de corrélation est introduite où chaque élément correspond à un coefficient de corrélation.

Les modèles stochastiques utilisés nécessitent de générer des mouvements browniens qui sont centrés et qui ont pour matrice de variance-covariance, la matrice de corrélation des actifs. A l'aide de la décomposition de Choleski, nous pouvons simuler le vecteur brownien.

Les hypothèses de modélisation sont désormais présentées, nous allons, dans le prochain chapitre, exposer les résultats obtenus à l'aide de l'outil de projection.

Chapitre 8

Le résultat de la projection de référence en date initiale

Dans ce chapitre, nous allons décrire les résultats issus du modèle de projection dans une situation de référence. La loi de rachats en fonction de l'ancienneté (construite dans la deuxième partie) sera la loi de rachats de référence.

8.1 La provision *Best Estimate* et la *Present Value of Future Profits*

Tout d'abord, nous présentons les estimations de la provision *Best Estimate* et de la *Present Value of Future Profits*.

8.1.1 La provision *Best Estimate*

La provision *Best Estimate* est évaluée en actualisant les flux futurs. Dans le modèle, les flux sont les prestations décès et rachats, les frais et les commissions. La provision *Best Estimate* est évaluée à **18 060** millions d'euros soit 756 millions de moins que la valeur de marché de l'actif en représentation des provisions mathématiques (18,8 milliards).

Afin de s'assurer que l'approximation de la provision *Best Estimate* est cohérente, nous proposons de visualiser ci-dessous la convergence des 1 000 BE simulés lors de la projection :

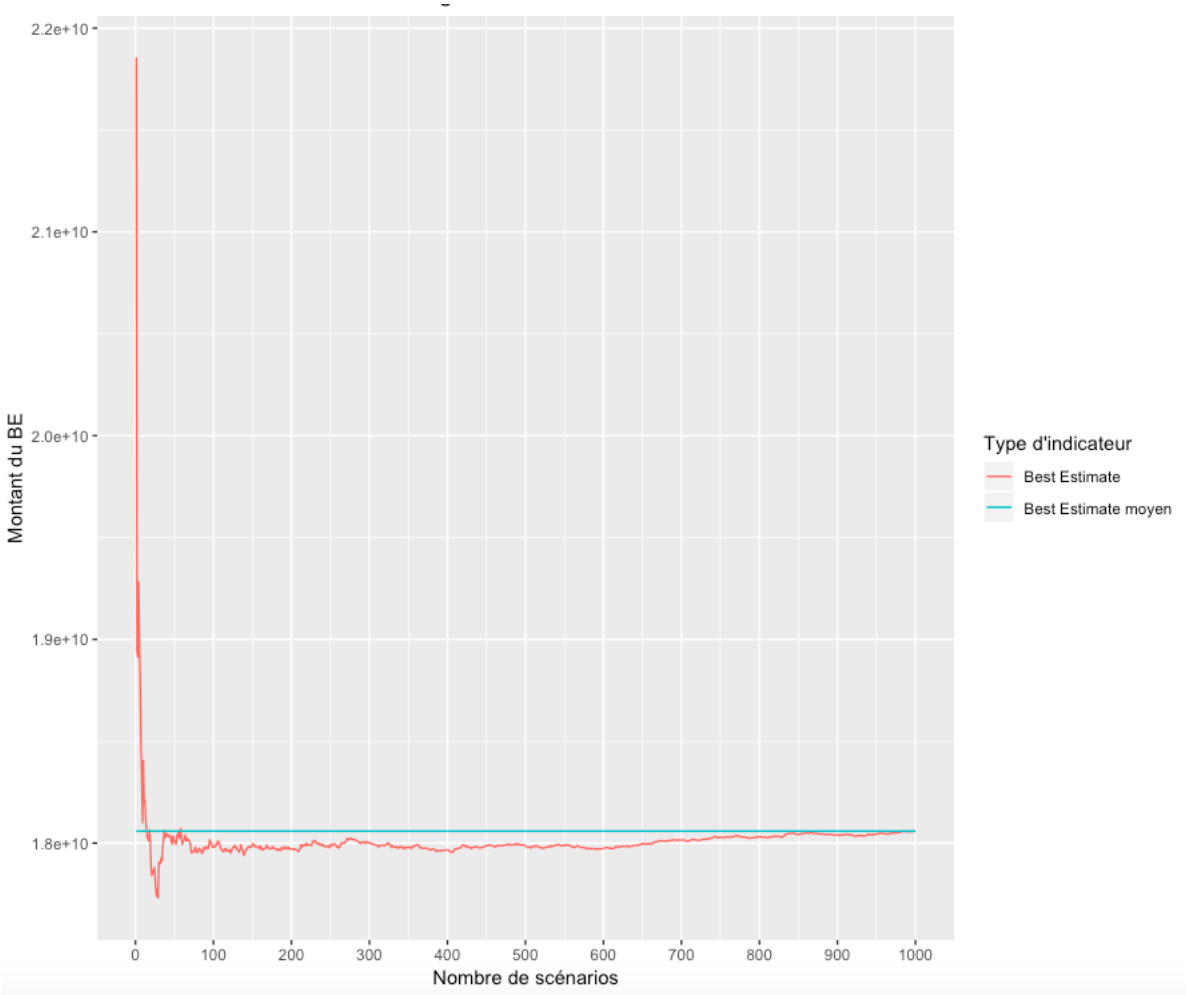

FIGURE 8.1 – Convergence du BE.

Le graphe retrace la moyenne des BE des scénarios de 1 à n avec n qui varie de 1 à 1000. La convergence du BE est réalisée rapidement, l'approximation du BE est validée.

Dans le référentiel Solvabilité II, les fonds propres économiques de l'assureur sont évalués en soustrayant la valeur de la provision *Best Estimate* à la valeur de marché de l'actif (en représentation des provisions mathématiques et des fonds propres sociaux) et en y ajoutant les fonds propres sociaux en valeur de marché :

$$\begin{aligned} Fonds_Propres_{SII} &= VM_{provisions_mathématiques} + VM_{fonds_propres_sociaux} - BE \\ &= 18\,800 \text{ M} € + 1\,900 \text{ M} € - 18\,100 \text{ M} € \\ &= 2\,600 \text{ M} € \end{aligned}$$

8.1.2 La *Present Value of Future Profits*

La *Present Value of Future Profits* est évaluée en actualisant les résultats futurs de l'assureur. Pour rappel, le résultat se décompose en un résultat technique et en un résultat financier :

— La somme des résultats techniques actualisés est 83 M€ ;

— La somme des résultats financiers est 679 M€.

Finalement, la PVFP est égale à :

$$\begin{aligned} PVFP &= 83\,\text{M}\,€ + 682\,\text{M}\,€ - min(0, fuite_de_modèle) \\ &= 765\,\text{M}\,€ - 8,5\text{M}\,€ \\ &= \mathbf{757}\,\text{M}\,€ \end{aligned}$$

Comme pour la provision *Best Estimate*, nous proposons de vérifier la cohérence des 1 000 PVFP simulées :

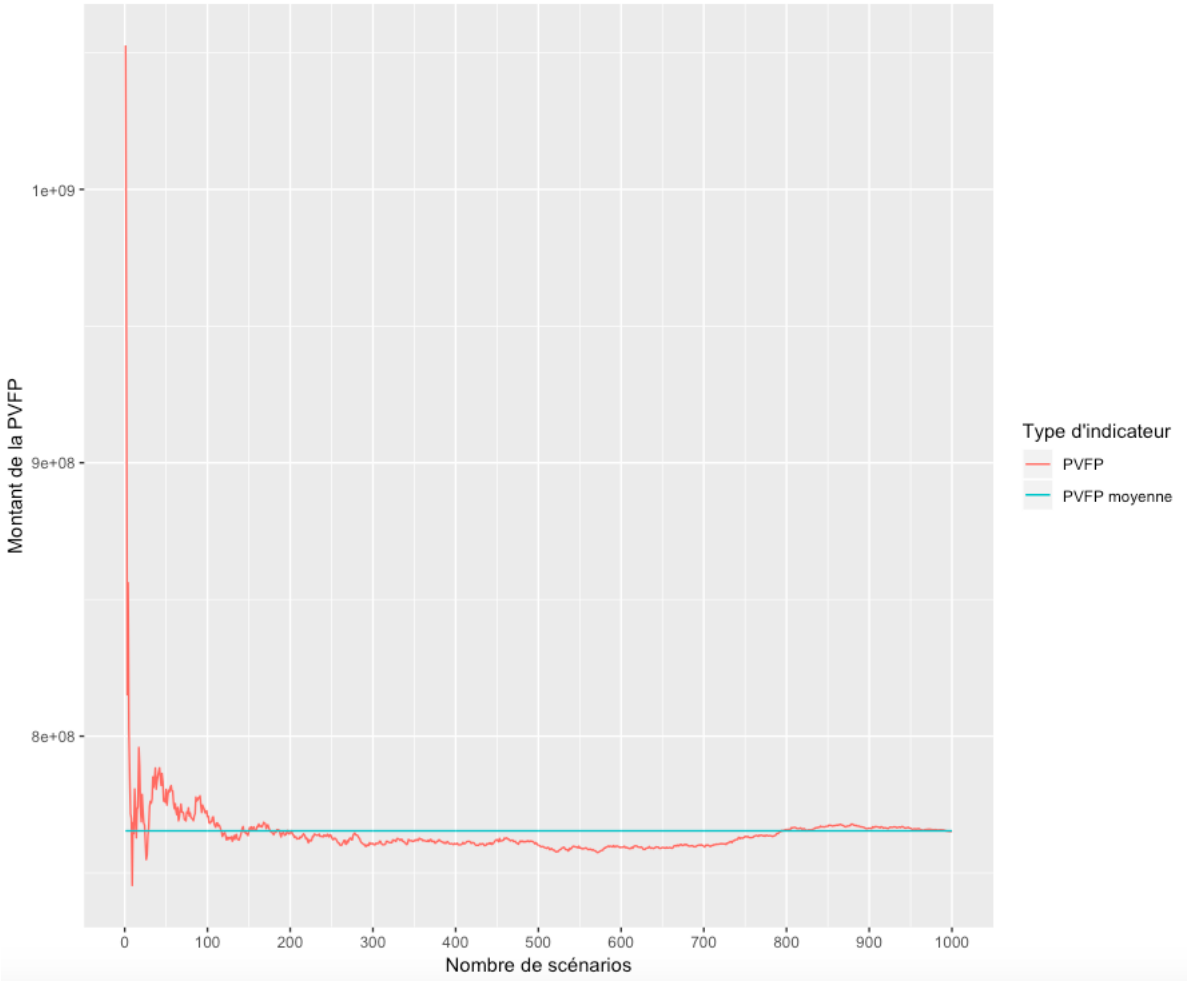

FIGURE 8.2 – Convergence du PVFP.

La convergence des moyennes des PVFP a lieu rapidement, l'approximation est validée.

8.1.3 La fuite de modèle

Afin de constater que la projection des flux de trésorerie est cohérente dans le modèle utilisé, nous avons tout d'abord effectué une projection avec un scénario déterministe. L'écart de convergence est le suivant :

$$VM_0 - BE_{déterministe} - PVFP_{déterministe} = 0$$

La fuite de modèle est nulle donc celui-ci a une projection cohérente des flux de trésorerie.

Dans le cas de la projection stochastique, le modèle génère une fuite de -8,5 M € soit -0,04% de la valeur de marché de l'actif en représentation des provisions mathématiques. Bien que cet écart soit négligeable, il a été alloué à la *Present Value of Future Profits* étant donné que celui-ci est négatif.

Nous allons désormais présenter les résultats du *Solvency Capital Requierement* (SCR) et du ratio de couverture.

8.2 Le SCR et le ratio de couverture

Les chocs des sous-modules de risque ont été appliqués au bilan afin d'estimer le *Solvency Capital Requierement*. Des fuites de modèle peuvent apparaître après avoir choqué le bilan. Lors du calcul du SCR, pour chaque sous-module de risque, nous prenons en compte cet écart dans le calcul.

Nous obtenons les résultats (en millions d'euros) suivants :

Sous-Module	Δ VM	Δ BE	SCR
Taux Baisse	-154	-337	183
Taux Hausse	1 050	60	990
Action	1 270	-170	1 440

TABLE 8.1 – SCR des sous-modules de risque marché.

Sous-Module	Δ VM	Δ BE	SCR
Mortalité	0	-16	16
Longévité	0	12	0
Rachat	0	-141	141
Frais	0	-42	42
Catastrophe	0	-56	56

TABLE 8.2 – SCR des sous-modules de risque vie.

Avec $\Delta VM = VM$ - $VM_{choqué}$ et $\Delta BE = BE$ - $BE_{choqué}$

Module	SCR
Marché	2 120
Vie	191
Global	2 170

TABLE 8.3 – SCR des modules de risque.

Le ratio de couverture obtenu est **121%**.

8.3 Le bilan prudentiel

Dans la situation de référence, l'assureur est solvable avec un ratio de couverture supérieur à 100 %. Son portefeuille de contrats en euros génère un bénéfice de 754 millions d'euros.

La valeur de la compagnie se détermine ainsi :

$$\begin{aligned} Valeur_compagnie &= PVFP + Fonds_Propres_{sociaux} \\ &= 0,75 \text{ Mds } € + 1,88 \text{ M } € \\ &= \mathbf{2.6} \text{ Mds } € \end{aligned} \tag{8.1}$$

Nous utiliserons cette situation comme situation de référence pour étudier la sensibilité des différents indicateurs économiques au changement de loi de rachats structurels. Le bilan prudentiel obtenu est le suivant :

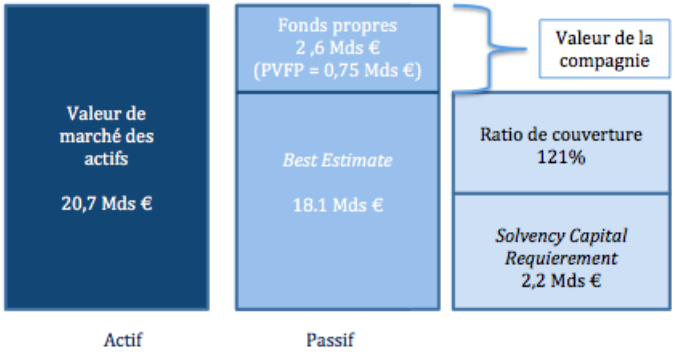

FIGURE 8.3 – Bilan prudentiel au 31/12/2017.

Dans la prochaine partie, les différentes lois de rachats construites grâce aux techniques de *machine learning* seront mise en oeuvre et les sensibilités des différents indicateurs économiques au changement de loi de rachats structurels seront estimées.

Quatrième partie

L'application des méthodes à un contrat d'assurance vie en euros

Chapitre 9

L'utilisation des techniques de *machine learning* pour construire les nouvelles lois de rachats structurels

Dans ce chapitre, les différentes techniques de *machine learning* seront mises en oeuvre dans la construction de la loi de rachats structurels puis les résultats des modèles seront comparés entre eux. Au terme de ce chapitre, les lois de rachats ainsi construites pourront être utilisées afin d'évaluer l'impact du changement de loi de rachats structurels sur les indicateurs économiques de l'assureur.

9.1 La phase d'apprentissage et le choix des paramètres

9.1.1 La phase d'apprentissage

• La préparation des bases pour l'apprentissage :

Avant d'effectuer la phase d'apprentissage, la base de rachat couvrant la période 2010 à 2017 est séparée en trois bases : une base d'entraînement, une base de validation et une base de test.

L'échantillon utilisé pour la base d'entraînement et de validation est constitué des données de 2010 à 2016. Les modèles sont ensuite testés sur l'année 2017. Les deux premières bases sont obtenues par tirage aléatoire sans remise en attribuant à la base d'entraînement 75% des données et 25% pour la base de validation. La structure de ces dernières est vérifiée afin de s'assurer que les modalités contenues dans la base « mère » soient également présentes dans les deux bases.

Les échantillons d'entraînement et de validation sont composés respectivement de 1 264 137 et 421 379 contrats. La base de test contient 347 758 contrats.

• Le choix des variables explicatives :

Initialement la base de rachat contient les variables :

Variables	Description	Nature
Statut du contrat	Indique si le contrat est sorti ou non	Classe
Date d'effet d'entrée	La date de souscription du contrat	Date
Date d'effet de sortie	La date de sortie du contrat	Date
Sexe	Le sexe de l'individu	Classe
Age	L'âge de l'individu	Numérique
Date de deces	La date de décès en cas de décès	Date
PM d'ouverture	La PM de début d'année	Numérique
PM de fermeture	La PM de fin d'année	Numérique
Rachat partiel	Le montant de rachat partiel dans l'année	Numérique
Rachat total	Le montant de rachat total dans l'année	Numérique
Versement	Le montant de versement dans l'année	Numérique
Arbitrage	Le montant arbitré dans l'année	Numérique
Anciennete	L'ancienneté fiscale du contrat	Numérique

FIGURE 9.1 – Variables contenues dans la base initiale.

Avant d'effectuer l'analyse d'importance des variables, un premier traitement des variables explicatives est effectué. Les variables « Statut du contrat », « Date d'effet d'entrée », « Date d'effet de sortie », « Date de décès » sont retirées pour leur redondance avec les autres variables explicatives et leur impertinence.

Ensuite, la variable « PM moyenne » (Provision mathématique moyenne) est créée à partir des variables « PM d'ouverture » et « PM de fermeture » en utilisant la même formule que pour l'Assiette de rachat calculée dans la partie (5.1.3). Cette nouvelle variable numérique est transformée en une variable de type « classe » constituée de tranches homogènes de PM (en nombre de contrats). De plus, à l'aide des variables « Rachat partiel » et « Rachat total », la variable « Type de rachat » est construite et contient les valeurs 0, 1 et 2 signifiant respectivement « Pas de rachat », « Rachat partiel » et « Rachat total ».

Enfin, les variables « Versement » et « Arbitrage » (qui sont toutes les deux des variables numériques) sont modifiées en deux variables binaires indiquant s'il y a eu arbitrage ($ARB_P.0$ et $ARB_P.1$) et versement ($VERS_P.0$) et $VERS_P.1$) dans l'année.

Les modèles *Random Forest* et *XGBoost* permettent de classer les variables selon leur pouvoir explicatif. Les résultats peuvent être visualisés à la page suivante :

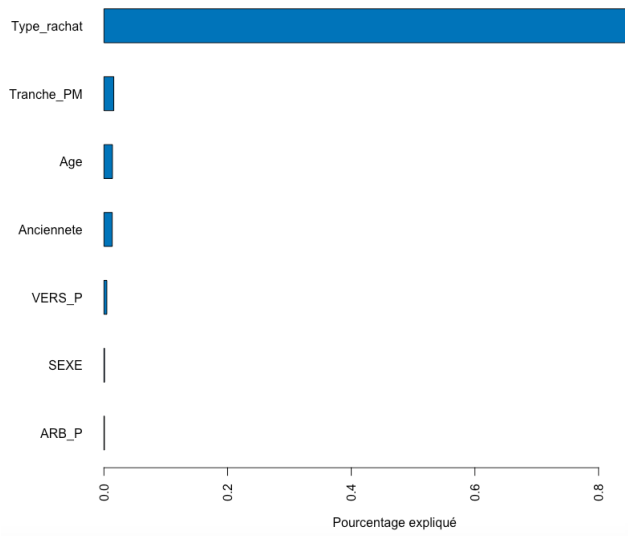

FIGURE 9.2 – Importance des variables Random Forest.

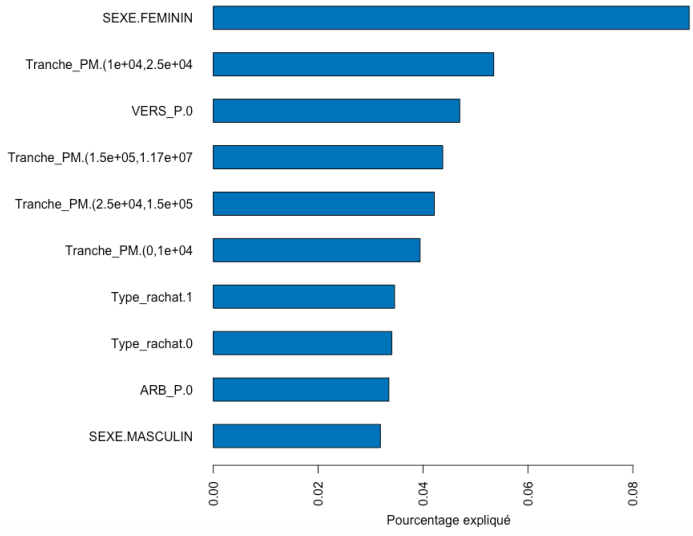

FIGURE 9.3 – Importance des variables XGBoost.

Le modèle *Random Forest* ne considère (à 95%) qu'une seule variable explicative : le type de rachat. Contrairement au modèle *XGBoost* qui utilise toutes les informations des variables. Le problème qui se pose est que le modèle *Random Forest* prédira environ le même taux de rachat pour un même type de rachat alors que toutes les autres variables seront différentes. Le modèle *XGBoost* propose une importance des variables différentes. Ce modèle capte moins bien les interactions entre les variables catégorielles car il considère chaque modalité comme des variables à part entière et donc ne prend pas en compte le caractère ordinal pour certaines d'entre elles (par exemple l'ancienneté ou l'âge).

Il apparaît que chaque variable a une importance plus ou moins significative dans au moins un des deux modèles. Les variables sont donc conservées dans la prédiction hormis la variable « Type de rachat » qui est clairement corrélée avec le taux de rachat.

9.1.2 Le choix des paramètres

La prochaine étape est la phase de calibrage, c'est à dire l'estimation des paramètres. L'objectif de cette partie est de trouver le jeu de paramètre « optimal » qui minimisera l'erreur de prédiction. Pour cela, la fonction *h2o.grid* compris dans le package H2O est utilisée afin d'obtenir un jeu de paramètre optimal pour chaque modèle.

◦ La fonction *h2o.grid* :

La « Grid Search » est une fonction du package H2O qui permet de créer des modèles de manière aléatoire. L'utilisateur spécifie exactement l'espace de l'hyperparamètre (jeu de paramètre à tester), H2O échantillonnera uniformément à partir de l'ensemble de toutes les combinaisons de valeurs d'hyperparamètres possibles. L'utilisateur spécifie également un critère d'arrêt, qui contrôle quand la « Grid Search » est terminée (sinon la recherche ne s'arrête pas). Il peut indiquer à la « Grid Search » de s'arrêter en spécifiant un nombre maximal de modèles ou le nombre maximal de secondes autorisé pour la recherche.

Une fois la « Grid Search » terminée, l'utilisateur peut interroger l'objet de *h2o.grid* et trier les modèles par une mesure de performance particulière (par exemple la RMSE). Tous les modèles sont stockés dans le *cluster* H2O et sont accessibles par identifiant de modèle.

Le temps de calcul de la fonction de recherche est très coûteux notamment quand le nombre de paramètres est important. Les paramètres optimaux ont été déterminés.

Le jeu de paramètre optimal pour le modèle *Random Forest* est :

— mtries = 2
— max_depth = 18
— ntrees = 105

Le jeu de paramètre optimal pour le modèle *GBM* est :

— learn_rate = 0,1
— max_depth = 7
— ntrees = 65

Le jeu de paramètre optimal pour le modèle *XGBoost* est :

— col_sample_rate = 0,519
— learn_rate = 0,1189
— max_depth = 7
— ntrees = 36
— sample_rate = 0,6651

Le jeu de paramètre optimal pour le modèle *deeplearning* est :

— hidden = [64,64]
— input_dropout_ratio = 0
— rate = 0,01
— rate_annealing = 1e-07
— epochs = 10

9.2 Les résultats des modèles d'apprentissage

Dans cette section la modélisation prédictive des taux de rachat issus des quatre modèles sera effectuée.

9.2.1 Les *Random Forest*

Tout d'abord, un premier modèle ensembliste est proposé comme modèle de prédiction : les *Random Forest*. En pratique, la fonction *h2o.randomForest* du package H2O est utilisée. Les graphes représentant l'importance des variables dans la prédiction des taux de rachat et la RMSE en fonction du nombre d'arbres sont présents ci-dessous.

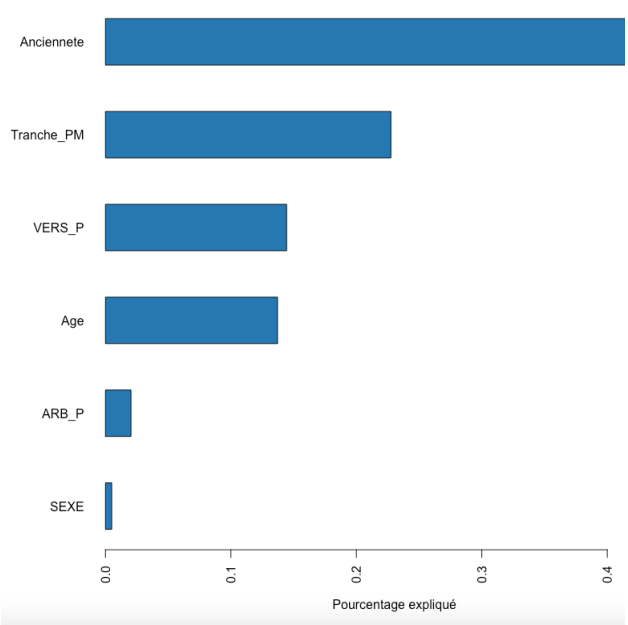

FIGURE 9.4 – Importance des variables RF.

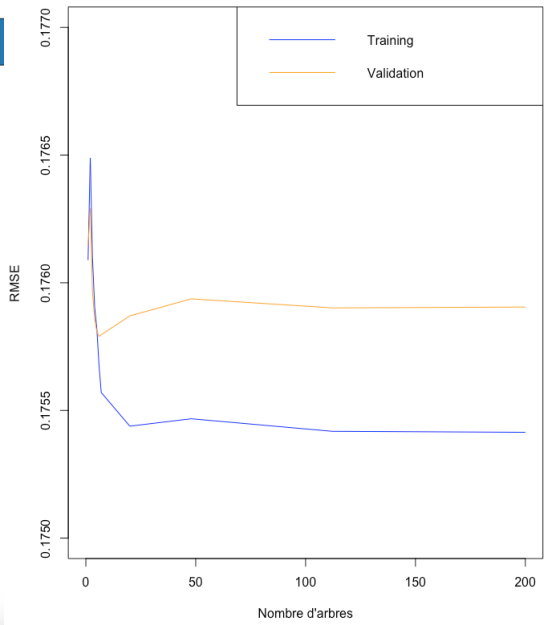

FIGURE 9.5 – RMSE en fonction du nombre d'arbres.

Tout d'abord l'ordre d'importance des variables explicatives est modifié dans ce modèle avec l'absence de la variable « Type de rachat ». L'ancienneté est devenue la variable qui explique le mieux le taux de rachat. De plus, le sexe et l'arbitrage ne sont pas très significatifs dans la prédiction du taux de rachat.

Le graphe de droite représente la RMSE sur la base d'apprentissage (*Training*) et de validation (*Validation*) en fonction du nombre d'arbres utilisés dans le modèle. Plus le nombre d'arbres augmente, plus la qualité de prédiction du modèle augmente. Une décroissance de l'erreur est naturellement observée. Celle-ci se stabilise pour un peu plus de 100 arbres avec un RMSE de 0,1754 (*Training*).

9.2.2 Le *Gradient Boosting Machine*

Un autre modèle ensembliste est ensuite proposé. Il reprend le principe général du boosting : le *Gradient Boosting Machine*. Pour l'implémentation de ce modèle, la fonction *h2o.gbm* est utilisée. Les résultats sont les suivants :

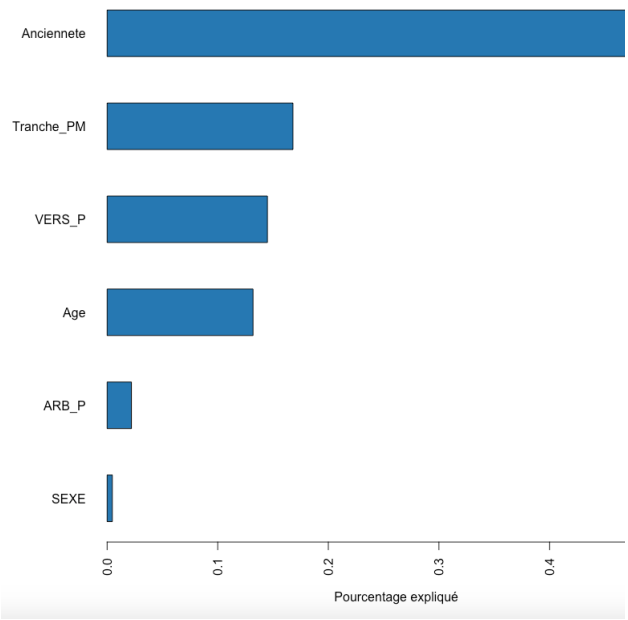

FIGURE 9.6 – Importance des variables GBM.

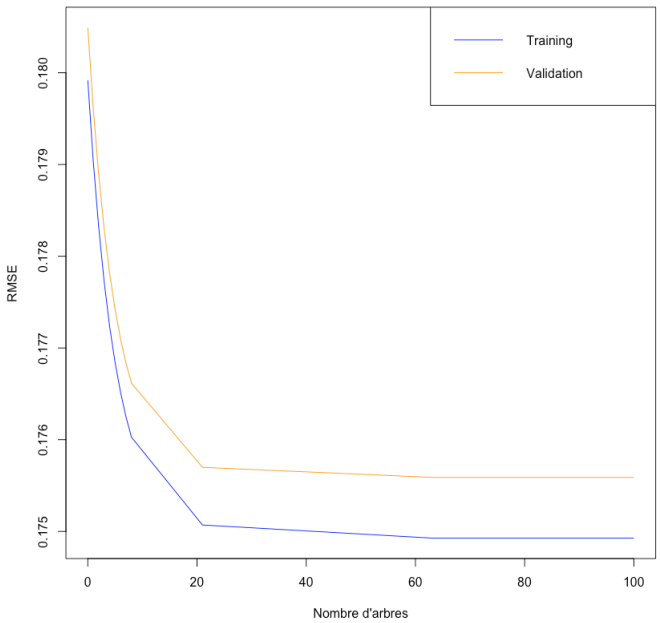

FIGURE 9.7 – RMSE en fonction du nombre d'arbres.

L'importance des variables explicatives dans la prédiction est pratiquement la même que pour le modèle *Random Forest*. L'ancienneté a un peu plus de pouvoir prédictif. De plus, la RMSE se stabilise à 0,1749 à partir de 60 arbres.

La différence entre ces deux modèles d'agrégation d'arbres (*GBM* et *Random Forest*) est que le modèle *GBM* apporte plus d'importance à l'ancienneté au détriment de la PM et que ce dernier a une erreur de prédiction plus faible que *Random Forest*.

9.2.3 *XGBoost*

Les résultats d'un dernier modèle ensembliste sont présentés : le modèle *XGBoost*. Les graphes représentant l'importance des variables et la valeur de la RMSE en fonction du nombre d'arbres se trouvent à la page suivante :

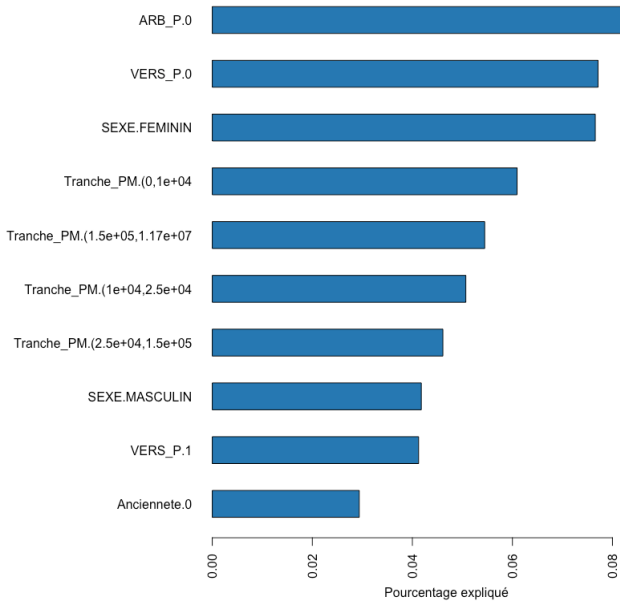

FIGURE 9.8 – Importance des variables XGBoost.

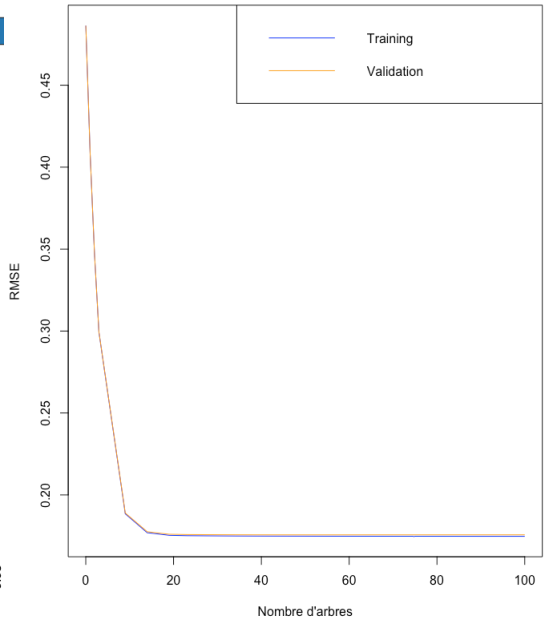

FIGURE 9.9 – RMSE en fonction du nombre d'arbres.

Pour ce modèle, un tout autre ordre d'importance des variables explicatives est observé. Le modèle semble privilégier les modalités représentant l'absence de versement et d'arbitrage ainsi que le sexe féminin. Une grande différence est constatée entre ce modèle et les modèles *Random Forest* et *GBM* dans l'importance des variables explicatives. Le modèle *XGBoost* capte mal les interactions entre les variables catégorielles de type ordinal. De plus, la RMSE se stabilise à 0,1749 à partir d'environ 35 arbres.

9.2.4 Le *Deeplearning*

Pour finir, les résultats issus d'un modèle neuronal sont présentés, le *Deeplearning*. Les résultats sont les suivants :

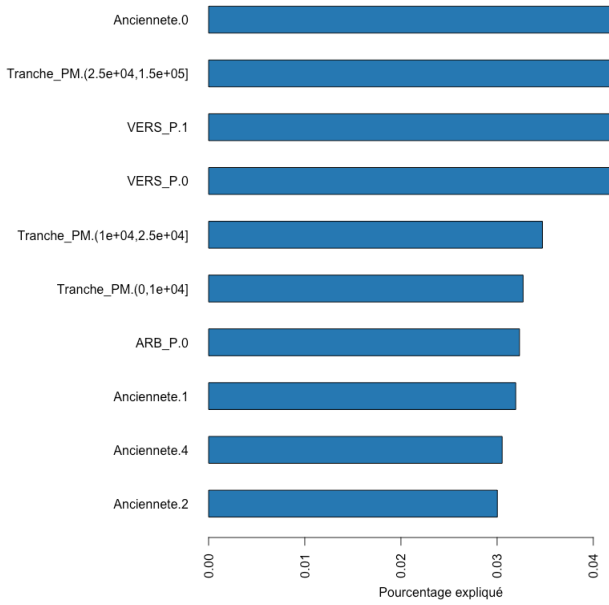

FIGURE 9.10 – Importance des variables Deeplearning.

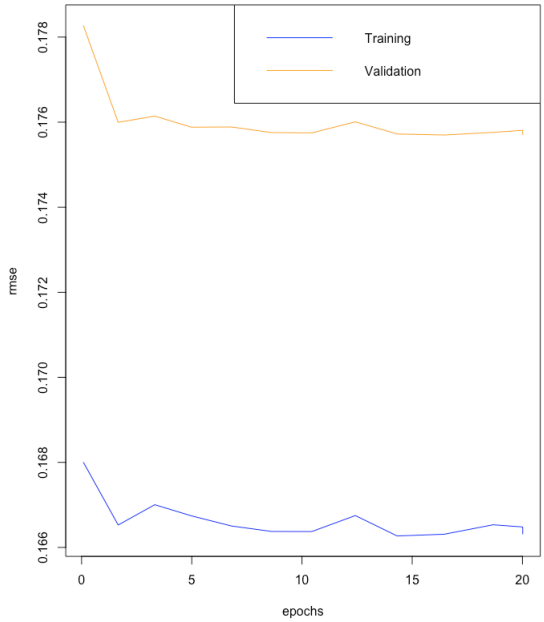

FIGURE 9.11 – RMSE en fonction du nombre d'arbres.

L'ancienneté, la PM et le versement sont retrouvés dans ce modèle comme variables qui ont le plus d'importance dans la prédiction du taux de rachat. Les résultats obtenus sont analogues à ceux observés pour les deux premiers modèles ensemblistes.

Le graphe de droite permet de constater que la RMSE ne varie pas beaucoup en fonction du nombre d'itérations (*epochs*) et que la RMSE minimale sur la base d'apprentissage est 0,1664.

9.2.5 La comparaison des modèles

Les 4 modèles sont désormais comparés avec leurs erreurs de prédiction sur la base de test.

Le taux de rachat moyen réel observé sur l'année 2017 est 4,6 %. Les résultats des taux de rachat issus des modèles sont les suivants :

Modèle	RMSE	Taux de rachat moyen
Random Forest	0,1843	4,53%
Gradient Boosting	0,18416	4,56%
Xgboost	0,18428	4,53%
Deeplearning	0,18431	4,49%

La loi de rachats en fonction de l'ancienneté propose un taux de rachat moyen de 4,84% sur l'année 2017. Les taux de rachat moyens prédits par les modèles sont tous plus faibles que le taux de rachat moyen réel mais sont plus proches du taux réel observé que le taux moyen proposé par la loi de rachats en fonction de l'ancienneté. En ne prenant en compte que l'ancienneté comme variable explicative, le taux de rachat est surestimé.

Parmi les 4 modèles, l'erreur la plus faible est celle du *Gradient Boosting Machine*. Néanmoins, l'erreur entre les modèles est très proche. De par la complexité dans le calibrage des modèles, il est difficile d'affirmer que le *Gradient Boosting Machine* est le modèle à privilégier. De plus nous observons que c'est aussi ce modèle qui a les taux de rachat prédits les plus proche en moyenne du taux de rachat réel moyen.

L'ancienneté est la variable qui explique le plus la prédiction du taux de rachat. Toutefois, elle n'est pas la seule variable qui a de l'importance. La tranche de PM, la présence de versements et l'âge ont également de l'importance dans la prédiction. De plus, l'introduction de ces variables avec les techniques de *machine learning* fournissent des taux de rachat plus proches des taux réels que les taux issus de la loi en fonction de l'ancienneté. Cette analyse vient renforcer l'idée de prendre en compte plusieurs variables explicatives.

Dans le prochain chapitre, l'impact de ces nouvelles lois sur la solvabilité et la richesse de l'assureur anonyme sera estimé et analysé.

Chapitre 10

La sensibilité des indicateurs de solvabilité et de richesse au changement de la loi de rachats structurels

Après avoir décrit les hypothèses et les étapes de la modélisation de l'outil de projection utilisé, nous présentons les résultats des différents indicateurs économiques pour les différentes lois de rachats utilisées. Dans un premier temps nous nous intéresserons à la loi de rachats en fonction de l'ancienneté adaptée puis nous énoncerons les résultats obtenus avec les lois construites à l'aide des techniques de *machine learning*.

L'objectif de ce chapitre est de constater l'impact du changement de la loi de rachats structurels sur les indicateurs économiques de l'assureur.

Le portefeuille étudié est celui présenté dans la deuxième partie. La projection est effectuée en *run-off*[1] sur les *model points* construits dans la troisième partie. La dernière année de projection, le portefeuille est liquidé.

10.1 L'effet de la loi de rachats adaptée sur les indicateurs économiques

Nous étudions l'impact du changement de loi de rachats structurels dans le modèle de projection en comparant les résultats obtenus avec ceux issus de la situation de référence, c'est à dire lorsque la loi de rachats structurels en fonction de l'ancienneté est utilisée.

Dans un premier temps, nous analysons les résultats obtenus avec la loi de rachats adaptée en fonction de l'ancienneté. Pour rappel, cette loi a été construite à l'aide de la loi de rachats en fonction de l'ancienneté et d'une fonction paramétrique qui caractérise le gain ou la perte pour les assurés (lors d'un rachat) avec la mise en place de la « Flat Tax ».

Les valeurs des indicateurs économiques dans la situation de référence et celles avec le changement de loi sont les suivantes :

1. Sans projeter de primes durant toute la projection.

Indicateur	Situation de référence	Loi de rachats adaptée	Ecart ref
Best Estimate	18 059 millions €	18 060 millions €	- 1 million €
PVFP	757 millions €	755 millions €	2 millions €
SCR	2 176 millions €	2 178 millions €	- 2 millions €
Ratio de solvabilité	121,2 %	121,0 %	0,2 %
Taux de rachat moyen annuel	3,86 %	3,87 %	-0,01 %

TABLE 10.1 – Comparaison des indicateurs avec la loi de rachats adaptée.

Le taux de rachat moyen annuel correspond au taux observé sur toute la durée de la projection (le taux à 30 ans est toujours égal à 50%, il a été retiré de la moyenne). Pour rappel, à partir de 16 ans d'ancienneté le taux de rachat est constant. Le taux moyen est légèrement plus élevé en utilisant la loi de rachats adaptée. Les contrats ayant moins de 8 ans d'ancienneté fiscale ont en effet un taux de rachat supérieur pour la loi adaptée (notamment ceux qui ont moins de 4 ans). Au-delà de cette ancienneté, le taux de rachat de la loi de référence est légèrement plus élevé. Les taux de rachat moyens observés à chaque pas de temps dans la projection sont représentés ci-dessous :

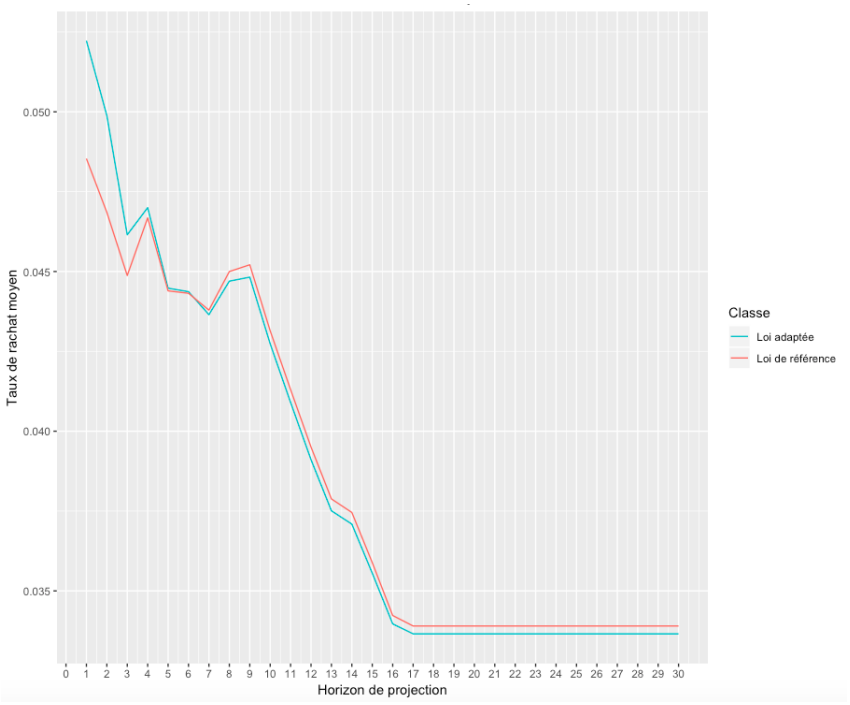

FIGURE 10.1 – Comparaison des taux de rachat moyens issus de la projection.

La loi de rachats adaptée est au dessus de la loi de référence les 4 premières années. Au delà de ce pas de projection, l'ancienneté moyenne du portefeuille est au dessus de 8 ans et donc les taux de rachat de la loi de référence sont légèrement plus élevés que les taux proposés par la loi adaptée.

Pour la dernière année de projection, nous avons fixé le taux de rachat structurel à 50%. La chronique des flux de prestations au titre du rachat structurel est représentée ci-dessous :

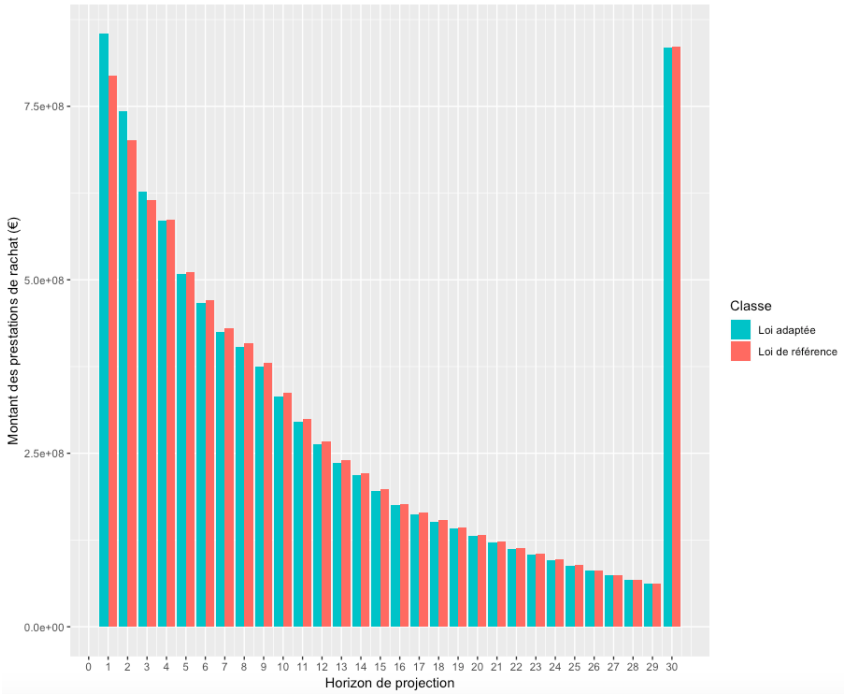

FIGURE 10.2 – Chronique des flux de prestation du rachat.

Le flux de prestations à 30 ans correspondant à la loi de rachats adaptée est moins important que celui de la loi de référence. La provision mathématique d'ouverture à 30 ans est donc inférieure dans le cas de la loi adaptée. Les prestations de rachat totales issues de la loi de référence sont plus importantes que celles issues de la loi adaptée. En effet, avec cette dernière, les prestations sont plus importantes en début de période donc la revalorisation s'effectue sur moins de provisions mathématiques par rapport à la loi de référence.

Le *Best Estimate* a augmenté légèrement en utilisant la loi de rachats adaptée. Sur le graphique représentant les taux de rachat moyens annuels, nous pouvions anticiper cette hausse. Les taux de rachat associés à cette loi sont supérieurs aux taux de rachat de la loi de référence en début de projection et avec l'effet de l'actualisation, les prestations ont plus de poids les premières années. Toutes choses égales par ailleurs, les fonds propres ont diminué et le SCR a augmenté : le ratio de solvabilité de la loi de rachats adaptée a donc diminué de 20 points de base[2]. De plus, on constate logiquement une baisse (2 millions €) de la PVFP. En effet, le contexte de taux à la fin 2017 reste bénéfique pour l'assureur qui prévoit des profits futurs sur les contrats en euros, pour lesquels les taux minimum garantis sont nuls. L'assureur génère du profit sur son portefeuille de contrats et donc est plus sensible à la hausse des rachats.

Nous allons maintenant étudier l'impact des lois de rachats construites à l'aide des techniques de *machine learning*.

2. Un point de base = 0,01 %.

10.2 L'effet des lois *machine learning* sur les indicateurs économiques

L'application des lois construites dans le chapitre précédent nous permet de constater l'impact de la prise en compte de plusieurs variables explicatives du rachat structurel sur les différents indicateurs économiques. Pour rappel, la prédiction du taux de rachat dans le modèle de projection s'effectue à chaque pas de temps et prend en compte l'évolution de la provision mathématique, l'âge et l'ancienneté par *models points* à chaque pas. Chacun de ces derniers est de plus caractérisé par son sexe et sa tendance à effectuer des versements ou des arbitrages. La base de rachats utilisée pour l'apprentissage des modèles ne contient pas de contrats où l'ancienneté fiscale est supérieure à 16 ans.

Nous comparons les résultats des projections avec ceux où la loi de rachats en fonction de l'ancienneté est utilisée (loi de référence). Les résultats du *Best Estimate*, de la PVFP et du SCR sont exprimés en millions d'euros.

Indicateur	Référence	*RF*	*GBM*	*XGB*	*DL*
Best Estimate	18 059	18 185	18 197	18 186	18 191
PVFP	757	631	618	629	624
SCR	2 176	2 198	2 198	2 197	2 197
Ratio de solvabilité	121,2 %	114,3 %	113,7 %	114,2 %	114,0 %
Taux de rachat moyen annuel	3,86 %	4,5 %	4,9 %	4,8 %	5,0 %

TABLE 10.2 – Comparaison des indicateurs avec les lois *machine learning*.

Nous constatons que la moyenne des taux de rachat est nettement supérieure avec les lois *machine learning* par rapport à la loi de référence. Pour visualiser ce résultat, les taux de rachat moyens observés à chaque pas de temps de chaque loi sont représentés ci-dessous :

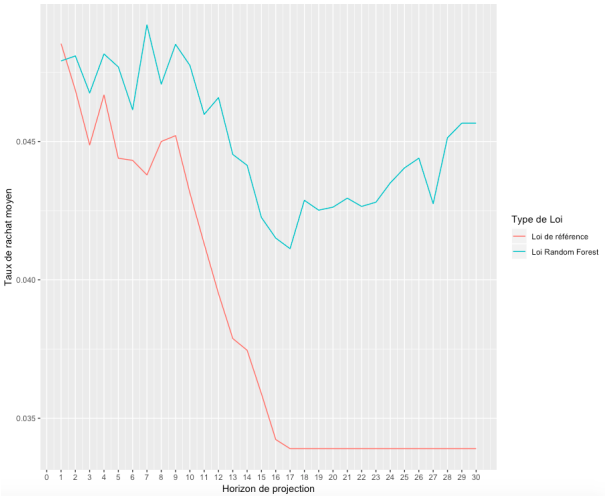

FIGURE 10.3 – Comparaison des taux de rachat moyens *Random Forest* avec la loi de référence.

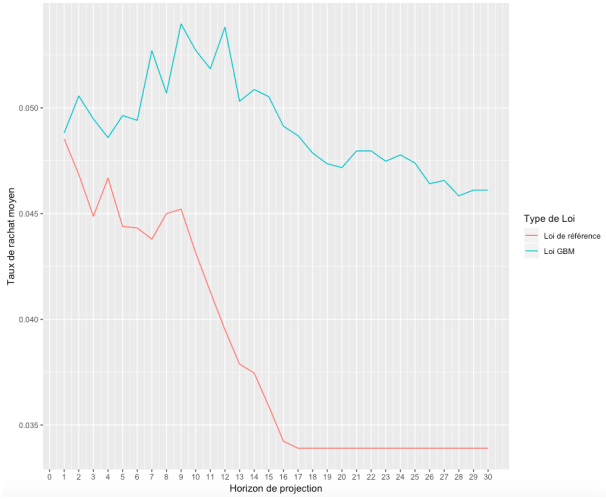

FIGURE 10.4 – Comparaison des taux de rachat moyens *GBM* avec la loi de référence.

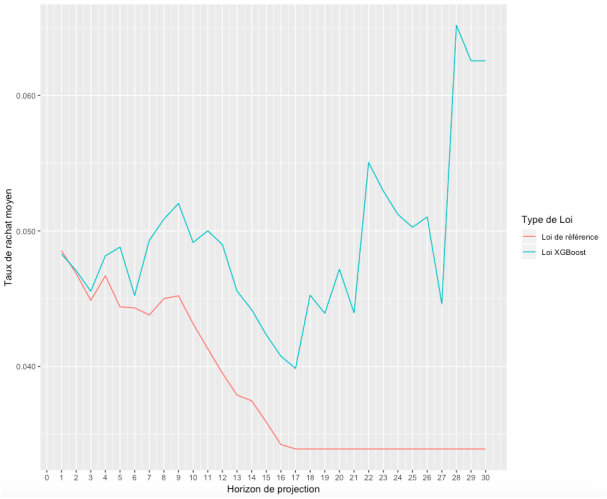

FIGURE 10.5 – Comparaison des taux de rachat moyens *Xgboost* avec la loi de référence.

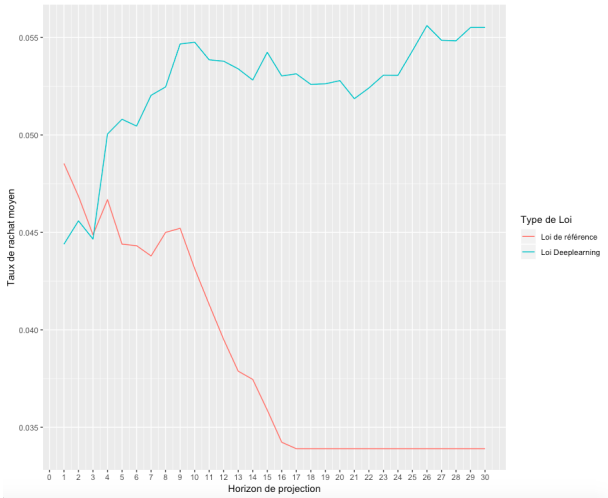

FIGURE 10.6 – Comparaison des taux de rachat moyens *Deeplearning* avec la loi de référence.

En moyenne, les taux de rachat des lois *machine learning* sont toujours plus élevés que ceux de la loi de référence. En début de projection, les différents taux sont assez proches et on observe que la loi *Deeplearning* fournit des taux inférieurs.

On peut faire un lien entre le pouvoir explicatif des variables dans les différents modèles et la représentation des taux de rachat moyens. L'importance des variables pour le modèle *Random Forest* et *GBM* est similaire. L'allure de la courbe représentant leur loi de rachats structurels est également proche. De plus, l'importance des variables dans les modèles *Xgboost* et *Deeplearning* sont différentes et nous pouvons constater que les taux de rachat moyens évoluent différemment. Le pouvoir explicatif des variables dans les modèles joue un rôle majeur dans la construction des lois.

Au cours de la projection, la loi de référence fournit un taux constant de 3,4% au delà de 16 ans d'ancienneté. Les modèles de *machine learning* effectuent leur apprentissage en prenant en compte, en plus de l'ancienneté, d'autres variables explicatives. Nous avons constaté que les taux de rachat en fonction de chacune des variables sont globalement supérieurs à 3,4%. Comme les modèles donnent des poids plus ou moins importants à ces variables, une fois que l'ancienneté des *models points* atteint 16 ans, les taux de rachat prédits sont donc nettement supérieurs à celui de la loi de référence.

La comparaison des taux de rachat moyens annuels entre les modèles *machine learning* et la loi de rachats en fonction de l'ancienneté a donc permis de constater qu'il y avait plus de prestations au titre du rachat structurel les premières années dans le cas des taux prédits (sauf dans le cas du *Deeplearning* les trois premières années de projection). En utilisant les lois *machine learning*, les prestations de rachats totales sont toutes inférieures à celles issues de la loi de référence. [3]

Les *Best Estimate* associés aux lois *machine learning* ont tous augmenté par rapport à celui obtenu avec la loi de référence. En effet, les prestations dans le cas des lois *machine*

3. Les chroniques des flux de prestations de rachat peuvent être consultées dans l'annexe C.

learning sont globalement toutes supérieures les premières années par rapport à celles issues de la loi de référence. Avec l'effet de l'actualisation, les prestations ont plus de poids en début de projection. Les SCR ont tous augmenté ce qui induit une diminution des ratios de solvabilité. De plus, on constate une baisse des PVFP qui s'explique de la même manière que dans la section précédente : un contexte économique favorable pour notre assureur qui est donc sensible à la hausse des rachats.

Nous remarquons que le taux de rachat moyen prédit par le modèle *Deeplearning* est supérieur à celui du *Gradient Boosting Machine* mais son *Best Estimate* est inférieur. Ceci s'explique par l'impact de l'actualisation sur les prestations de rachats structurels. En effet, on peut constater qu'en début de projection les taux de rachat prédits par le *Gradient Boosting Machine* sont supérieurs à ceux du *Deeplearning*.

L'utilisation des lois *machine learning* comme loi de rachats structurels aux dépends de la loi de rachats en fonction de l'ancienneté induit une baisse de la rentabilité et de la solvabilité de l'assureur.

Conclusion

Dans ce mémoire, nous avons proposé deux méthodes de modélisation des rachats structurels en alternative à l'utilisation de la loi de rachats en fonction de l'ancienneté du contrat. Chacune de ces lois a été ensuite appliquée pour estimer le ratio de couverture et la *Present Value of Future Profits* d'un assureur.

La première méthode a consisté à adapter la loi de rachats en fonction de l'ancienneté de l'assureur en utilisant une fonction paramétrique représentant le gain ou la perte pour l'assuré générée par la mise en oeuvre de la « Flat Tax ». En effet, avant 8 ans d'ancienneté du contrat, ce prélèvement forfaitaire unique est beaucoup plus favorable que les taux historiques. Il est donc probable que les comportements de rachat évoluent. La valeur des indicateurs économiques obtenus avec cette méthode est très proche des montants obtenus avec la loi en fonction de l'ancienneté.

La seconde méthode, dont l'objectif est d'affiner les lois de rachats structurels utilisées, repose sur l'utilisation des techniques de *machine learning*. Elles ont permis de prendre en compte d'autres facteurs explicatifs que l'ancienneté dans l'estimation des taux de rachat. Nous avons pu constater, grâce au pouvoir explicatif des variables, que l'ancienneté n'explique pas totalement la prédiction des taux de rachat. L'idée d'utiliser d'autres variables explicatives pour estimer le taux de rachat semble être pertinente.

Nous avons étudié quatre modèles d'apprentissage : trois méthodes ensemblistes et une méthode connexionniste. Le *Gradient Boosting Machine* s'avère être le modèle le plus performant en termes d'erreur quadratique moyenne. Les lois construites par les techniques de *machine learning* et la loi de rachats en fonction de l'ancienneté ont été confrontées sur la prédiction des taux de rachat sur l'année 2017. Nous avons constaté que la loi en fonction de l'ancienneté surestimait en moyenne les taux de rachat réels et que les lois *machine learning* proposaient une estimation des taux de rachat plus proche de la réalité pour l'année 2017. Lorsque nous avons effectué la projection du portefeuille, l'utilisation des lois *machine learning* a induit une baisse de la solvabilité et de la rentabilité de l'assureur.

L'approche *machine learning*, pour la construction de la loi de rachats structurels, semble donc être une bonne alternative à l'utilisation de la loi en fonction de l'ancienneté. Nous devons toutefois être vigilants sur la robustesse des résultats qui dépendent d'un nombre important d'hypothèses dans notre étude. Les techniques de *machine learning* présentent par ailleurs l'inconvénient de la difficulté d'interprétation des réponses. De plus, leur mise en oeuvre pourrait représenter un coût opérationnel important pour les assureurs et nécessite d'avoir une base d'apprentissage conséquente.

Notre étude scientifique présente quelques limites. En effet, certains facteurs explicatifs de la décision du rachat structurel (catégorie socio-professionnelle, le changement de situation de l'assuré, etc) n'ont pas été retenus pour des raisons d'indisponibilité des données. De plus, les données que nous avons utilisées pour estimer le taux de rachat structurel ne peuvent être complètement décorrélées de l'environnement économique. Eric Sammuel NANA NJOYA proposa dans son mémoire (2016) d'estimer la probabilité de rachat en prenant en compte, en plus des facteurs structurels, des facteurs économiques. Des travaux complémentaires à notre étude pourraient être mis en oeuvre pour prédire le comportement de rachat dynamique en prenant en compte la conjecture économique actuelle et estimer ainsi un taux de rachat global (au lieu de sommer les taux de rachat structurel prédits avec les taux conjoncturels fournis par la loi proposée par les ONC comme dans ce mémoire). De plus, notre étude pourrait être réutilisée dès lors qu'un historique de rachat avec la nouvelle fiscalité sera assez important pour appliquer les techniques de *machine learning*. Il serait donc intéressant de faire une étude sur les recettes fiscales générées par la nouvelle fiscalité et ainsi d'avoir une approche macro-économique de cette nouvelle réforme de la fiscalité de l'assurance vie.

Liste des abréviations, des sigles et des symboles

ACPR : Autorité de Contrôle Prudentiel et de Résolution

BE : *Best Estimate*

BSCR : *Basic Solvency Capital Requirement*

CoC : *Cost of Capital*

BSCR : *Basic Solvency Capital Requirement*

EIOPA : *European Insurance and Occupational Pensions Authority*

FP : Fonds Propres

GBM : *Gradient Boosting Machine*

GSE : Générateur de Scénarios Économiques

LTGA : *Long-Term-Guarantees-Assessment*

MCR : *Minimum Capital Requirement*

ONC : Orientations Nationales Complémentaires

PB : Participation aux Bénéfices

PFU : Prélèvement Forfaitaire Unique

PVFP : *Present Value of Future Profits*

RC : Rachat Conjoncturel

RMSE : *Root Mean Square Error*

RPG : Rétropropagation du Gradient

SCR : *Solvency Capital Requirement*

TME : Taux Mensuel des emprunts d'État

TMG : Taux Minimum Garanti

TS : Taux Servi

UC : Unité de compte

XGBoost : *Extreme Gradient Boosting*

Annexe A : Illustration des différents cas de rachats selon avant ou après la réforme

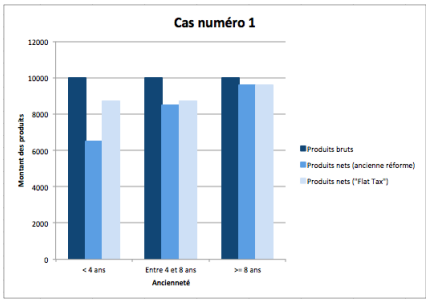

FIGURE 10.7 – Cas 1 : Nouveaux contrats

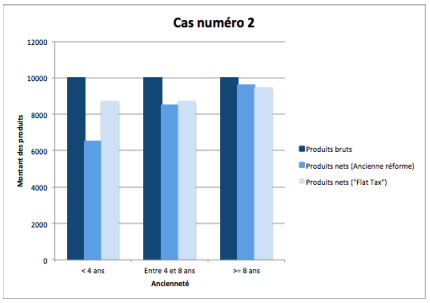

FIGURE 10.8 – Cas 2 : Nouveaux contrats

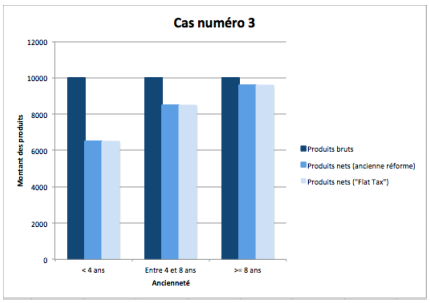

FIGURE 10.9 – Cas 3 : Anciens contrats

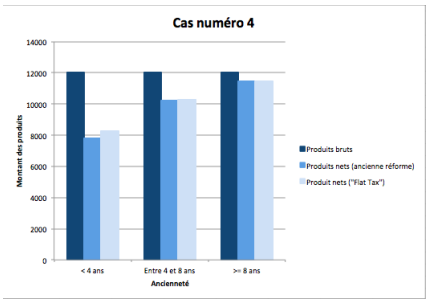

FIGURE 10.10 – Cas 4 : Anciens contrats

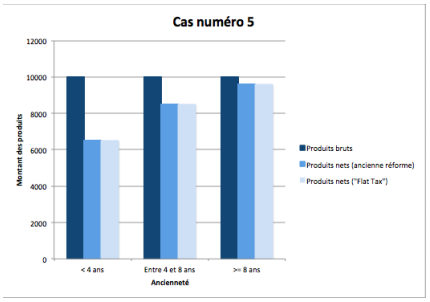

FIGURE 10.11 – Cas 5 : Anciens contrats

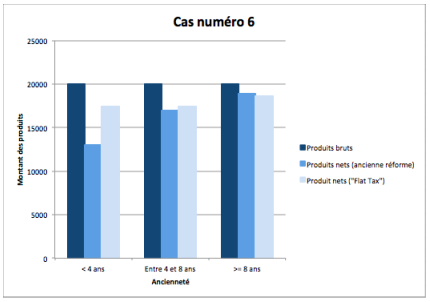

FIGURE 10.12 – Cas 6 : Anciens contrats

Annexe B : Matrices de corrélation pour les calculs de SCR

Matrice de corrélation des modules de SCR

	$SCR_{marché}$	SCR_{cred}	SCR_{vie}	$SCR_{santé}$	$SCR_{non-vie}$
$SCR_{marché}$	1				
SCR_{cred}	0,25	1			
SCR_{vie}	0,25	0,25	1		
$SCR_{santé}$	0,25	0,25	0,25	1	
$SCR_{non-vie}$	0,25	0,5	0	0	1

TABLE 10.3 – Matrice de corrélation des modules de risque composant le SCR

Matrice de corrélation des sous modules de SCR marché

	Taux	Action	Immobilier	*Spread*	Change	Concentration
Taux	1					
Action	0 / 0,5	1				
Immobilier	0 / 0,5	0,75	1			
Spread	0 / 0,5	0,75	0,5	1		
Change	0,25	0,25	0,25	0,25	1	
Concentration	0	0	0	0	0	1

TABLE 10.4 – Matrice de corrélation des sous-modules de risque composant le SCR marché

Matrice de corrélation des sous-modules de SCR souscription vie

	Mortalité	Longévité	Invalidité	Rachat	Frais	Révision	Cat
Mortalité	1						
Longévité	-0,25	1					
Invalidité	0,25	0	1				
Rachat	0	0,25	0	1			
Frais	0,25	0,25	0,5	0,5	1		
Révision	0	0,25	0	0	0,5	1	
Cat	0,25	0	0,25	0,25	0,25	0	1

TABLE 10.5 – Matrice de corrélation des sous-modules de risque composant le SCR souscription vie

Annexe C : La chronique des flux de prestations issus des lois *machine learning*

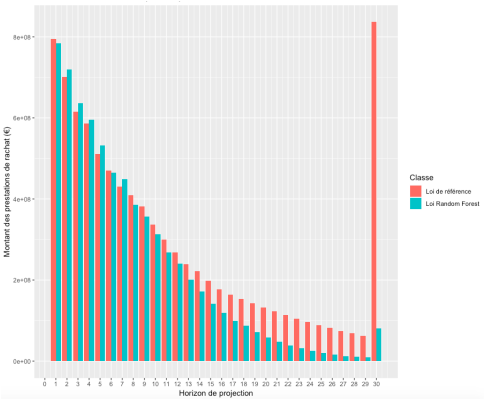

FIGURE 10.13 – Chonique des prestations RF

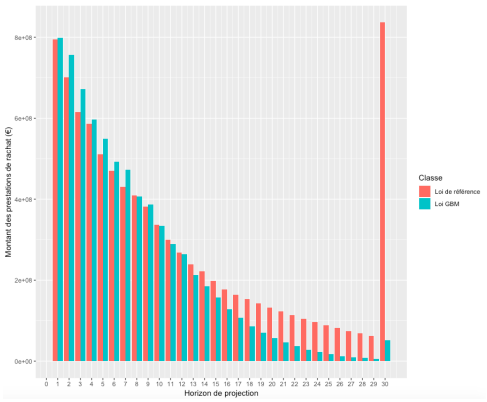

FIGURE 10.14 – Chronique des prestations GBM

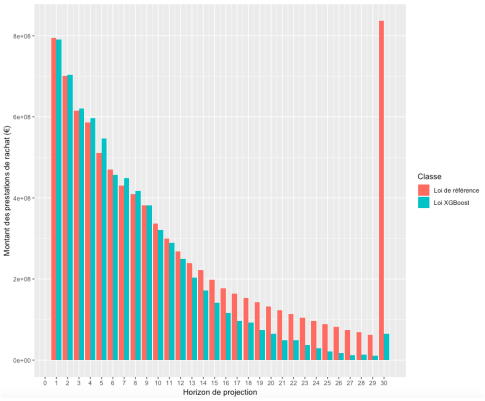

FIGURE 10.15 – Chronique des prestations XGBoost

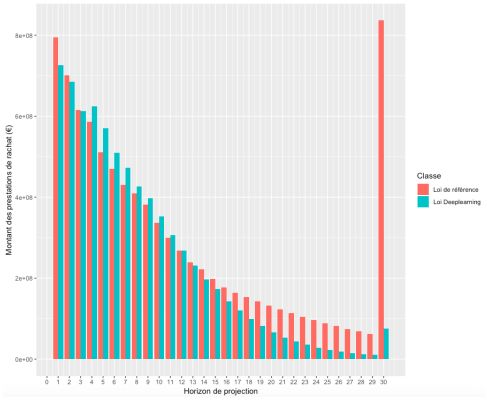

FIGURE 10.16 – Chronique des prestations Deeplearning

Bibliographie

[1] E. Biernat and M. Lutz. *Data Science : Fondamentaux et études de cas*, 2015.

[2] L. Breiman and J. Friedman. *Classification and Regression Trees*, 1984.

[3] CFO Forum. *Market Consistent Embedded Value Principles*, 2016.

[4] T. Chen and C. Guestrin. *XGBoost : A scalable Tree Boosting System*, 2016.

[5] Commission des finances. *Projet de Loi Finances 2018*, 2017.

[6] Laurent Devineau. *Agrégation des risques (Cours Euria Master 2)*, 2018.

[7] EIOPA. *Directive 2008/138/CE du Parlement Européen et du Conseil*, 2008.

[8] EIOPA. *Règlement délégué 2015/35*, 2015.

[9] FIXAGE. *Développement d'un modèle Actif-Passif stochastique*, 2017.

[10] FIXAGE. *Loi de Finances 2018 : la Flat Tax*, 2018.

[11] J. Friedman. *Greedy Function Approximation : A Gradient Boosting Machine*, 1999.

[12] Armel Kamal. *Générateur de scénarios économiques (Cours Euria Master 2)*, 2018.

[13] M. Malal. *La détection et la limitation des "écarts de convergence" dans un modèle de projection stochastique en univers risque neutre.*, 2016.

[14] X. Milhaud. *Segmentation et modélisation des comportements de rachat en Assurance Vie*, 2011.

[15] E. Nana. *Prédiction des comportements de rachat en épargne individuelle : une approche machine learning*, 2016.

[16] D. Pointin. *Les assureurs peuvent-ils continuer à vendre des contrats en euros ?*, 2017.

[17] Franck Vermet. *Apprentissage statistique : une approche connexionniste (Cours Euria Master 1)*, 2017.