

Mémoire présenté le :
pour l'obtention du diplôme
de Statisticien Mention Actuariat
et l'admission à l'Institut des Actuaires

Par : Hawa BALLO

Titre du mémoire :
**Impacts de la réforme des retraites de 2010 sur le
risque arrêt de travail et projection d'impacts**

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

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

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

signature

Entreprise : AXA FRANCE


Nom :

*Membres présents du jury de la
filiale*

Signature :

*Directeur de mémoire en
entreprise :*

Nom : LUZON Annaël

Signature : 

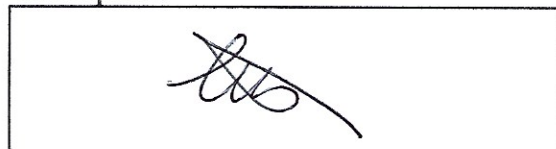
Invité :

Nom :

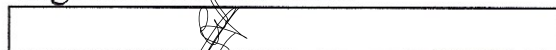
Signature :

*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



Résumé

Le suivi des garanties incapacité et invalidité couvertes par la prévoyance collective représente un enjeu de taille pour un grand assureur comme AXA. Ces risques sont sensibles à plusieurs facteurs externes (lois, réglementations, état de santé..). Au sein de la direction technique d'AXA France, plusieurs études ont été réalisées dans le but d'identifier et de calibrer ces facteurs externes. Cela permet d'avoir une meilleure connaissance du risque, donc d'adapter les paramètres de tarification afin de proposer une souscription adaptée à la réalité du risque. C'est dans cette optique que s'inscrivent les travaux menés dans le cadre de ce mémoire. En effet, le présent mémoire porte sur l'étude des impacts de la réforme des retraites de 2010 sur la garantie arrêt de travail (incluant l'invalidité). Cette réforme est à l'origine du décalage générationnel de l'âge de départ à la retraite de 2 ans de manière progressive. Elle entraîne mécaniquement un vieillissement de la population active, ce qui serait révélateur d'une hausse de la sinistralité en arrêt de travail. Tout au long de cette étude, il a été question de montrer, dans quelle(s) mesure(s) les effets de la réforme des retraites de 2010 sur la garantie arrêt de travail pourraient être mis en évidence. L'objectif sous-jacent est double : avoir une estimation des impacts de la réforme passée dans un premier temps puis à partir de cet impact pouvoir anticiper une future réforme qui se présenterait sous le même format. Compte tenu de l'âge moyen de départ à la retraite dans les pays de l'OCDE, du vieillissement de la population et de l'évolution des taux d'activité, il est attendu que l'âge de départ à la retraite subisse à nouveau un report dans les années à venir. Les travaux s'articulent au tour de cinq grandes parties. La première est la présentation des garanties de la prévoyance collective concernées par l'étude. Le système de retraite français est ensuite décrit à travers une

présentation de ses principes de base, un déroulé de l'historique des différentes réformes puis un focus sur la réforme de 2010 qui sera l'occasion de motiver le bien fondé de l'étude. La troisième partie consistera à décrire les données utilisées et les différents traitements et hypothèses ayant permis d'obtenir les statistiques descriptives. Cette partie est très importante puisqu'elle conditionne la suite. Elle permettra de confirmer la pertinence de la tentative d'analyse et de déceler des premiers éléments de réponse à la problématique. La démarche adoptée et les indicateurs utilisés pour mesurer l'impact de la réforme seront exposés dans le quatrième chapitre. Ce sera aussi l'occasion de présenter les projections des impacts d'une future réforme des retraites. Pour terminer, une attention particulière sera accordée à la population des assurés de plus de 58 ans dont les taux d'incidence feront l'objet d'une modélisation approfondie. Il sera question de mesurer l'effet de la réforme sur l'évolution des taux d'incidence à l'aide des modèles de séries temporelles. Le choix de ces modèles réside en leur capacité à décrire et prédire les séries indexées par le temps et perturbées par des événements spéciaux.

Mots clés : prévoyance collective, incapacité, invalidité, durée, provisionnement, retraite, réforme, âge de départ à la retraite, décalage, vieillissement, taux d'activité, classification, taux d'incidence, série temporelle, analyse d'intervention, perturbation, événements spéciaux

Abstract

The monitoring of disability and invalidity cover provided by the group provident scheme represents a major challenge for a leading insurer such as AXA. These risks are sensitive to several external factors (laws, regulations, state of health, etc.). AXA France's technical department has conducted several studies to identify and calibrate these external factors. This provides a better understanding of the risk and enables the pricing parameters to be adapted to propose an underwriting policy that is adapted to the reality of the risk. It is with this in mind that the work carried out in the context of this thesis takes place. Indeed, this paper focuses on the study of the impacts of the 2010 pension reform on the work stoppage guarantee (including disability). This reform is at the origin of the generational shift of the retirement age by 2 years in a progressive way. It mechanically leads to an aging of the working population, which would be indicative of an increase in the sick leaves claim rate. Throughout this study, the question was to show to what extent the effects of the 2010 pension reform on the sick leaves demand could be highlighted. The underlying objective is twofold : first, to estimate the impacts of the past reform and then, on the basis of this impact, to be able to anticipate a future reform in the same format. Given the average retirement age in OECD countries, population aging, and changes in participation rates, it is expected that the retirement age will be postponed again in the coming years. The work is divided into five main parts. The first is a presentation of the group pension benefits covered by the study. The French pension system is then described through a presentation of its basic principles, a review of the history of the various reforms and a focus on the 2010 reform, which will provide an opportunity to motivate the validity of the study. The third part will consist in describing the data

used and the different treatments and hypotheses that made it possible to obtain the descriptive statistics. This part is very important since it will determine the next steps. It will confirm the relevance of the attempted analysis and will allow us to detect the first elements of response to the problem. The approach adopted and the indicators used to measure the impact of the reform will be presented in the fourth chapter. It will also be an opportunity to present projections of the impacts of a future pension reform. Finally, the focus will be made on the population of insured persons over 58 years of age, whose incidence rates will be the subject of in-depth modeling. The effect of the reform on the evolution of incidence rates will be measured using time series models. The choice of these models lies in their ability to describe and predict time-indexed series that are affected by special events.

Keywords : sick leaves, disability, duration, reserving, retirement, reform, aging, participation rate, classification, incidence rate, time series, intervention analysis, disruption, special events

Synthèse

La réforme des retraites de 2010 est à l'origine du relèvement de deux ans de l'âge de départ à la retraite. Cette nouvelle mesure a favorisé l'allongement de carrière, ce qui est susceptible d'avoir des conséquences sur l'exposition de l'assureur aux risques incapacité et invalidité. L'objectif de ce mémoire est d'identifier les différents impacts de cette nouvelle réglementation sur la sinistralité du portefeuille AXA.

Les garanties incapacité et invalidité tiennent une place de taille dans les contrats collectifs de prévoyance. Elles permettent de compenser la diminution de revenus des salariés suite à un arrêt de travail à travers le versement d'indemnités journalières (incapacité) ou de pensions (invalidité). Le calcul et le versement du complément de revenus dépend de plusieurs paramètres : l'ancienneté dans l'entreprise, la franchise, le salaire, le niveau de couverture, la catégorie ou le taux d'invalidité, etc. Les facteurs démographiques influent énormément sur le suivi de ces risques.

La réforme des retraites par le biais du recul des âges réglementaires de départ à la retraite de 2 ans entraîne un vieillissement de la population active. Les salariés travaillent plus longtemps, par conséquent, la période de couverture des garanties incapacité et invalidité s'étend. A l'extension de la période de couverture, s'ajoute la dégradation de l'état de santé avec l'âge, ce qui justifie que l'on puisse émettre une telle hypothèse : l'allongement de l'âge de départ à la retraite entraînerait une hausse de la sinistralité. La première partie de ce mémoire reprend le fonctionnement général de la prévoyance collective. Le système de retraite français connaît depuis sa création en 1673, des évolutions réglementaires ayant permis de s'adapter aux besoins socio-économiques de la population française. La seconde partie retrace les différentes évolutions du système et présente la réforme des

retraites de 2010 car celle-ci se distingue de toutes les réformes qui l'ont précédée. Elle a suscité beaucoup d'attention depuis sa mise en place en 2011. Le recul de 2 ans s'est opéré de manière progressive, à raison de 5 mois par génération :

Année de naissance	ÂGE MINIMUM LÉGAL	ÂGE DU TAUX PLEIN
Avant 01/07/1951	60 ans	65 ans
1951	60 ans et 4 mois	65 ans et 4 mois
1952	60 ans et 9 mois	65 ans et 9 mois
1953	61 ans et 2 mois	66 ans et 2 mois
1954	61 ans et 7 mois	66 ans et 7 mois
À partir de 1955	62 ans	67 ans

Âges de départ à la retraite par génération suite à la réforme de 2010

Plusieurs études ont mis en lumière les conséquences sur l'activité des seniors, mais aussi les phénomènes dits de déversement vers d'autres dispositifs. En effet, DUBOIS et KOUBI ont montré que la réforme augmente l'offre de travail des seniors, le taux d'activité à 60 ans a augmenté de 24 points pour les hommes et 22 points pour les femmes. Les projections de population active de l'Insee entre 1990 et 2070 indiquent une croissance de la tranche des personnes de 60 ans et plus (18,9% en 1990 et 34,5% en 2070). La comparaison des taux d'activité en France et dans l'Union Européenne entre 2000 et 2017 révèle que le niveau d'activité en France est beaucoup plus faible que ce qui est observé dans les pays voisins. Cependant, quoique sensible, un rapprochement se crée entre les deux évolutions au fil des années, expliqué par l'augmentation progressive de l'âge de départ à la retraite en France. Ces résultats justifient l'utilité des travaux réalisés dans ce mémoire et leur apport à l'assurance prévoyance collective.

Les différentes sources de données et les traitements effectués figurent dans la troisième partie. Les travaux ont été réalisés avec les données issues de la base de provisionnement interne de AXA

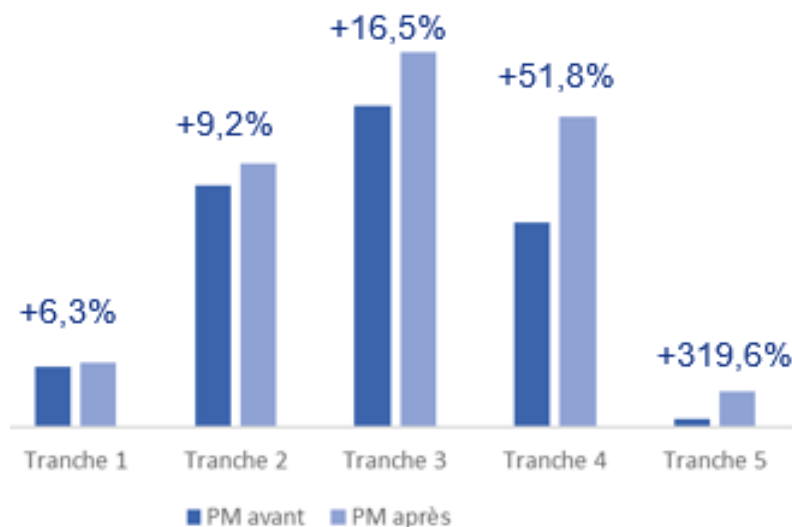
(base MRM) pour un historique de 9 années et les données issues des DSN¹ pour les survenances 2018 et 2019 uniquement compte tenu de la récence du dispositif. En complément aux variables de la base MRM, d'autres variables ont été créées ou corrigées pour répondre au mieux aux besoins de l'étude. C'est le cas de la variable "indemnité" qui a été ramenée à une base journalière car initialement exprimée dans des périodicités différentes selon les délégataires. La variable "sexe" par exemple est renseignée avec 4 codes, ce qui est aberrant. A l'aide du prénom de l'assuré et du fichier des prénoms de l'Etat civil, le sexe a été rétabli. Ce travail de définition et de correction des variables se fait en amont des statistiques descriptives et de la modélisation. Au terme de ce travail préparatoire, les statistiques du portefeuille de sinistres ont été produites. Plusieurs paramètres sont étudiés : la répartition du volume de sinistres, la répartition des âges à la survenance, l'évolution de la durée moyenne des arrêts et enfin les indemnités journalières. Les statistiques descriptives ont fourni des éléments de réponse concernant les conjectures émises. En effet, à partir de 2016, il a été constaté le pic de sinistralité le plus important entre deux survenances consécutives sur toute la période d'observation. Entre 2015 et 2016, 28% de sinistres supplémentaires ont été indemnisés. Cette hausse du volume d'arrêts est expliquée en majeure partie par une sur-représentation des personnes de 60 ans et plus constatée sur le portefeuille pendant la même période. La composition du portefeuille par âge révèle une majorité féminine, ce qui est en phase avec les taux d'absentéismes plus élevés chez les femmes d'après les baromètres AXA. Pour identifier d'éventuels impacts, l'évolution de la durée moyenne en fonction des années est d'une grande utilité. Pour ce faire, les durées restantes probables ont été calculées pour les sinistres encore en cours à la date d'extraction de la base (6% en incapacité et 48% en invalidité). Ainsi, les comparaisons entre survenances sont cohérentes et le volume de données utilisé reste satisfaisant. Dans le but d'améliorer les comparaisons des durées, des classes d'âges ont été créées, reflétant ainsi la structure des données et les besoins de l'étude. La durée moyenne en incapacité baisse au fur et à mesure des années, les seniors restent en arrêt maladie plus longtemps que la population globale. En invalidité, la durée moyenne des seniors est en croissance, 1,79 années en 2011 contre 2,71 années en 2019. Pour le reste

1. Déclaration Sociale Nominative

de la population, la durée moyenne est quasiment stable. La réforme des retraites est un facteur important de la hausse de la durée d'invalidité des seniors. L'évolution du montant de l'indemnité journalière moyenne par tranche d'âge montre que les seniors ont des indemnités plus élevées que le reste de la population, ce qui est intuitif car les salaires sont généralement plus élevés en fin de carrière. Le vieillissement de la population active implique à la fois une sur-fréquence et une montée en charge des arrêts des seniors.

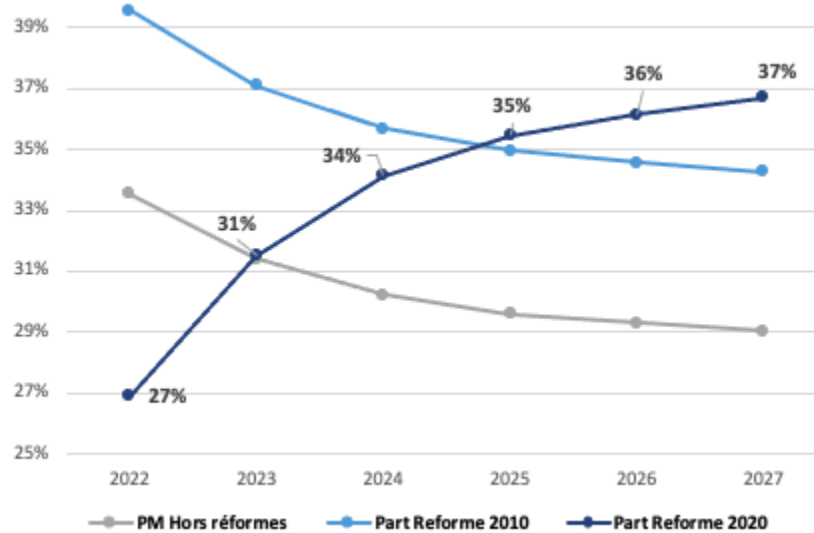
Le quatrième chapitre constitue le coeur du mémoire. Les différentes étapes et la méthodologie adoptée pour l'analyse des impacts de la réforme y sont recensés. La première étape consiste à déterminer, pour chaque assuré de la base, l'âge de départ à la retraite avant et après réforme. Suite à quoi deux catégories d'impacts se déclinent : les arrêts survenus à cause du prolongement de l'âge de départ à la retraite et les arrêts qui se sont allongés pour le même motif. Les impacts sont alors présentés pour chaque décalage et selon différents indicateurs. La réforme aura engendré 20,79% d'invalides additionnels et seulement 1,7% en incapacité. Le décalage de 2 ans englobe 91% des sinistres additionnels en invalidité et 71% en incapacité. En terme de durée, le supplément dû à la réforme est généralement contenu dans le décalage. Par exemple, en repoussant l'âge de départ de 4 mois, la durée additionnelle en invalidité est de 3.99 mois ce qui représente 99% du prolongement. La proportion du prolongement consommée décroît avec le décalage, variant de 28 à 72% en incapacité et 69 à 99% en invalidité. Les invalides impactés par la réforme durent en moyenne 28,5 mois tandis que ceux qui ne le sont pas durent 11 mois. La durée moyenne des seniors invalides étant de 21 mois, il est clair que cette durée est portée essentiellement par la réforme. La prestation moyenne versée au titre de l'invalidité est de 20 9698€ porté par les sinistres impactés par la réforme pour lesquels 28 349€ ont été versés contre seulement 11 104€ pour le reste de la population. En incapacité, le constat est le même : la réforme contribue à l'augmentation des indemnités versées. Les provisions mathématiques se sont alourdies suite à l'allongement de la période de couverture du risque invalidité. Toutes survenances confondues, la réforme aura entraîné un accroissement de 25% au niveau de la provision mathématique en invalidité. Le surplus est principalement porté par la tranche des 58 ans et plus pour laquelle

la PM après réforme représente 319,6% de la PM avant réforme. Le graphique suivant présente l'impact de la réforme sur la PM par tranche d'âge



PM avant/après réforme par tranches d'âge à la survenance

En 2019, un projet de loi visant à réformer le système de retraite a été initié. Le projet prévoit un report de l'âge de départ à la retraite de 2 ans sous le même schéma que la réforme de 2010. Une projection d'impacts a donc été faite pour une mise en place de la nouvelle réforme entre 2022 et 2027 comme suggéré par le rapport DELEVOYE. Les nouveaux sinistres inhérents à la nouvelle réglementation ont été simulés sur la base de ce qui a été observé suite à la réforme de 2010 en appliquant des correctifs liés à l'évolution démographique. C'est ce qui a permis d'établir un impact à +20,85% pour l'invalidité et 1,84% en incapacité. Sur le stock d'invalides, l'impact de la réforme de 2010 diminue progressivement, cela, compensé par la nouvelle réforme dont l'effet croît en fonction des années.



Évolution de l'impact de chaque réforme sur la PM du stock de sinistres en invalidité

Le dernier chapitre porte sur la modélisation des taux d'incidence des seniors. Dans cette partie, l'objectif principal est de décrire et prédire les taux d'incidence des seniors à l'aide de séries temporelles. Les modèles de séries temporelles sont très efficaces pour étudier des données indexées par le temps. Par ailleurs, ils permettent aussi de prendre en compte l'intervention d'événements spéciaux, ce qui en fait un choix privilégié pour notre étude. La modélisation des processus ARMA n'est possible que sur des séries stationnaires. La série d'incidence présente une tendance et une saisonnalité, par conséquent elle n'est pas stationnaire. La première étape de la modélisation consiste alors à stationnariser la série en appliquant la méthode de différenciation. S'en suivent les étapes d'identification, d'estimation, de validation du modèle et enfin les prévisions. L'analyse de l'intervention consiste à introduire dans le modèle ARMA classique une variable exogène de type indicatrice qui représente la perturbation que subit la série. L'indicatrice prend la forme d'un saut ou d'une impulsion. Les notions de filtres linéaires et de fonctions d'intervention ont été présentées. Dans le cadre de la réforme, la perturbation est intervenue en janvier 2016, qui correspond à l'entrée en vigueur du recul de 2 ans (ce décalage concerne près de 97% des personnes en arrêt maladie). La fonction de transfert choisie a la forme d'une courbe en croissance exponentielle, à l'image de l'évolution constatée du volume de sinistres. Suite à l'introduction de l'intervention les

qualités prédictives du modèle se sont améliorées, ce qui témoigne du bénéfice de cette méthode. Finalement, la réforme aura impacté les taux d'incidence à hauteur de 21,62% à la hausse entre 2016 et 2018. Les prévisions sur le long terme indiquent un gain de 10 points sur les taux d'incidence d'ici 2030.

Synthesis

The 2010 pension reform led to a two-year increase in the retirement age. This new measure has encouraged career lengthening, which is likely to have consequences on the insurer's exposure to disability and invalidity risks. The purpose of this memorandum is to identify the various impacts of this new regulation on the claims experience of the AXA portfolio.

Disability and invalidity cover plays a major role in group sick leaves insurance contracts. They make it possible to compensate for the reduction in employee income following a work stoppage through the payment of daily allowances (disability) or pensions (invalidity). The calculation and payment of the supplementary income depends on several parameters : seniority in the company, deductible, salary, level of coverage, category or rate of disability, etc. Demographic factors greatly influence the monitoring of these risks. Pension reform through the lowering of the statutory retirement ages by 2 years is leading to an ageing of the working population. Employees are working longer, and as a result, the coverage period for disability and invalidity benefits is extending. In addition to the extension of the coverage period, the deterioration in health status with age justifies the hypothesis that a longer retirement age would lead to an increase in claims. The first part of this dissertation looks at the general functioning of group provident schemes. Since its creation in 1673, the French pension system has undergone regulatory changes that have made it possible to adapt to the socio-economic needs of the French population. The second part retraces the different evolutions of the system and presents the 2010 pension reform, as it differs from all the reforms that preceded it. It has attracted a great deal of attention since its implementation in 2011. The 2-year rollback has been gradual, at a rate of 5 months per generation :

Année de naissance	ÂGE MINIMUM LÉGAL	ÂGE DU TAUX PLEIN
Avant 01/07/1951	60 ans	65 ans
1951	60 ans et 4 mois	65 ans et 4 mois
1952	60 ans et 9 mois	65 ans et 9 mois
1953	61 ans et 2 mois	66 ans et 2 mois
1954	61 ans et 7 mois	66 ans et 7 mois
À partir de 1955	62 ans	67 ans

Âges de départ à la retraite par génération suite à la réforme de 2010

Several studies have highlighted the consequences on the activity of senior citizens, but also the phenomena known as spillage to other devices. Indeed, DUBOIS et KOUBI have shown that the reform increases the labour supply of seniors, the activity rate at 60 years old has increased by 24 points for men and 22 points for women. INSEE projections of the active population between 1990 and 2070 indicate a growth in the 60+ age group (18.9% in 1990 and 34.5% in 2070). A comparison of activity rates in France and the European Union between 2000 and 2017 reveals that the level of activity in France is much lower than that observed in neighboring countries. However, although noticeable, a rapprochement between the two evolutions is being created over the years, explained by the gradual increase in the retirement age in France. These results justify the usefulness of the work carried out in this paper and its contribution to group provident insurance.

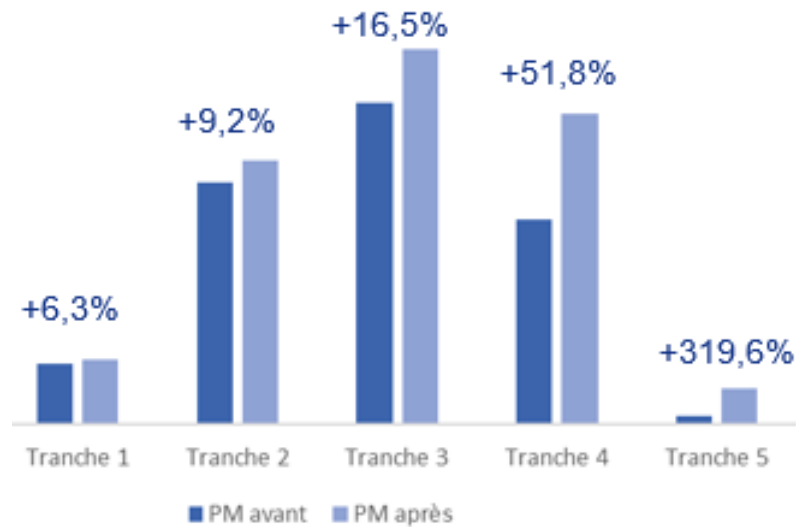
The various data sources and the processing carried out are presented in the third part. The work was carried out using data from AXA's internal provisioning database (MRM database) for a 9-year history and data from the DSN² for the 2018 and 2019 events only, given the recency of the scheme. In addition to the variables in the MRM database, other variables have been created or corrected to best meet the needs of the study. have been corrected. This is the case of the "indemnity" variable, which was reduced to a daily basis because it was initially expressed at different intervals depending on the delegates. The "sex" variable, for example, is filled in with

2. Déclaration Sociale Nominative

4 codes, which is absurd. With the help of the first name of the insured and the civil status file, the sex has been restored. This work of definition and correction of variables is done upstream of descriptive statistics and modeling. At the end of this preparatory work, the statistics of the claims portfolio were produced. Several parameters were studied : the distribution of the volume of claims, the distribution of ages at occurrence, the evolution of the average duration of stoppages and finally the daily allowances. Descriptive statistics provided elements of response concerning the conjectures made. Indeed, from 2016 onwards, the highest peak in claims between two consecutive occurrences was observed over the entire observation period. Between 2015 and 2016, 28% more claims were compensated. This increase in the volume of stoppages is mainly explained by an over-representation of people aged 60 and over in the portfolio during the same period. The composition of the portfolio by age reveals a majority of women, which is in line with the higher absenteeism rates among women according to AXA's barometers. In order to identify possible impacts, the change in the average length of time over the years is very useful. To do this, the probable remaining durations were calculated for claims still in progress at the date of extraction from the database (6% incapacity and 48% disability). Thus, comparisons between occurrences are consistent and the volume of data used remains satisfactory. In order to improve comparisons of durations, age classes have been created, reflecting the data structure and the needs of the study. The average duration of disability decreases over the years, with older people remaining on sick leave longer than the overall population. In terms of disability, the average length of time seniors are on disability is increasing, 1,79 years in 2011 versus 2,71 years in 2019. For the rest of the population, the average duration is virtually stable. Pension reform is a major factor in the increase in the length of disability among seniors. The change in the average daily allowance per age group shows that senior citizens have higher allowances than the rest of the population, which is intuitive as wages are generally higher at the end of their careers. The aging of the working population implies both an over-frequency and an increase in the number of sick leaves demands among seniors.

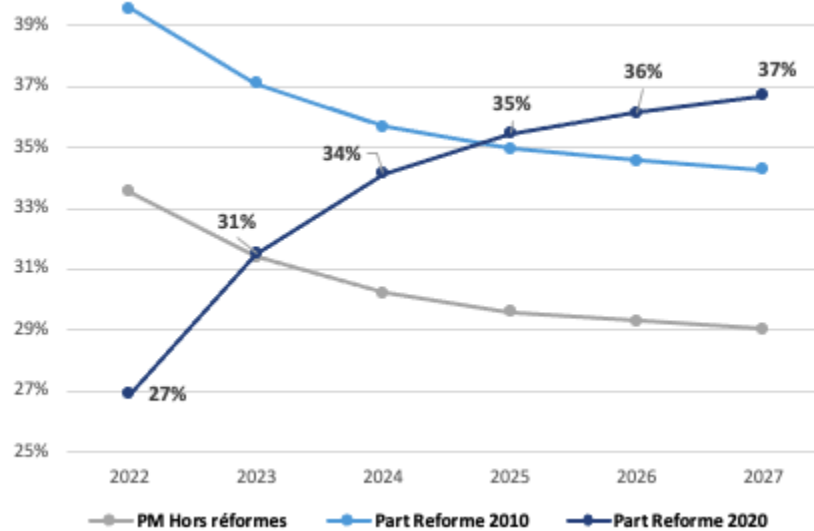
The fourth chapter is the heart of the brief. It lists the various stages and the methodology adopted for analyzing the impacts of the reform. The first step is to determine the pre- and post-

reform retirement age for each insured person of the data base. Then two categories of impacts are identified : sick leaves that occurred because of the extension of the retirement age and those that were extended for the same reason. Impacts are then presented for each lag and according to different indicators. The reform will have resulted in 20,79% additional disability claims and only 1,7% incapacity. The 2-year time lag covers 91% of additional disability claims and 71% of incapacity claims. In terms of duration, the supplement due to the reform is generally contained in the time lag. For example, by postponing the age of departure by four months, the additional duration of disability is 3.99 months, which represents 99% of the extension. The proportion of the extension consumed decreases with the time lag, ranging from 28 to 72% on disability and 69 to 99 % on incapacity. Disabled people impacted by the reform last an average of 28.5 months, while those who are not affected by the reform last 11 months. As the average duration of disabled seniors is 21 months, it is clear that this duration is essentially due to the reform. The average benefit paid for disability is 20,9698 € carried by the claims impacted by the reform for which 28,349 € were paid compared to only 11,104 € for the rest of the population. For those on incapacity, the finding is the same : the reform contributes to the increase in the compensation paid out. Mathematical reserves have increased following the extension of the period of coverage of the disability risk. All years taken together, the reform have led to an increase of 25% in the mathematical provision for disability. The surplus is mainly carried by the 58 years old and over bracket for which the MP after reform represents 319.6% of the MP before reform. The following graph shows the impact of the reform on the MP by age group :



PM before/after the reform by age brackets at occurrence

In 2019, a bill to reform the pension system was initiated. The bill provides for a 2-year postponement of the retirement age under the same scheme as the 2010 reform. An impact projection has therefore been made for the implementation of the new reform between 2022 and 2027 as suggested by the report DELEVOYE. The new claims inherent in the new regulations were simulated on the basis of what was observed following the 2010 reform by applying corrections linked to demographic changes. This resulted in an impact of +20,85% for disability and 1,84% for incapacity. On the stock of disabled people, the impact of the 2010 reform is gradually decreasing, which is compensated by the new reform whose effect increases with the years.



Evolution of the impact of each reform on the MPof the disability claims stock

The last chapter deals with the modeling of incidence rates for seniors. In this section, the main objective is to describe and predict senior citizen incidence rates using time series. Time series models are very effective for studying time-indexed data. Moreover, they also allow the intervention of special events to be taken into account, making them a preferred choice for our study. ARMA process modeling is only possible on stationary series. The incidence series shows a trend and seasonality, therefore it is not stationary. The first step of the modeling consists then in stationaryizing the series by applying the differentiation method. This is followed by identification, estimation, model validation and finally forecasting. The analysis of the intervention consists in introducing into the classical ARMA model an exogenous variable of indicator type which represents the perturbation that the series undergoes. The indicator takes the form of a jump or an impulse. The notions of linear filters and intervention functions were presented. Within the framework of the reform, the disturbance occurred in January 2016, which corresponds to the entry into force of the 2-year delay (this delay concerns nearly 97% of people on sick leave). The transfer function chosen has the shape of an exponentially growing curve, reflecting the observed change in the volume of claims. Following the introduction of the intervention, the predictive qualities of the model have improved, which testifies to the benefits of this method. Finally, the reform have impacted

incidence rates by 21,62% between 2016 and 2018. Long-term forecasts indicate a 10-point gain in incidence rates by 2030.

Remerciements

Je tiens en premier lieu à remercier toute l'équipe de la Direction Technique et Innovation de AXA France pour leur accueil et leur accompagnement tout au long de mon apprentissage. Je remercie chaleureusement ma tutrice en entreprise, Annael Luzon, pour son encadrement, ses nombreux conseils et le temps qu'elle m'a consacré.

Je souhaiterais exprimer ma reconnaissance à Catalina Sepulveda et Jean-Charles Martin-Vivier pour m'avoir fait confiance en me permettant d'effectuer mon alternance au sein de l'équipe Business Support France Etudes. Je tiens également à remercier Delphine Clot et El Hadj Ndiaye pour le temps, l'écoute et leurs précieux conseils.

J'adresse mes sincères remerciements à Olivier Lopez, mon tuteur pédagogique, pour son encadrement et la pertinence de ses conseils tout au long des travaux du mémoire. Je remercie également l'ensemble du corps éducatif de l'ISUP pour la qualité des enseignements transmis et leur sollicitude. En particulier, Olivier Wintenberger, professeur de séries temporelles à l'ISUP pour m'avoir encouragée et accompagnée dans l'exploration de ce sujet.

Je ne saurais terminer sans remercier, du fond du coeur, ceux qui m'ont toujours soutenue tout au long de ma scolarité, mes parents, mon frère et ma soeur ; ce sont des personnes qui ont cru en moi et pour lesquelles j'ai toujours voulu donner le meilleur de moi. Mention spéciale à mon défunt père qui n'a ménagé aucun effort pour la réussite de mes études.

A tous ceux et celles qui, de près ou de loin, ont suivi avec intérêt ce sujet de mémoire, j'adresse mes remerciements.

Table des matières

	Page
1 La prévoyance collective	23
1.1 Incapacité	24
1.1.1 Définition	24
1.1.2 Prestations de la Sécurité Sociale	24
1.2 Invalidité	26
1.2.1 Définition	26
1.2.2 Prestations de la Sécurité Sociale	27
1.3 Prestations de l'employeur	30
2 La retraite en France	32
2.1 Présentation et évolution du système de retraite	32
2.1.1 Présentation du régime Français	32
2.2 Réforme 2010	36
2.2.1 Présentation de la réforme	36
2.2.2 Enjeux pour la prévoyance collective	39
2.2.3 Interaction avec d'autres réformes	45
3 Présentation des données et statistiques descriptives	47
3.1 Présentation des bases de données	47
3.2 Traitement des données	49
3.2.1 Périmètre de l'étude et hypothèses	49
3.2.2 Variables créées	50
3.2.3 Variables corrigées	52
3.3 Statistiques descriptives	53
3.3.1 Répartition du volume de sinistres	53
3.3.2 Répartition des âges à la survenance	56
3.3.3 Durée des arrêts	57
3.3.4 Indemnité journalière	63

4	Impacts de la réforme sur la sinistralité du portefeuille	64
4.1	Construction de la base d'analyse et hypothèses	64
4.1.1	Détermination de l'âge de départ à la retraite	65
4.1.2	Identification des arrêts impactés par la réforme des retraites	67
4.2	Présentation des résultats	69
4.2.1	Impacts sur le volume	69
4.2.2	Impacts sur les durées	71
4.2.3	Impacts sur les prestations versées	73
4.3	Provisions mathématiques	75
4.3.1	Exigences réglementaires	75
4.3.2	Calcul des provisions mathématiques	78
4.3.3	Effets du décalage d'âge de départ à la retraite sur la PM en invalidité	80
4.3.4	Projet de réforme 2020	84
5	Modélisation des taux d'incidence	88
5.1	Choix de la modélisation	88
5.1.1	Quantité modélisée	88
5.1.2	Modèles utilisés	91
5.2	Les processus ARMA	92
5.2.1	Rappels mathématiques	92
5.2.2	Mise en oeuvre de la modélisation	96
5.3	Analyse d'intervention	101
5.3.1	Filtres linéaires	101
5.3.2	Présentation du modèle d'intervention	105
5.3.3	Fonctions d'intervention type	107
5.4	Application aux séries d'incidence	110
5.4.1	Résultats de la modélisation ARMA	110
5.4.2	Résultats de la modélisation ARMA avec intervention	112
5.4.3	Conclusions et limites de la modélisation	117
Appendices		
A	Algorithme de Jenks	129
B	Statistiques descriptives	131
B.0.1	Volume	131
B.0.2	Durées moyennes	131
B.0.3	Passage de l'incapacité vers l'invalidité	132

C	Analyse des résidus	134
C.1	Tests de bruit blanc	134
C.1.1	Fonction d'autocorrélation empirique (ACF)	134
C.1.2	Test du Porte-manteau	135
C.2	Tests de normalité des résidus	135
	Bibliographie	138

Chapitre 1

La prévoyance collective

Le régime complémentaire est le régime qui vient, comme son nom l'indique, en complément du régime de base de la Sécurité Sociale. Il est destiné à renforcer les garanties de la SS qui ne sont pas toujours suffisantes. A ce titre, les régimes de prévoyance complémentaires peuvent couvrir l'arrêt de travail. L'arrêt de travail est une période pendant laquelle le salarié se voit dans l'impossibilité d'exercer son activité professionnelle à la suite d'une maladie ou d'un accident de la vie privée, d'un accident du travail ou d'une maladie professionnelle. Il doit être prescrit par un médecin qui aura constaté cette incapacité. Un assuré est dit en état d'incapacité de travail lorsque, par suite de maladie ou d'accident, il est contraint d'interrompre totalement ou partiellement son activité. Il en résulte une diminution de gain ou de salaire à laquelle la sécurité sociale et le régime de prévoyance complémentaire se proposent de remédier. L'incapacité peut être constatée par le médecin traitant mais également par la sage-femme dans sa limite de compétence professionnelle et pour une durée fixée par décret. Le terme "incapacité de travail" recouvre deux états : l'incapacité temporaire et l'incapacité permanente.

1.1 Incapacité

1.1.1 Définition

On parle d'incapacité, ou d'incapacité temporaire, lorsque l'arrêt de travail n'est pas définitif. On peut distinguer l'incapacité temporaire totale, durant laquelle le salarié est dans l'impossibilité absolue, complète et continue de travailler, de l'incapacité temporaire partielle, durant laquelle le salarié peut reprendre partiellement son activité avec accord de la Sécurité Sociale. Dans les deux cas, la sécurité sociale verse des indemnités journalières (réduites pour l'incapacité temporaire partielle), après un délai de carence de trois jours. Ce délai permet de réduire le risque d'aléa moral, qui veut que les personnes modifient leur comportement dès lors qu'elles se savent assurées. Dans le cas de l'arrêt de travail, il permet donc de dissuader les assurés de recourir à un arrêt de complaisance. La durée maximale de l'état d'incapacité est de trois ans, après quoi l'assuré peut être classé en invalidité s'il est toujours dans l'impossibilité d'exercer son activité.

1.1.2 Prestations de la Sécurité Sociale

Arrêt de travail résultant d'une maladie ou d'un accident de la vie privée

L'assuré est indemnisé sous la forme d'indemnités journalières (IJ), à la condition qu'il exerce une activité professionnelle au moment de son arrêt. De plus, l'interruption de l'activité professionnelle doit être constatée par la prescription d'un arrêt de travail signée par un médecin, et peut faire l'objet d'un contrôle. Pour être indemnisé pendant les six premiers mois de son arrêt de travail, l'assuré doit remplir l'une des deux conditions suivantes :

- son salaire total sur les six mois précédant l'arrêt doit être supérieur ou égal à 1 015 fois le SMIC horaire¹ ;
- son temps total de travail sur les trois mois civils précédant l'arrêt doit être supérieur ou égal à 200 heures.

1. Le SMIC (salaire horaire minimum légal) en 2020 est de 10,15 €

Pour être indemnisé au-delà des six premiers mois d'arrêt de travail, l'assuré doit être salarié depuis au moins un an au moment de son arrêt de travail, et remplir l'une des deux conditions suivantes :

- son salaire total sur les douze mois précédant l'arrêt doit être supérieur ou égal à 2 030 fois le SMIC horaire, dont au moins 1 015 fois le SMIC horaire durant les six premiers mois ;
- son temps total de travail sur les douze mois civils précédant l'arrêt doit être supérieur ou égal à 800 heures, dont au moins 200 heures durant les trois premiers mois.

Le calcul de l'indemnité journalière fait appel à la notion de tranches de salaires définies par la Sécurité Sociale. La première tranche TA correspond à la part du salaire comprise entre 1 et une fois le PMSS². Le montant de l'indemnité journalière est alors égal à :

$$IJ = \text{BASE} \times 50\% \times \frac{1}{90},$$

La *base* correspond à la tranche A du salaire moyen des 3 mois précédents l'arrêt. Le montant de l'IJ est donc limité au PASS³ / 720, soit 57,13 € en 2020. Si l'assuré possède au moins trois enfants à charge, ce montant est majoré à compter du 31^e jour d'arrêt :

$$IJ = \text{BASE} \times \frac{2}{3} \times \frac{1}{90},$$

montant limité au PASS / 540, soit 76,17 € pour 2020. Enfin, une revalorisation est prévue lorsque la durée de l'arrêt dépasse trois mois.

Les indemnités journalières perçues sont soumises à la CSG¹ au taux de 6,2 % et à la CRDS² au taux de 0,5 %, ainsi qu'à l'impôt sur le revenu, et valident les droits à la retraite.

Arrêt de travail résultant d'un accident de travail ou d'une maladie professionnelle

Contrairement aux maladies et accidents de la vie privée, il n'y a ici aucun délai de carence, et les indemnités sont versées au 1^{er} jour d'arrêt. Le risque d'aléa moral est ici quasiment nul, car

2. Le PMSS (Plafond Mensuel de Sécurité Sociale) est de 3428 €

3. Le PASS (Plafond Annuel de Sécurité Sociale) est de 41 136 €

l'accident de travail ou la maladie professionnelle doit être constatée par un médecin du travail. L'assuré en incapacité reçoit des indemnités journalières, dont le montant dépend de son salaire journalier de base. Ce dernier est égal à 1/30 du dernier salaire mensuel brut, limité à 0,834% du PASS, soit 343,07 € pour 2020.

Durant les 28 premiers jours d'arrêt, l'assuré perçoit une indemnité journalière d'un montant égal à 60% de son salaire journalier. Après cette période, le montant de l'indemnité passe à 80% du salaire journalier.

Comme dans le cas précédent, les indemnités journalières perçues sont soumises à la CSG au taux de 6,2 % et à la CRDS au taux de 0,5 %, et valident les droits à la retraite, mais ne sont pas soumises à l'impôt sur le revenu.

Tableau récapitulatif des prestations d'incapacité

MOTIF DE L'ARRÊT	CONTEXTE PROFESSIONNEL	CONTEXTE NON PROFESSIONNEL
DÉLAI DE CARENCE		3 jours
INDEMNITÉS (EN % DU SALAIRE)	<ul style="list-style-type: none"> • Du 1^{er} jour au 28^e jour : 60% • Au delà de 28 jours : 80% 	<ul style="list-style-type: none"> • Du 1^{er} jour au 30^e jour : 50% • Au delà de 30 jours : 50%, 66.66% si plus de 3 enfants

1.2 Invalidité

1.2.1 Définition

On parle d'invalidité, ou d'incapacité permanente, lorsque la maladie ou l'accident qui a occasionné l'arrêt de travail entraîne une réduction de la capacité de travail. L'état d'invalidité est apprécié soit après consolidation ou stabilisation de l'état de santé, soit à l'expiration de la période d'invalidité. Dans le cas des maladies et accidents de la vie privée, la Sécurité Sociale définit trois

types d'invalidité :

CATÉGORIE	DÉFINITION
1 ^{ère} catégorie	Invalides capables d'exercer une activité rémunérée
2 ^{ème} catégorie	Invalides absolument incapables d'exercer une activité professionnelle quelconque
3 ^{ème} catégorie	Invalides incapables d'exercer une profession et dans l'obligation de recourir à l'assistance d'une tierce personne pour effectuer les actes ordinaires de la vie

Dans le cas des maladies professionnelles et des accidents du travail, la Caisse d'Assurance Maladie détermine un taux d'incapacité permanente, après avis du médecin conseil. Ce taux dépend de la nature de l'infirmité, de l'état général, de l'âge, des facultés physiques et mentales, et des aptitudes et qualifications professionnelles de l'assuré. Le montant de la pension d'invalidité versée par la Sécurité dépend de ce taux.

1.2.2 Prestations de la Sécurité Sociale

Arrêt de travail résultant d'une maladie ou d'un accident de la vie privée

L'assuré est indemnisé sous la forme d'une rente versée mensuellement à terme échu, jusqu'à ses 60 ans, la pension de vieillesse prenant le relais à la date anniversaire. Les conditions d'ouverture des droits sont les suivantes :

- Une capacité de travail réduite d'au moins des 2/3,
- Une ancienneté au moment de l'arrêt de travail d'au moins un an,
- Un salaire total sur les douze mois précédant l'arrêt supérieur ou égal à 2 030 fois le SMIC horaire, dont au moins 1 015 fois le SMIC horaire durant les six premiers mois.

Le montant de la rente correspond au produit du taux de pension et du salaire de base. Le taux de pension dépend de la catégorie d'invalidité, et le salaire de base mensuel est égal à 1/12 de la moyenne des dix meilleures tranches TA des salaires annuels.

CATÉGORIE	DÉFINITION	MONTANT MENSUEL MINIMUM EN 2020	MONTANT MENSUEL MAXIMUM EN 2020
1 ^{ère} catégorie	30% du salaire de base	292,8 €	1 028,40 €
2 ^{ème} catégorie	50% du salaire de base	292,8 €	1 714 €
3 ^{ème} catégorie	50% du salaire de base + majoration forfaitaire pour tierce personne	292,8 € + 1 125,29 €	1 714 € 1 125,29 €

La rente perçue est revalorisée chaque année au 1^{er} janvier, soumise à la CSG au taux de 6,6 % et à la CRDS au taux de 0,5 %, ainsi qu'à l'impôt sur le revenu, à l'exception de la majoration forfaitaire pour tierce personne. En cas de modification de l'état d'invalidité, la pension peut être revalorisée et l'assuré peut être classé dans une autre catégorie.

Arrêt de travail résultant d'un accident de travail ou d'une maladie professionnelle

Si le taux d'incapacité permanente de l'assuré est inférieur à 10%, il perçoit une indemnité forfaitaire en capital, dont le montant varie selon la valeur de ce taux, comme présenté dans le tableau ci-après. Cette indemnité n'est soumise ni à la CSG, ni à la CRDS, ni à l'impôt sur le revenu.

TAUX D'INVALIDITÉ	MONTANT DE L'INDEMNITÉ
1%	418,96 €
2%	680,96 €
3%	995,08 €
4%	1 570,57 €
5%	1 989,64 €
6%	2 460,85 €
7%	2 984,21 €
8%	3 560,36 €
9%	4 188,63 €

Source : *Service Publique*

Si le taux d'incapacité permanente de l'assuré est supérieur ou égal à 10%, il perçoit une rente annuelle viagère, dont le montant est égal au produit du salaire annuel de base par un taux de rente. Le salaire annuel de base est égal à la somme des douze salaires mensuels précédant l'arrêt. Le taux de la rente est égal au taux d'incapacité permanente réduit de moitié pour la partie de ce taux ne dépassant pas 50% et augmenté de moitié pour la partie supérieure à 50%. Illustrons ceci par deux exemples :

- Si le taux d'incapacité permanente est égal à 30%, le taux de la rente sera :

$$30\% \times \frac{1}{2} = 15\%$$

- Si le taux d'incapacité permanente est égal à 70%, le taux de la rente sera :

$$50\% \times \frac{1}{2} + (70\% - 50\%) \times \frac{3}{2} = 55\%$$

Cette rente, dont le montant est revalorisé chaque année au 1^{er} janvier, n'est soumise ni à la CSG, ni à la CRDS, ni à l'impôt sur le revenu. En outre, dans le cas où le taux d'incapacité

dépasse 80% et oblige l'assuré à recourir à l'assistance d'une tierce personne pour effectuer les actes ordinaires de la vie, le montant de la rente est majoré de 40%. Le montant de la rente est susceptible de varier de deux façons : la première, fondée sur des motifs économiques, se traduit par la revalorisation périodique des rentes ; la seconde, basée sur la constatation d'une modification de l'état d'incapacité de l'assuré, donne lieu à l'engagement de la procédure de révision qui peut aboutir, soit au maintien de la rente, soit à la suppression de celle-ci, soit à la modification de son taux.

Tableau récapitulatif des prestations d'invalidité

MOTIF DE L'ARRÊT	CONTEXTE PROFESSIONNEL	CONTEXTE NON PROFESSIONNEL
FACTEUR DÉTERMINANT	Taux d'invalidité	Catégorie d'invalidité
PRESTATION	<ul style="list-style-type: none"> • < 10% : montant forfaitaire • ≥ 10% : rente 	<ul style="list-style-type: none"> • Cat 1 : 30% du salaire • Cat 2 : 50% du salaire • Cat 3 : 50% du salaire + majoration

1.3 Prestations de l'employeur

L'Accord National Interprofessionnel de Mensualisation du 10 décembre 1977, obligatoire pour tous les salariés, prévoit un complément d'indemnisation par l'employeur. Les risques couverts par cet accord sont la maladie et l'accident, qu'ils soient de nature professionnelle ou non. L'assuré doit réunir quatre conditions pour pouvoir prétendre à percevoir ces prestations :

- avoir au moins trois ans d'ancienneté dans l'entreprise,
- fournir un certificat médical dans les 48 heures suivant l'arrêt,
- être pris en charge par la Sécurité Sociale,

- effectuer ses soins en France ou dans l'Union Européenne

Les prestations sont versées dès le premier jour d'arrêt en cas d'accident du travail ou de maladie professionnelle, alors que dans les autres cas, un délai de carence de dix jours est appliqué. La durée d'indemnisation dépend de l'ancienneté de l'assuré dans l'entreprise au moment de son arrêt, et est divisée en deux parties d'égales durées : une première durant laquelle l'employeur complète l'indemnisation de la Sécurité Sociale à hauteur de 90% du salaire brut, et une seconde à hauteur de 2/3 du salaire brut. Les durées d'indemnisations sont synthétisées dans le tableau suivant :

ANCIENNETÉ	90% DU SALAIRE BRUT	2/3 DU SALAIRE BRUT
Entre 3 et 8 ans	30 jours	30 jours
Entre 8 et 13 ans	40 jours	40 jours
Entre 13 et 18 ans	50 jours	50 jours
Entre 18 et 23 ans	60 jours	60 jours
Entre 23 et 28 ans	70 jours	70 jours
Entre 28 et 33 ans	80 jours	80 jours
33 ans et plus	90 jours	90 jours

Les prestations versées par la Sécurité Sociale et l'employeur étant faibles, les assurés se retrouvent avec des ressources d'un niveau insuffisant en cas d'arrêt de travail. C'est pourquoi ils font appel à une couverture supplémentaire, de manière individuelle ou collective, par le biais d'un organisme d'assurance complémentaire, afin de conserver un niveau de vie satisfaisant en cas d'arrêt de travail. Les couvertures varient selon les entreprises et les assureurs, qui peuvent inclure une période de franchise ou une durée maximale d'indemnisation dans les contrats. En général, les garanties sont exprimées en pourcentage du salaire de l'assuré, et dépendent de son degré d'incapacité ou d'invalidité.

Chapitre 2

La retraite en France

Le système de retraite français permet à l'ensemble des travailleurs de s'arrêter de travailler tout en continuant à percevoir un revenu appelé "pension". Le système de retraite était initialement destiné à assurer un minimum vital aux travailleurs trop âgés pour continuer à occuper un emploi. Ce système a beaucoup évolué au cours du temps et constitue le premier poste de dépenses de la protection sociale, en France. En effet, les pensions de vieillesse et de survie ont été estimées par la DREES¹ à 300 milliards d'euros, en 2015. Par ailleurs, 16 millions de personnes vivant en France ou à l'étranger possèdent, d'après la DREES, le statut de retraités de droit direct à fin 2015, et leur retraite est gérée par différents organismes qui dépendent du statut professionnel de la personne active.

2.1 Présentation et évolution du système de retraite

2.1.1 Présentation du régime Français

Éléments fondateurs et principes

Trois éléments principaux fondent le système de retraite français. Le premier concerne le caractère obligatoire dans le sens où toute personne entrée dans la vie active cotise pour sa retraite. De ce fait, des cotisations salariales et patronales sont versées, chaque année, directement par l'em-

ployeur au titre de la future pension de retraite de ses salariés. Le deuxième élément correspond au caractère d'un système par répartition. Ce deuxième élément s'appuie sur le principe de la solidarité inter-générationnelle car le montant total des cotisations, versées chaque année par les personnes en activité professionnelle, sert à payer les pensions des personnes retraitées, la même année. Les actifs financent ainsi les retraites de leurs aînés et s'ouvrent parallèlement des droits pour leur propre retraite. Enfin, le troisième élément concerne le caractère d'un système dit "contributif", où la pension perçue à la retraite est calculée en fonction du montant total des cotisations qui ont été versées durant toute la carrière professionnelle de la personne et proportionnellement à son revenu. En général, une personne active cotise automatiquement à deux caisses de retraite qui dépendent de son statut et qui correspondent à la caisse de retraite de base et la caisse de retraite complémentaire obligatoire. Des cotisations sont prélevées sur le revenu par les deux caisses et au moment du départ à la retraite, une pension est versée par les deux caisses. Il est à noter que le système de retraite en France est le seul à rendre obligatoire la retraite complémentaire contrairement aux autres pays de l'Union Européenne, où seul le régime de base est obligatoire. Par ailleurs, le système de retraite français comprend une dimension familiale importante car il prend en compte certains événements de la vie au moment du calcul des pensions de la retraite tels que les congés de maternité et met à disposition certains mécanismes tels que les rentes de réversion qui permettent au conjoint survivant de bénéficier, sous conditions, d'une partie de la retraite du conjoint décédé. Le système de retraite en France prend également en compte les aléas de la vie active, comme les périodes d'invalidité ou de chômage qui pourraient empêcher la personne active de cotiser pour sa retraite. De surcroît, il est possible de se constituer un complément de retraite en adhérant à un régime de retraite supplémentaire d'entreprise, proposé par l'employeur dans le cadre d'un contrat collectif ou en souscrivant à un contrat de retraite individuelle auprès d'un organisme d'assurance, par exemple. Enfin, le système de retraite en France peut être résumé par les trois piliers présentés dans le graphique suivant :

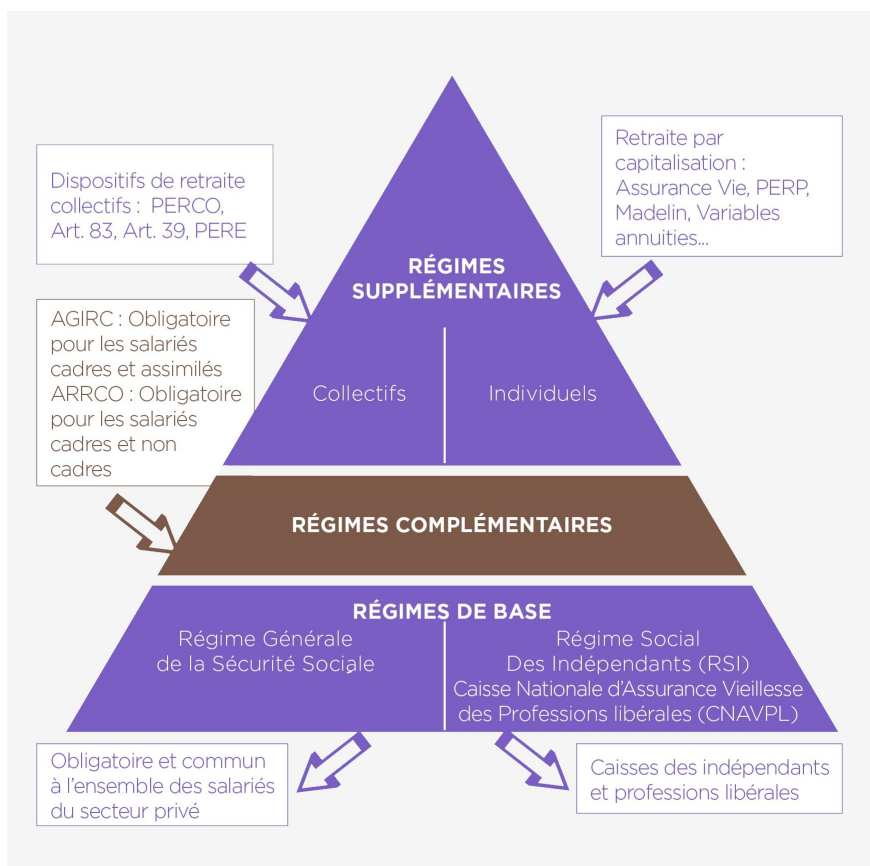


Figure 2.1 Organisation du système de retraite en France

Historique des réformes

Après la Seconde guerre mondiale, les réformes successives des retraites renforcent les garanties proposées par le système : extension de la retraite à l'ensemble des salariés (1945), instauration du minimum vieillesse (1956), mode de calcul des pensions plus généreux (1971), retraite complémentaire obligatoire (1972). L'âge de départ à la retraite est abaissé de 65 à 60 ans en 1982. Mais le Livre Blanc sur les retraites publié en 1991 fait le constat du vieillissement de la population et anticipe les déséquilibres à venir dans le financement des retraites. Le rapport envisage l'allongement de la durée de cotisation et un mode de calcul des pensions et de leur revalorisation moins généreux. Il préfigure les réformes à venir : il s'agit de sauvegarder le système en assurant son financement. La réforme Balladur de 1993 promulgue ainsi trois mesures importantes pour le régime général :

- allongement progressif de la durée de cotisation pour bénéficier d'une retraite à taux plein de 37,5 ans à 40 ans,
- calcul du salaire moyen de référence, base du calcul de la pension, sur les 25 meilleures années et non sur les 10 meilleures,
- revalorisation de la pension sur la base de l'évolution des prix et non plus des salaires ; cette mesure se traduit par une baisse importante des pensions de retraite, les prix évoluant généralement moins vite que les salaires.

Deux ans plus tard, le Premier Ministre Alain Juppé souhaite étendre la réforme Balladur à la fonction publique mais recule face à la mobilisation sociale. En 1999, le gouvernement de Lionel Jospin crée le Fonds de réserve des retraites, qui a vocation à faire face au choc démographique lié à l'arrivée à l'âge de la retraite des baby-boomers. Il est alimenté par des recettes exceptionnelles, des recettes fiscales et les excédents éventuels des régimes sociaux. Fin 2011, le fonds dispose de 35,1 milliards d'euros, loin des 150 milliards originellement prévus à l'horizon 2020. Le gouvernement installe également le Conseil d'Orientation des Retraites (COR), instance consultative rattachée au Premier Ministre qui a pour missions essentielles d'assurer le suivi de l'évolution des régimes et de formuler des propositions pour garantir leur solidité financière et leur fonctionnement solidaire. Les publications et études montrent que le financement à moyen et à long terme des retraites n'est pas assuré. Le Rapport Charpin, sur la base de projections financières sur l'évolution du système de retraite à l'horizon 2040, préconise dès 1999 de porter la durée de cotisation pour une retraite à taux plein à 42,5 ans en 2020. La réforme Fillon de 2003 s'en inspire et fait de la durée d'assurance le paramètre essentiel d'ajustement du système de retraite : elle pose le principe que la durée d'assurance nécessaire pour bénéficier d'une pension à taux plein évolue pour maintenir constant jusqu'en 2020, le rapport constaté entre cette durée et la durée moyenne de retraite, entendue comme l'espérance de vie à l'âge de 60 ans telle qu'estimée 5 ans auparavant. Cette réforme crée les conditions d'un accroissement futur de la durée de cotisation à 41 ans en 2012, et éventuellement plus au-delà de cette date. D'autre part, elle aligne la durée de cotisation pour une

retraite à taux plein (40 ans) des fonctionnaires sur celles des salariés du privé à l'horizon 2008. Elle crée également un dispositif dédié aux carrières longues, ainsi que deux nouveaux produits d'épargne retraite : le PERCO (Plan d'Épargne Retraite Collective), souscrit dans le cadre de l'entreprise, et le PERP (Plan d'Épargne Retraite Populaire), par capitalisation individuelle.

2.2 Réforme 2010

2.2.1 Présentation de la réforme

La loi 2010-1330 portant réforme des retraites est promulguée le 10 novembre 2010 avec l'objectif affiché de réduire le déficit du système de retraite français estimé à 32 milliards d'euros et d'atteindre l'équilibre à l'horizon 2018. Le budget de l'État consacré aux retraites a en effet atteint 275 milliards d'euros en 2010, soit 13 % du PIB. La réforme s'appuie sur des projections du COR estimant un déficit annuel du système de retraite français de l'ordre de 70 à 80 milliards d'euros en 2030. Avec l'augmentation de l'espérance de vie chaque année de 2 à 3 mois en moyenne à partir de 60 ans et la pyramide des âges actuelle, le ratio entre les actifs cotisants et les retraités passerait de 1,2 en 1950 à 1,8 en 2010. Plusieurs dispositifs sont mis en place, tels que :

- Alignement du taux de cotisation du public sur le privé en dix ans,
- Allongement de la durée de cotisation à 41,5 ans à l'horizon 2020,
- Prise en compte des conséquences de l'invalidité sur des critères individuels,
- Maintien du dispositif pour carrière longue,
- Utilisation anticipée du Fonds de réserve pour les retraites.

Les deux principales mesures de la réforme sont le report de 2 ans de l'âge minimum de liquidation de pension de retraite de 60 à 62 ans, ainsi que de l'âge de départ à taux plein qui passe de 65 à 67 ans. Ce décalage s'applique au régime général, aux régimes alignés et aux régimes de la fonction publique selon le calendrier suivant :

Année de naissance	ÂGE MINIMUM LÉGAL	ÂGE DU TAUX PLEIN
Avant 01/07/1951	60 ans	65 ans
1951	60 ans et 4 mois	65 ans et 4 mois
1952	60 ans et 8 mois	65 ans et 9 mois
1953	61 ans	66 ans et 2 mois
1954	61 ans et 4 mois	66 ans et 7 mois
1955	61 ans et 4 mois	67 ans
À partir de 1956	62 ans	67 ans

Table 2.1 Âges de départ à la retraite par génération après la réforme de 2010

La réforme des retraites de 2010 est unique dans le paysage français des retraites des 25 dernières années car c'est la seule qui ait programmé une augmentation rapide et substantielle des âges légaux. La dernière réforme modifiant les âges légaux (en l'occurrence l'âge d'ouverture des droits des salariés du secteur privé) remonte en effet à 1983. L'âge du taux plein avait alors été réduit de 65 ans à 60 ans et la réforme n'avait pas fait l'objet d'une étude approfondie. Le tableau 2.2 illustre, pour une date de naissance au 1^{er} janvier pour chaque génération (sauf la première qui débute en juillet), l'âge minimum légal et la date de départ à la retraite avant et après la réforme.

DATE DE NAISSANCE	ÂGE LÉGAL DE DÉPART		DATE DE DÉPART	
	Avant	Après	Avant	Après
1 ^{er} Juillet 1951	60 ans	60 ans et 4 mois	1 ^{er} Juillet 2011	1 ^{er} Novembre 2011
1 ^{er} Janvier 1952	60 ans	60 ans et 8 mois	1 ^{er} Janvier 2012	1 ^{er} Septembre 2012
1 ^{er} Janvier 1953	60 ans	61 ans	1 ^{er} Janvier 2013	1 ^{er} Janvier 2014
1 ^{er} Janvier 1954	60 ans	61 ans et 4 mois	1 ^{er} Janvier 2014	1 ^{er} Mai 2015
1 ^{er} Janvier 1955	60 ans	61 ans et 8 mois	1 ^{er} Janvier 2015	1 ^{er} Septembre 2016
1 ^{er} Janvier 1956	60 ans	62 ans	1 ^{er} Janvier 2016	1 ^{er} Janvier 2018

Table 2.2 Evolution de l'âge légal de départ et de la date de départ avant et après la réforme par génération

Fin 2011, la loi de financement de la sécurité sociale pour 2012 accélère le calendrier initialement prévu. Ainsi, l'âge légal de départ à la retraite pour les assurés nés à partir de 1952 est relevé de 5 mois (cf. tableau 2.3) par génération au lieu de 4 (cf. tableau 2.1). Dès 2016, l'âge légal de départ à la retraite atteindra alors 62 ans pour les personnes nées à partir de 1956. L'âge d'obtention du taux plein évolue également selon les mêmes modalités.

Année de naissance	ÂGE MINIMUM LÉGAL	ÂGE DU TAUX PLEIN
Avant 01/07/1951	60 ans	65 ans
1951	60 ans et 4 mois	65 ans et 4 mois
1952	60 ans et 9 mois	65 ans et 9 mois
1953	61 ans et 2 mois	66 ans et 2 mois
1954	61 ans et 7 mois	66 ans et 7 mois
À partir de 1955	62 ans	67 ans

Table 2.3 Nouvelles mesures : Âges de départ à la retraite par génération suite à la réforme de 2010

Selon l'étude de ARNAUD et al. publiée en 2019 [2], le recul de l'âge minimum de départ à la retraite « a entraîné une baisse des effectifs des nouveaux retraités en 2011 et, dans une moindre mesure, en 2012. En revanche, après avoir été successivement positif en 2013, négatif en 2014 et 2015, l'effet du calendrier de la réforme redevient positif à partir de 2016. Ces variations dépendent du nombre de personnes qui atteignent l'âge d'ouverture des droits¹ une année donnée. ... Cette même réforme a instauré un recul progressif de l'âge d'annulation de la décote, qui a une incidence sur le nombre de nouveaux retraités à partir de l'année 2016 et jusqu'en 2023. La contribution de cette mesure à la variation du nombre de nouveaux retraités dépend du nombre de personnes qui atteignent l'âge d'annulation de la décote une année donnée».

2.2.2 Enjeux pour la prévoyance collective

Dans cette partie, il est question motiver l'étude des impacts du recul de l'âge de départ à la retraite introduit par la réforme des retraites de 2010 sur les garanties incapacité et invalidité de la prévoyance collective. Cela se fera à travers une brève présentation des résultats d'études réalisées, corroborée par une présentation d'indicateurs socio démographiques.

Résultats d'études sur les impacts du décalage d'âge sur l'activité des seniors

Les réformes des retraites changent le comportement d'activité des seniors, c'est ce qu'une étude menée par la Direction des Statistiques Démographiques et Sociales [7] a mis en évidence en 2007. En effet, il a été constaté que la participation au marché de l'emploi entre 1970 et 2005 était très faible au delà de 65 ans. L'explication fournie est la suivante : «.. *Cette baisse s'explique en partie par des changements dans la structure des emplois ... , mais surtout par la mise en place de dispositifs favorisant le retrait de l'activité des seniors. Le recours à ces dispositifs a fortement été encouragé par les institutions et les partenaires sociaux, en réponse à la montée générale du chômage et à la détérioration du marché du travail des travailleurs âgés ... Ainsi, les années 1970 ont vu le développement des préretraites pour les 60-64 ans, entériné par l'abaissement de l'âge légal de la retraite à 60 ans en 1983. A la fin des années 1970, les cessations anticipées d'activité se sont étendues aux moins de 60 ans.*». Ainsi, dans un contexte où l'âge de départ à la retraite au taux plein était fixé à 65 ans, un lien a été établi entre cette réforme de l'âge de départ à la retraite et la modification du comportement d'activité au delà de 65 ans. Bien qu'ancienne, les conclusions de cette étude peuvent aisément se transposer à la réforme de 2010.

L'impact de la réforme sur l'activité des seniors a fait l'objet d'une étude [10] réalisée en 2017 par DUBOIS et KOUBI. Le but de cette étude a été d'évaluer l'impact sur l'activité des seniors de l'augmentation des âges légaux de départ à la retraite programmée par la réforme de 2010. Les auteurs affirment qu' «*en effet, la réforme a un effet direct d'augmentation de l'offre de travail des seniors car ceux-ci sont incités à prolonger leur activité à des âges où ils étaient naguère inactifs.*». Les conclusions de l'étude ne se limitent pas uniquement à l'effet sur le taux d'activité. Les répercussions du surcroît d'activité sur d'autres dispositifs sont aussi mis en exergue : «*Les seniors touchés par la réforme peuvent se retrouver au chômage ou, pour diverses raisons et notamment leur état de santé, dans des dispositifs hors activité (**phénomène de déversement vers d'autres dispositifs**)*»

Les mêmes auteurs DUBOIS et KOUBI se sont intéressés en 2017 à l'analyse des relations entre le report de l'âge de retraite et les taux d'emplois des seniors dans le cas particulier de la réforme des

retraites de 2010 [11]. Ils ont conclu qu'«entre les premières générations concernées par la réforme de 2010 et celles immédiatement antérieures, le taux d'activité à 60 ans a augmenté fortement : de 24 points pour les hommes et de 22 points pour les femmes. Avant la réforme, le taux d'activité à 60 ans était de 32% pour les hommes et de 43% pour les femmes. ... Cette hausse du taux d'activité à 60 ans s'est traduite surtout par un accroissement de l'emploi. Pour la même catégorie de personnes, la probabilité d'occuper un emploi a progressé de 17 points pour les hommes et de 16 points pour les femmes,... L'effet dominant de la réforme à court terme aurait été de figer les situations atteintes à de la soixantaine dans l'attente du nouvel âge d'accès à la retraite : **c'est surtout par l'allongement de la durée d'emploi des personnes encore en emploi entre 58 ans et 60 ans que la réforme aurait permis d'accroître l'emploi global.**»

Dans la littérature, plusieurs études réalisées sur un périmètre plus vaste (hors de France) ont abouti à des conclusions similaires concernant le sujet des impacts du recul de l'âge de la retraite sur les taux d'activité. On peut citer, entre autres, STAUBLI et ZWEIMÜLLER [14] en 2013 pour l'analyse des impacts du relèvement réglementaire de l'âge de départ à la retraite par sexe (2 ans pour les hommes et 3,25 ans pour les femmes) sur le taux d'emploi en Australie.

Indicateurs socio-démographiques

Les valeurs observées et projetées de la démographie et de la population active française permettent de faire des conjectures quant aux évolutions des âges réglementaires de départ à la retraite et leur impact sur les taux d'activité.

Le graphique 2.2 montre l'évolution de la population française (observée et projetée) entre 1990 et 2070. Il indique une population vieillissante à travers la tranche d'âge des 60 ans qui s'agrandit au fil des années (18,9% en 1990 à 34,5% en 2070).

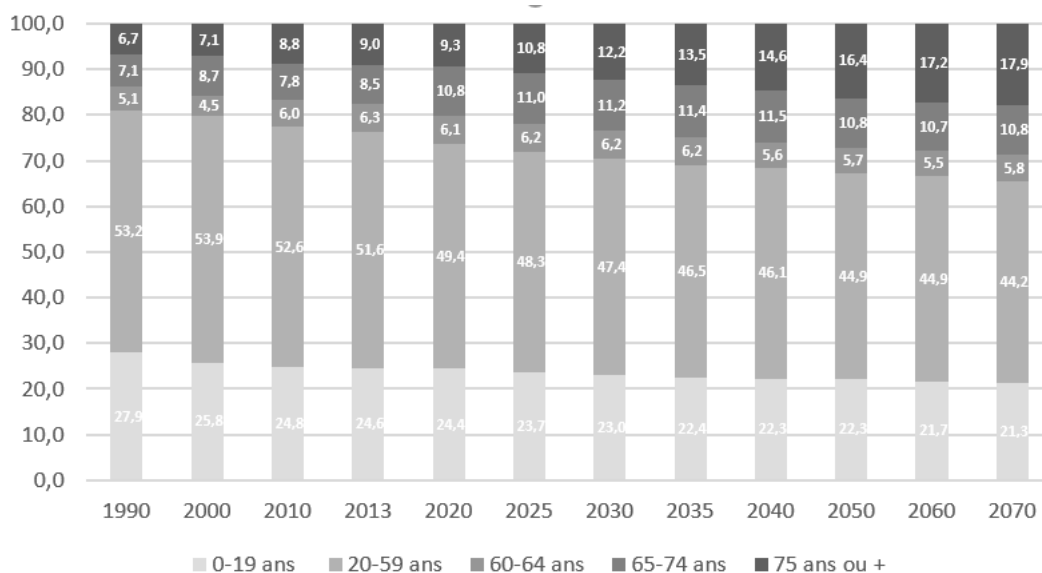


Figure 2.2 Évolution de la population de la France de 1990 à 2070 (scénario central de projection).

Source : *Projections de population à l'horizon 2070*, Insee

L'Insee utilise un modèle de micro-simulation dynamique nommé **Destinie**¹ permettant de donner, à l'aide d'un échantillon d'individus, l'évolution de la démographie française jusqu'à l'horizon 2070. Selon le scénario central, la population française atteindrait 76,5 millions d'habitants au 1er janvier 2070. L'augmentation toucherait principalement les habitants âgés de 65 ans ou plus dont la part augmenterait de 10,8 millions d'habitants par rapport à 2013. D'après [4], *En 2070, la France compterait plus d'une personne âgée de 65 ans ou plus pour deux personnes âgées de 20 à 64 ans.* *La part des personnes âgées continuerait d'augmenter sur le long terme, les âges de fin de carrière en seraient impactés à la hausse.* »

L'hypothèse selon laquelle les âges de fin de carrière augmenteront est corroborée par l'évolution de la population active. La figure 2.3 montre que le nombre d'actifs de 55 ans et plus continuerait d'augmenter fortement jusqu'en 2015 puis, à un rythme moins soutenu jusqu'en 2030. Ce vieillissement de la population active est principalement causé par la présence des générations nombreuses nées entre 1946 et 1970 dans cette classe d'âge. Il a commencé à la fin des années 1990 au moment

1. Les hypothèses et modules du modèles sont disponibles sur le site de l'Insee

où les premières générations du baby-boom fêtaient leur 55ème anniversaire. Il se poursuivrait jusqu'en 2030, et serait renforcé par une activité plus forte des seniors consécutive aux réformes des retraites.

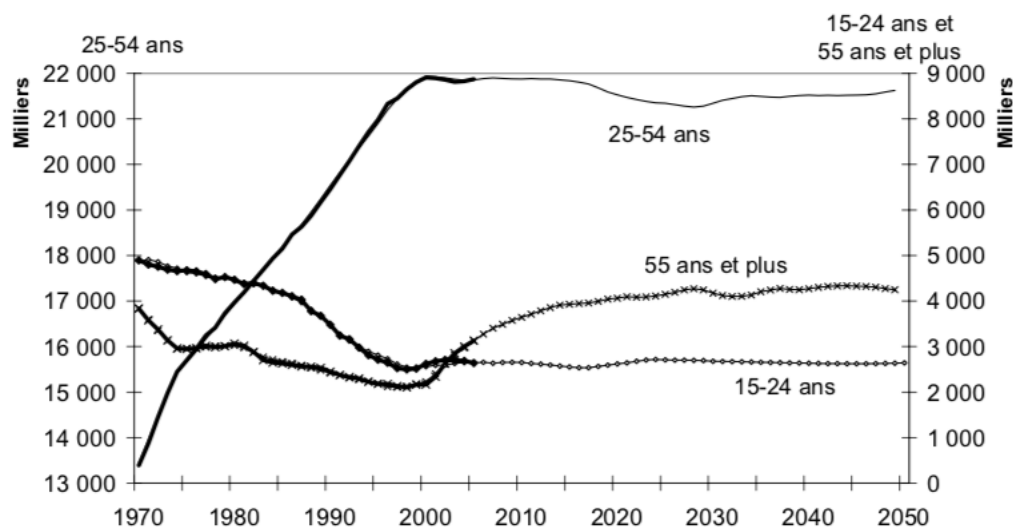


Figure 2.3 Population active observée et stimulée par âge en France métropolitaine.
Source : *Projections de population active 2006-2050, Enquêtes Emploi, Insee.*

Depuis 2000, les taux d'activité et des 60-64 ans se sont accrus en France comme dans l'ensemble de l'UE. Entre 2000 et 2017, le taux d'activité des 60-64 ans a augmenté d'un peu plus de 20 points en France (passant de 10,8 à 31,5%) et dans l'UE à 27 (de 24,3 à 45,2%).

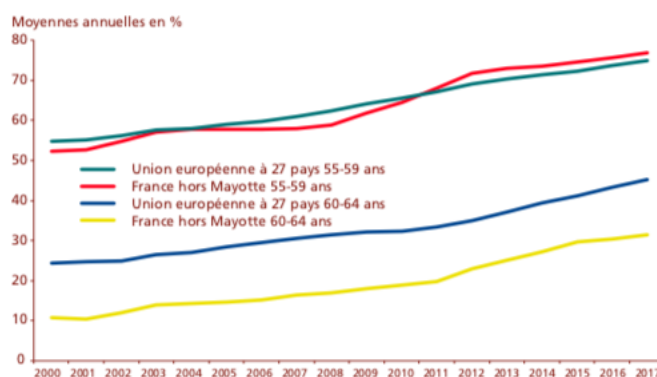


Figure 2.4 Taux d'activité des 55-59 ans et des 60-64 ans en France et dans l'Union Européenne à 27 pays depuis 2000.
Source : *Emploi et chômage des seniors en 2019, DARES*

Les parties précédentes ont permis de montrer que l'âge de départ à la retraite et le taux d'activité sont fortement corrélés positivement. En effet, la réforme induit une hausse de l'activité des seniors. C'est ce qui explique que l'on puisse, valablement, s'intéresser à l'étude des impacts du recul réglementaire de l'âge de départ à la retraite sur la prévoyance collective. L'exposition de l'assureur aux risques incapacité et invalidité en devient plus importante. Face à cette extension de la période de couverture de ces garanties, la loi n'indique pas expressément que les assureurs ont l'obligation de couvrir deux années de prestations supplémentaires. Cependant en pratique, la non application de cette disposition expose l'assureur à un risque de réputation, et à un risque légal : dans le cas où le contrat prévoit l'âge de 60 ans comme terme des prestations, la position d'un assureur refusant de verser des prestations après cette date sans avenant semble difficilement tenable, car la doctrine de la Sécurité Sociale considère comme non éligibles au régime de faveur de l'article L. 242-1 du code de la Sécurité sociale les contrats prévoyant des limites d'âge. D'autres contrats font référence à l'âge prévu à l'article L. 351-1 du code de la Sécurité sociale, qui a été modifié par la réforme. Dans ce cas, l'assureur est tenu contractuellement de poursuivre les prestations jusqu'à la liquidation de la retraite. En reportant de deux ans l'âge légal de départ à la retraite, la loi de novembre 2010 sur les retraites a posé la question de la prolongation des prestations incapacité et invalidité prévue par la loi Evin dans le cadre des contrats collectifs. La loi Evin oblige en effet les organismes assureurs à constituer intégralement les provisions et capitaux de couverture nécessaires au complet paiement des prestations et au maintien intégral des garanties décès, au bénéfice des assurés malades ou invalides, en cas de résiliation d'un contrat collectif. Le législateur a organisé un mécanisme transitoire d'étalement de la hausse des provisions en prévoyant :

- d'une part, que les organismes assureurs puissent étaler les provisions liées au changement de réglementation sur une période de six ans à partir de 2010 ;
- d'autre part, qu'en cas de résiliation ou non renouvellement du contrat durant cette période transitoire de six ans, une indemnité est due par l'établissement assuré à l'organisme assureur dans la mesure où ce dernier est tenu de maintenir le versement des prestations. Le montant de

l'indemnité de résiliation est égal à la différence entre le montant des provisions à constituer et le montant des provisions effectivement constituées. Conformément à la loi, cette mesure s'applique aux entreprises/établissements sans qu'il soit nécessaire de la notifier par avenant.

A compter du 1er mars 2010, selon l'article 67 de la loi du 24 décembre 2009 de financement de la Sécurité Sociale pour 2010, les invalides dont le niveau d'incapacité permet la poursuite d'une activité peuvent rester en emploi après l'âge légal de retraite tout en percevant leur pension d'invalidité de 1ère catégorie jusqu'à l'âge de retraite à taux plein. La perception de la pension de vieillesse allouée au titre d'inaptitude au travail n'est désormais concédée que sur demande expresse de l'assuré. Auparavant, la pension d'invalidité était automatiquement remplacée par la pension de vieillesse dès l'âge légal de retraite. Cette mesure augmente donc également les engagements d'invalidité de 1ère catégorie des assureurs, dont le terme coïncide généralement avec celui des pensions d'invalidité, qui peuvent désormais être prolongées jusqu'à l'âge de retraite à taux plein, soit 67 ans suite à la réforme.

2.2.3 Interaction avec d'autres réformes

Les autres réformes en cours interagissent peu avec l'augmentation des âges légaux, hormis le dispositif des carrières longues qui joue en sens contraire. La réforme des âges de 2010 est contemporaine d'autres mesures concernant la retraite et l'emploi des seniors, qui modifient également potentiellement le comportement d'activité des seniors. Il s'agit, dans le domaine des retraites, de l'augmentation de la durée d'assurance pour obtenir le taux plein et des différentes extensions du dispositif des carrières longues. Sur le marché du travail, il s'agit de la fermeture du dispositif des dispenses de recherche d'emploi. L'interaction avec la réforme des durées est de fait assez limité. Elle est un peu plus importante que les extensions du dispositif des carrières longues, mais seulement à partir de la génération née en 1953. Par ailleurs, ce dispositif joue dans le sens opposé à celui de la réforme des âges, en relâchant précisément la condition d'âge dans certaines situations. Une interaction faible à court terme avec l'augmentation des durées. Outre l'âge légal d'ouverture des droits, le départ à la retraite avec bénéfice du taux plein (absence de décote ou de surcote),

est conditionné par la validation d'un nombre défini de trimestres de durée d'assurance. Cette durée d'assurance requise a été modifiée à plusieurs reprises depuis 2003, afin de tenir compte de l'évolution démographique et de l'augmentation de l'espérance de vie.

La réforme de 2003 fixe ainsi le principe d'évolution par génération de la durée d'assurance nécessaire pour bénéficier d'une pension de retraite au taux plein avec pour objectif de maintenir constant, jusqu'en 2020, le rapport entre cette durée et l'espérance de vie à la retraite. Dans un premier temps, la réforme a fait converger, pour les générations 1943 à 1948, la durée d'assurance requise dans les régimes de la fonction publique vers celle requise au régime général (40 ans pour ces générations). Pour les générations de 1949 à 1952, elle a fait évoluer la durée jusqu'à 41 annuités au rythme d'un trimestre par génération. La réforme prévoyait en outre pour les générations 1953 à 1960, de la faire évoluer par décret, par des rendez-vous quadriennaux sur avis de la Commission de garantie des retraites et du Conseil d'orientation des retraites, sur le principe de maintenir le rapport entre la durée requise et la durée moyenne de retraite. La réforme 2010 a maintenu ce principe, tout en modifiant le principe de rendez-vous quadriennaux en confiant au COR l'obligation légale de prendre, chaque année, un avis technique sur l'évolution de la durée d'assurance, applicable pour la génération ayant 56 ans cette année-là.

La réforme 2014, enfin, fixe l'augmentation de la durée de cotisation d'un trimestre tous les trois ans entre 2020 et 2035, jusqu'à 172 trimestres pour la génération 1973. L'évolution du taux d'activité des séniors peut donc être impactée par la durée d'assurance requise qui évolue au cours des générations en particulier pour les générations ici étudiées. Néanmoins, cette évolution a moins d'impact sur l'âge de départ à la retraite de ces générations que la réforme de 2010, ces générations ayant, en particulier pour les hommes, souvent validé leur durée d'assurance avant d'avoir atteint l'âge minimal auquel elles peuvent liquider. Ainsi, parmi les hommes de la génération 1946 (la plus récente pour laquelle on dispose de données complètes concernant leur départ à la retraite), seuls 13% partent avec une décote ou sans décote mais au bénéfice de l'âge (Drees, 2016). Les femmes sont plus nombreuses (28%) et l'interaction entre les contraintes d'âge et de durée est potentiellement plus forte pour celles-ci.

Chapitre 3

Présentation des données et statistiques descriptives

Cette partie présente les données à partir desquelles les travaux ont été réalisés. Elle décrit les fichiers bruts utilisés, puis les différents traitements réalisés permettant d'obtenir des informations exploitables et enfin énumère les choix faits et les raisons qui y ont conduit.

3.1 Présentation des bases de données

Les données utilisées pour cette étude sont extraites de différentes bases de données de sinistres de l'entrepôt d'AXA et de sources extérieures. On distingue deux principales bases :

Base MRM

Suite à la réforme des retraites de 2010, l'outil MRM a été créé pour déterminer les provisions mathématiques des arrêts travail après expiration de la franchise. Les informations relatives à ces sinistres indemnisés sont extraites de deux sources de données différentes selon le type de gestion (directe ou déléguée). Ainsi, l'outil regroupe des données mensuelles pour tous les sinistres connus à différentes dates d'inventaire avec plusieurs variables descriptives. Les variables de la table qui seront utilisées dans le cadre de l'étude sont les suivantes :

CARACTÉRISTIQUES	VARIABLES
<i>Contrat</i>	<ul style="list-style-type: none"> • Numéro du contrat • Type de garantie : IT, IP ou IT-IP • Dates de début et de fin d'effet
<i>Sinistre</i>	<ul style="list-style-type: none"> • Gestion : direct ou déléguée • Statut : ouvert ou clos • Origine : Maladie, maladie professionnelle, accident du travail ou de la vie privée • Type d'arrêt : incapacité ou invalidité • Dates d'inventaire • Date de dernière échéance réglée • Date d'entrée en invalidité • Code de la catégorie de l'invalidité • Provision mathématique avant et après réforme • Montant de la prestation versée
<i>Assuré</i>	<ul style="list-style-type: none"> • Prénom, date de naissance, sexe • Régime, CSP, salaire • Identifiant unique associant le salarié à son sinistre

Table 3.1 Variables de la base de données MRM

Déclaration Sociale Nominative

La DSN est un fichier mensuel alimenté par les logiciels de paie destiné à communiquer les informations nécessaires à la gestion de la protection sociale des salariés aux organismes et administrations concernées. Depuis sa mise en place effective en janvier 2017, elle permet aux employeurs de signaler les arrêts de travail de leurs salariés tous les mois. Cette base contient tous les sinistres déclarés par les entreprises ce qui représente un volume de données beaucoup plus important que la base MRM. Cependant, l'objectif de cette étude et l'historique des données disponibles rendent son exploitation limitée. Elle servira à enrichir la base de données MRM avec les masses salariales disponibles, les domaines d'activité et les effectifs salariés inconnus de la base MRM.

La construction d'une base de données fiable et adaptée aux objectifs de l'étude est nécessaire pour comprendre et analyser les résultats. Pour se faire, une première étape consiste à déterminer un périmètre précis appliqué aux données brutes de la base principale.

3.2 Traitement des données

Dans le processus de création d'une base fiable, le traitement des données permet de gérer les incohérences et d'apporter des corrections aux bases brutes selon les besoins de l'étude. Après avoir clairement défini le périmètre de l'étude et les hypothèses sur le choix des observations retenues, les traitements ont consisté en la création et la correction des variables.

3.2.1 Périmètre de l'étude et hypothèses

Le périmètre de l'étude est un choix très important puisqu'il va conditionner l'ensemble des résultats. Différentes contraintes antagonistes existent : le nombre d'années d'observation doit être suffisant pour obtenir des volumes d'étude satisfaisants, des années trop anciennes peuvent intégrer des informations caduques. La base MRM regroupe des sinistres survenus entre 1971 et 2020. En raison du faible volume de données (moins de 1% de l'ensemble des données), les survenances antérieures à 2011 ne sont pas prises en compte. De plus, les données de l'année 2020 ne sont

pas encore suffisamment consolidées pour être exploitables. Le portefeuille sera donc restreint aux sinistres survenus entre le 01/01/2011 et le 31/12/2019, ce qui représente une période d'observation de 9 ans.

Les contrats concernés par la réforme sont les contrats avec des garanties IT, IP ou ITIP (avec ou sans garantie décès). Le champ a été réduit aux régimes obligatoires et facultatifs des salariés ainsi que certains régimes des travailleurs non salariés. Le périmètre regroupe uniquement les sinistres ayant donné lieu à indemnisation de AXA d'un montant journalier au moins égal à un euro. La base de donnée renferme des sinistres clos ou ouvert pour les risques incapacité temporaire et invalidité permanente.

Pour fiabiliser la base, un certain nombre d'hypothèses ont été prises. Elles sont répertoriées ci-dessous :

Regroupement de sinistres : Dans la base il existe, pour une même personne, des sinistres avec des dates de survenances qui se chevauchent (moins de 2 jours d'intervalle). Le sinistre qui a la date la plus ancienne est considérée comme un doublon et donc supprimé.

Suppression de sinistres : Les sinistres dont la date de survenance et la date de dernier inventaire sont les mêmes sont supprimés de la base.

Filtre selon l'âge à la survenance : Les sinistres pour lesquels l'âge à la survenance n'est pas compris entre 14 (âge minimum légal d'emploi) et 70 (âge de mise en retraite obligatoire) ans sont supprimés.

Correction de dates de passages en invalidité : Pour les sinistres IT qui évoluent vers une IP, au bout de 3 ans la date de passage en invalidité correspond à la date du 3e anniversaire de la date de survenance.

3.2.2 Variables créées

ID : Il s'agit d'un identifiant unique par assuré sinistré qui rassemble les informations personnelles du salarié (le numéro du contrat, la date de naissance et le prénom). Elle permettra d'avoir

le nombre de personnes concernées et le compteur de sinistres de l'assuré.

Indemnité journalière moyenne : Dans la base MRM chaque ligne correspond à la vision du sinistre à une date d'inventaire. Ainsi, selon les dates, le montant de la prestation (de même que le salaire et la date de dernière échéance réglée) peut être amené à évoluer. Cette nouvelle variable représente la moyenne des sommes versées sur une base journalière.

Remarque : Toutes les prestations n'ont pas la même périodicité de règlement et sur certaines observations le montant réglé peut être inférieur ou égal à un euro. Ces observations introduisent un biais et ne sont pas prises en compte dans le calcul de la moyenne.

Date de l'ancienneté : C'est la date qui servira au calcul de l'ancienneté en arrêt. Elle correspond à la date la moins récente entre la dernière date d'inventaire et la dernière date de règlement. Lorsque l'incapacité évolue vers une invalidité, c'est la date de passage à l'état d'invalidité qui est retenue.

Durée : C'est la distance entre la date de survenance du sinistre et la date de l'ancienneté.

Âge à la survenance : C'est la distance entre la date de naissance de l'assuré et la date de survenance du sinistre.

Provenance du sinistre : Indicateur de l'évolution de l'invalidité. C'est une variable catégorielle valant 1 si le sinistre provient d'une IT et 0 sinon.

Catégorie Socio-Professionnelle : Une correspondance a été faite entre la catégorie de personnel concernée par chaque contrat et la CSP qui n'est pas directement disponible dans la base. Trois catégories sont retrouvées : cadres, non cadres et ensemble du personnel.

Domaine d'activité : Chaque contrat a été rattaché au secteur d'activité concerné à l'aide des DSN. En effet, les domaines d'activité liées à chaque contrat sont renseignées dans la DSN. A partir d'une fusion sur le numéro de contrat entre les bases MRM et DSN, les 9 domaines d'activité ont été retrouvés. Cependant, pour 30% des sinistres l'information est manquante.

Durée estimée : La durée résiduelle a été estimée pour chaque sinistre en cours à partir de la formule suivante :

3.2.3 Variables corrigées

Certaines variables des bases brutes sont pertinentes pour notre étude mais présentent des anomalies ou sont manquantes. Il a donc fallu les reconstituer à l'aide de traitements et de statistiques externes collectées à partir de différentes sources.

Sexe : La base MRM contient des données provenant à la fois des bases de gestion directe et déléguée. La variable sexe est une variable catégorielle prenant 4 valeurs différentes : 0, 1, 2 et 3. Aucune règle de correspondance n'est établie, à titre d'exemple le prénom Jeanne est à la fois attribué aux quatre catégories alors qu'à priori il s'agit bien d'un prénom féminin. Le prénom des assurés étant une variable fiable et disponible pour chaque observation, il permettra de déterminer le genre. Le fichier des données nationales contenant tous les prénoms attribués aux enfants nés en France hors Mayotte entre 1900 et 2018 et les effectifs par sexe associés à chaque prénom, disponible sous ce [liens](#) permis de rétablir le sexe pour chaque individu. Les variables du fichier sont : le sexe, le prénom usuel, l'année de naissance et l'effectif. Pour chaque prénom, il est indiqué pour chaque année de naissance (de 1900 à 2018) et chaque sexe, le nombre de personnes inscrites à l'état civil sous ce prénom. Un poids de représentation selon le genre est alors attribué à chaque prénom. Par exemple, le prénom Jeanne est porté à 99,98% par des femmes contre 0.02% pour les hommes. Le sexe représenté à plus de 95% est retenu pour chaque prénom. Ainsi, tous les assurés de la base dénommés Jeanne sont des femmes.

Statut : L'information de la clôture du sinistre n'est pas toujours renseignée. En effet, certains sinistres sont toujours en statut "Ouvert" alors qu'aucune provision n'est calculée depuis la date d'ancienneté. Il s'agit d'erreurs de gestion qui sont corrigées en fonction de la distance entre la dernière date d'ancienneté et la date d'extraction de la base (31/03/2020). Ainsi,

deux cas sont à considérer :

- Date de dernière échéance réglée connue : Si la distance entre la date d'ancienneté et celle d'extraction est supérieure à une période alors le sinistre est considéré clos. La période correspond à la périodicité du règlement de l'indemnité : mensuelle, trimestrielle, semestrielle ou annuelle.
- Date de dernière échéance réglée inconnue : Si la distance entre la date d'ancienneté (qui est alors la date d'inventaire) et celle d'extraction est supérieure à 31 jours alors le sinistre est considéré clos car les données de la base sont mensuelles

3.3 Statistiques descriptives

L'objectif de cette partie est d'étudier les principales caractéristiques de la population sinistrée pour avoir une meilleure connaissance de celle-ci.

3.3.1 Répartition du volume de sinistres

Par survenance : Il est important d'analyser l'évolution du volume de sinistres en fonction des années pour pouvoir prétendre analyser un éventuel effet de la réforme des retraites.

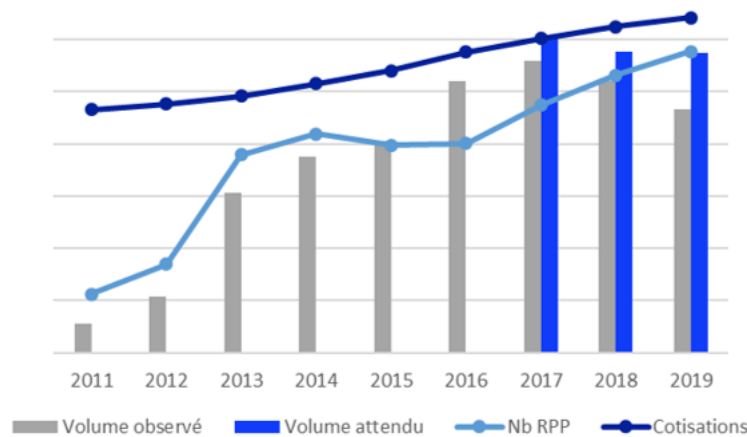


Figure 3.1 Evolutions du nombre de sinistres observé et attendu, du nombre de contrats (RPP) et du montant des cotisations par année de survenance

Le nombre de sinistres augmente naturellement compte tenu des évolutions de la démographie des entreprises mais aussi du volume de nouvelles affaires. L'étude de la cadence d'inventaire des sinistres a permis d'obtenir une estimation du volume attendu pour les survenances pour lesquelles les données ne sont pas encore exhaustives (2017 – 2019). Une correction a été apportée pour tenir compte de l'impact de l'augmentation des RPP à travers le montant des cotisations. Le portefeuille semble atteindre son régime de croisière (effectifs plus ou moins stables à partir d'une certaine date) et une stabilisation de la sinistralité se présente. En 2016, le volume de sinistres a augmenté de 28% par rapport à 2015. En parallèle, le décalage de l'âge légal de la retraite de 60 à 62 ans commence en 2016 pour les générations nées à partir de 1956 (cf. 2.2). A partir de 2017 le volume se maintient progressivement. Pour apprécier l'impact de la réforme sur la hausse de la sinistralité à partir de 2016, il est intéressant de regarder l'évolution du nombre de sinistres pour les effectifs sous risque âgés de plus de 60 ans.

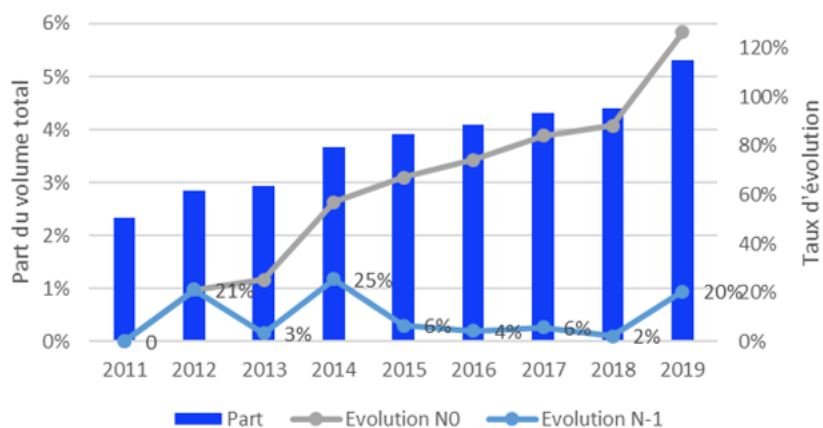


Figure 3.2 Évolutions de la part des assurés âgés de plus de 60 ans à la survenance du sinistre

La proportion des personnes âgées de plus de 60 ans à la survenance de leur arrêt est en croissance continue d'une année à l'autre (cf. courbe bleue). En 2016, le volume de sinistres a évolué de 74% par rapport à 2011 qui représente l'année de référence N_0 (cf. courbe grise). L'étude [10] réalisée par DUBOIS et KOUBI en 2017 citée plus haut permet de conforter ces

résultats. Les résultats de cette étude suggèrent que la réforme a eu un effet très incitatif sur l'activité des seniors. En effet, une des conclusions de l'étude indique qu'après la réforme, la probabilité d'être actif à 60 ans augmente de plus de 20 points. Cet accroissement se traduit essentiellement par de l'emploi (+17 points pour les hommes et +16 points pour les femmes). Le taux d'activité des seniors est en hausse selon l'Insee, ce qui entraîne un vieillissement de la population active.

Par sexe : La répartition des arrêts en fonction du sexe permet d'identifier des tendances dans le portefeuille de sinistres.

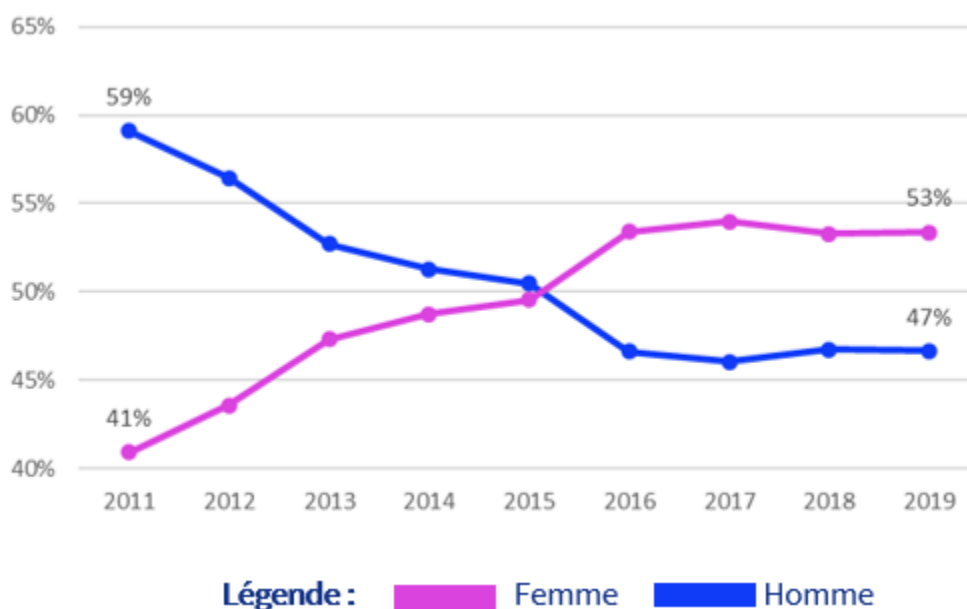


Figure 3.3 Composition du portefeuille par sexe

La proportion d'hommes dans le portefeuille diminue au fil des années. On a une population qui se « féminise », en effet les femmes sont de plus en plus nombreuses sur le marché de l'emploi. De manière générale, elles sont plus touchées par l'absentéisme. La maternité, l'accompagnement des enfants et des parents vieillissants, entre autres, les rendent plus exposées aux arrêts de travail.

3.3.2 Répartition des âges à la survenance

La répartition des arrêts en fonction de l'âge permet d'analyser la sinistralité.

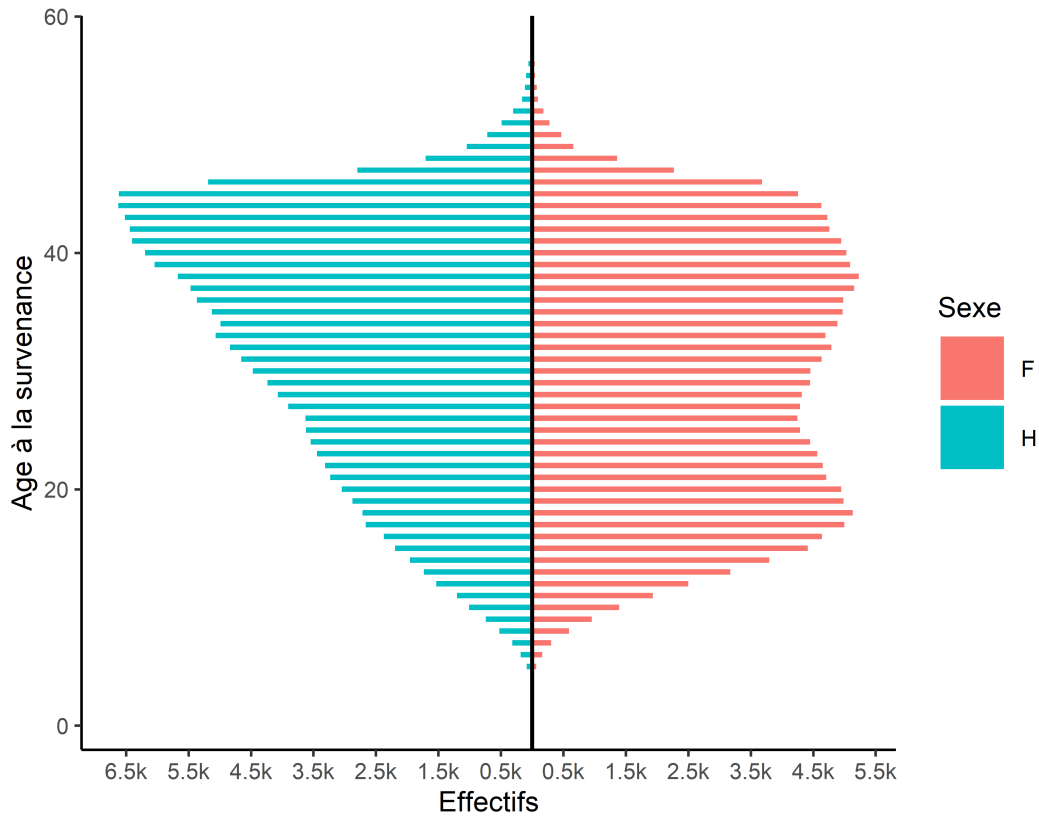


Figure 3.4 Pyramide des âges

La pyramide des âges s'apparente à celle de l'Insee sur la population française vue en 2020. La sinistralité augmente avec l'âge, notamment chez les hommes. Jusqu'à l'âge de 55 ans, les femmes sont plus nombreuses que les hommes et représentent 41 % des seniors. L'espérance de vie sans incapacité à la naissance est plus faible chez les hommes. En 2018 elle était de 63,4 ans chez les hommes contre 64,5 selon la Drees [9]. Les femmes sont en bonne santé plus longtemps que le sexe opposé. L'âge moyen du portefeuille est de 44 ans au global, 50,7 ans en invalidité et 43,4 ans en incapacité.

Construction des classes Pour mieux étudier l'échantillon et pouvoir faire des comparaisons cohérentes pour les différents indicateurs, la construction de classes d'âges est importante. Les

besoins de l'étude et la connaissance métier impliquent la création d'une classe pour les assurés de plus de 60 ans. Une marge de 2 ans est prise, entraînant la construction d'une tranche pour les assurés ayant 58 ans et plus au vu des conclusions de l'étude [10] citée plus haut (partie 2.2.2). Le nombre de classes ainsi que les valeurs limites de ces classes doivent être justifiées statistiquement. Pour ce faire, l'algorithme de Jenks basé sur les seuils naturels observés a permis de construire 4 classes en plus de celle déjà établie (la méthode est décrite en annexe A).

Suite à l'utilisation de l'algorithme de Jenks, les tranches d'âge contruites sont les suivantes :

- Tranche 1 : 14 – 32 ans
- Tranche 2 : 33 – 41 ans
- Tranche 3 : 42 – 49 ans
- Tranche 4 : 50 – 57 ans
- Tranche 5 : 58 – 70 ans

NB : Dans toute la suite, les seniors représentent les personnes âgées de 58 ans et plus.

3.3.3 Durée des arrêts

Pour identifier un éventuel impact de la réforme des retraite il est intéressant de regarder comment évoluent les durées moyennes en fonction des survénances. L'idée étant de pouvoir comparer l'évolution de la durée moyenne sur les différentes années de survénances, il faudrait idéalement considérer uniquement les arrêts clos. A la date d'extraction de la base d'étude une grande partie des sinistres étaient en cours comme l'indique le tableau 3.2. En retenant uniquement les sinistres clos, la perte de données serait trop importante notamment pour le risque invalidité où 48% des arrêts sont en cours. Pour palier à cette perte d'information, les durées restantes théoriques ont été estimées pour les arrêts en cours afin de pouvoir réaliser des comparaisons pertinentes. L'estimation des durées restantes se fait à travers les formules de provisionnement en incapacité (cf. 3.1) et en invalidité (cf. 3.2) en prenant un taux technique et un taux d'arrérages nuls. Les tables de

maintien en incapacité et en invalidité utilisées ont été construites en *best estimate* sur les données issues de la même base MRM. Les tables ayant été construites sur la même population étudiée, les durées calculées sont bien fiables et reflètent la réalité du portefeuille.

	2011	2012	2013	2014	2015	2016	2017	2018	2019	GLOBAL
INCAPACITÉ							2%	7%	28%	6%
INVALIDITÉ	34%	38%	38%	45%	51%	58%	63%	71%	75%	48%

Table 3.2 Proportion de sinistres encore ouverts à la date du 30 mars 2020 par année de survenance

$$\tilde{P}_{IT}(x, n) = \sum_{k=1}^{36-n} \frac{\tilde{l}_{x:n+k}}{\tilde{l}_{x:n}} \frac{1}{(1+i)^k} \cdot (1+T_a) \quad (3.1)$$

$$\tilde{P}_{IP}(x, n, \omega) = \sum_{k=1}^{\omega-(x+n)} \frac{\bar{l}_{x:n+k}}{\bar{l}_{x:n}} \frac{1}{(1+i)^k} \cdot (1+T_a) \quad (3.2)$$

Avec :

x = l'âge de l'assuré à la survenance ;

n = l'ancienneté ;

ω = l'âge de départ à la retraite ;

$\tilde{l}_{x:n}$ = le nombre d'individus en incapacité (resp. d'invalidité) d'ancienneté n mois (resp. n années)

et entré à l'âge x dans cet état ;

i =taux technique ;

T_a = taux d'arrérages ;

Les survenances 2011 à 2013 ne sont pas prises en compte pour l'incapacité à cause d'un manque de données susceptible de biaiser les résultats. En effet, les arrêts clôturés avant le 31 octobre 2013 ne figurent pas dans la base MRM, le volume de données en incapacité sur la période 2011 - 2013 est donc sous-estimé.

Incapacité

Le volume de sinistres en incapacité diminue en fonction de l'ancienneté. Le graphique 3.5 montre une décroissance exponentielle avec un pic à 36 mois. Les incapacités ayant atteint une durée 36 mois évoluent, dans 96% des cas, vers l'état d'invalidité. Le pic observé à 36 mois est donc expliqué par l'aggravation quasi probable de l'état d'incapacité.

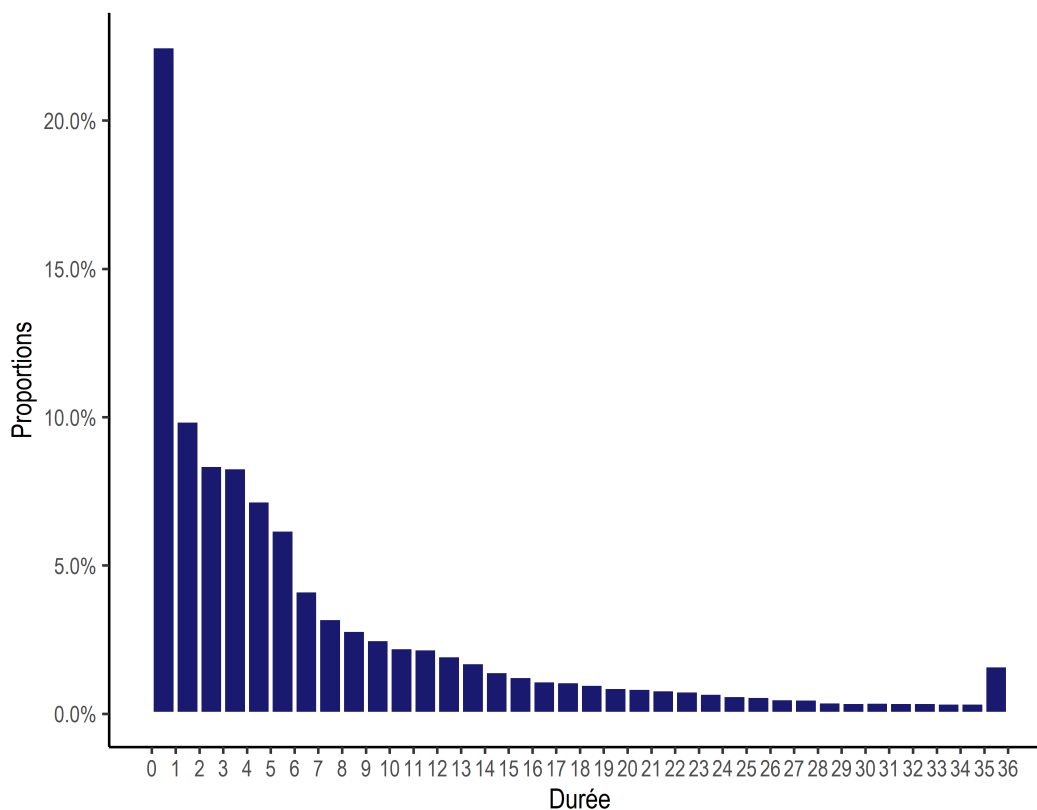


Figure 3.5 Distribution des arrêts selon l'ancienneté en incapacité

Les durées moyennes et médianes en incapacité sont exprimées en mois et les moyennes pondérées par les effectifs. Au vu du tableau 3.3, aucune tendance particulière (hausse ou baisse) ne se décline en terme de durée. Sur l'ensemble des sinistres la durée moyenne est de 6,7 mois.

	2014	2015	2016	2017	2018	2019
DURÉE MÉDIANE	4.1	4.4	3.3	3.4	3.8	3.7
DURÉE MOYENNE	6.9	7.0	5.9	5.8	5.9	5.5

Table 3.3 Évolution des durées médianes et moyennes en fonction des survenances

La figure 4.11 montre une croissance de la durée moyenne en fonction de l'âge. Les personnes âgées de 58 ans et plus (tranche 5) ont une durée moyenne de 7,5 mois. Le graphique 3.6 révèle que la durée moyenne de l'ensemble de la population est largement inférieure à celle des seniors pour toutes les années de survenance.

	TRANCHE 1	TRANCHE 2	TRANCHE 3	TRANCHE 4	TRANCHE 5	GLOBAL
DURÉE MOYENNE	4.2	6.2	7.4	8.3	7.5	6.7

Table 3.4 Durée moyenne par tranche d'âge

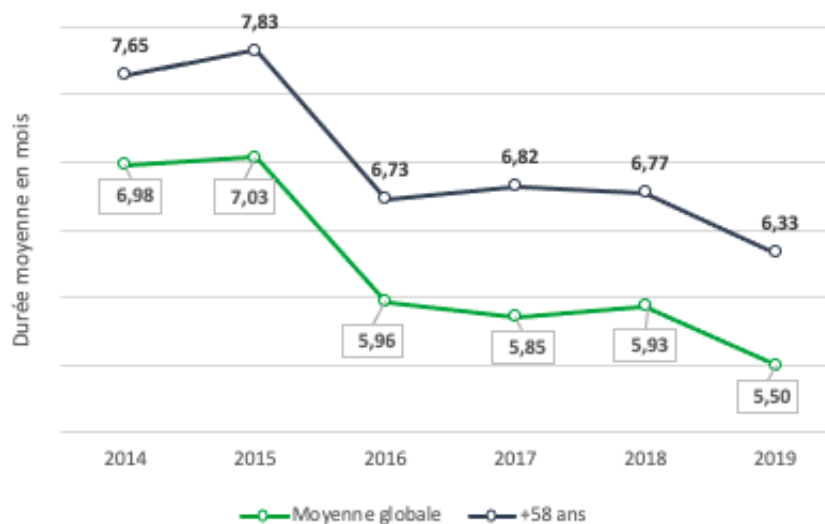


Figure 3.6 Comparaison de la durée moyenne des seniors et de l'ensemble de la population en incapacité

Invalidité

Le volume de sinistres en invalidité décroît avec l'ancienneté. Les durées moyennes (cf. 4.12) semblent plutôt stables en fonction des années, avec une faible tendance à la baisse.

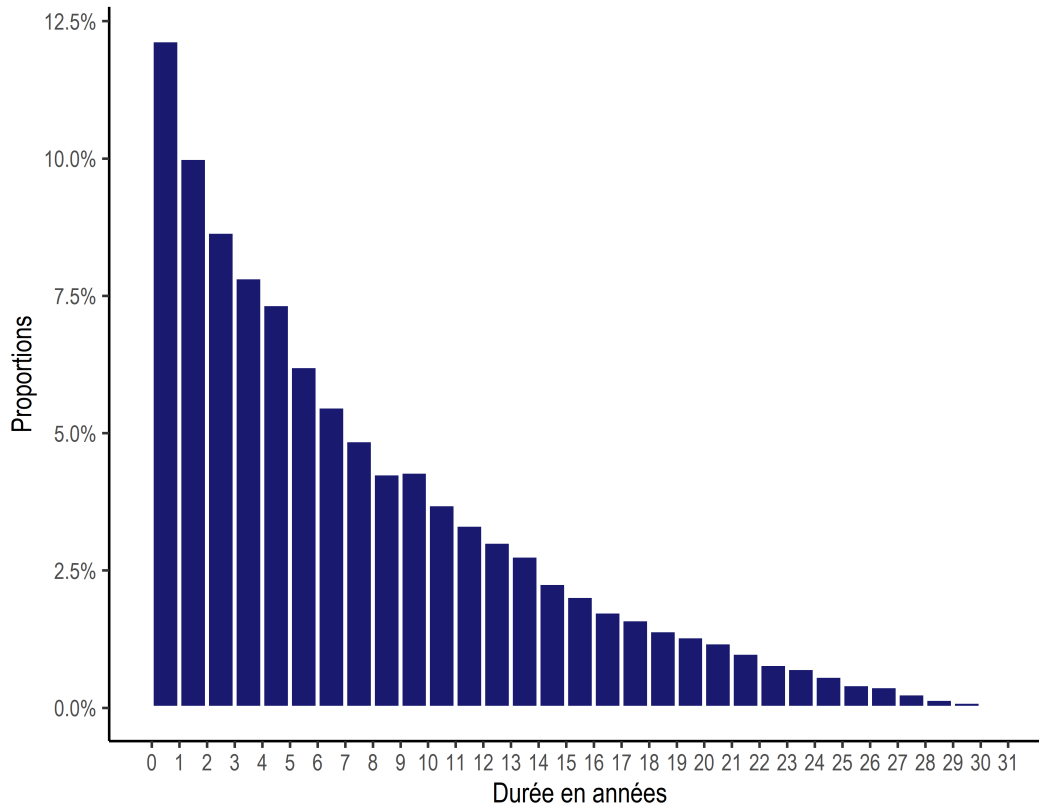


Figure 3.7 Distribution des arrêts selon l'ancienneté en invalidité

	2011	2012	2013	2014	2015	2016	2017	2018	2019
DURÉE MÉDIANE	5.58	5.42	4.83	5.57	5.98	6.00	5.92	5.76	5.70
DURÉE MOYENNE	7.09	7.07	6.79	7.12	6.97	6.96	6.75	6.60	6.49

Table 3.5 Évolution des durées médianes et moyennes en fonction des survenances

Contrairement à l'incapacité, ici la durée moyenne est décroissante en fonction de l'âge. Le tableau 3.6 indique que les plus de 58 ans passent en moyenne 2 ans en invalidité. La comparaison

de la durée moyenne des seniors et celle du reste de la population dévoile deux informations importantes. La durée moyenne des seniors augmente au fur et à mesure des années, ce qui découle du retard progressif de l'âge de départ à la retraite. A l'opposée, la durée moyenne du reste de la population présente une tendance baissière.

	TRANCHE 1	TRANCHE 2	TRANCHE 3	TRANCHE 4	TRANCHE 5	GLOBAL
DURÉE MOYENNE	16.29	13.65	9.62	5.53	2.00	6.94

Table 3.6 Durée moyenne par tranche d'âge

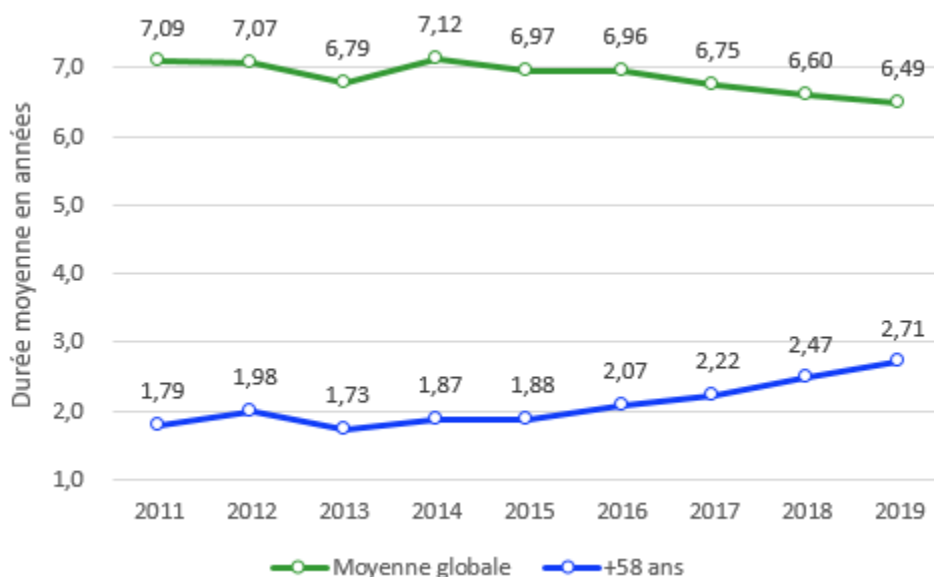


Figure 3.8 Comparaison de la durée moyenne des seniors et de l'ensemble de la population en invalidité

Passage incapacité vers invalidité

La durée passée dans l'état d'incapacité (IT) avant le passage vers l'état d'invalidité (IP) est fonction de l'âge. La table 3.7 présente le nombre de mois passés en incapacité puis le nombre d'années en invalidité pour chaque classe d'âge. En moyenne, les assurés passent 23 mois dans l'état d'incapacité avant d'être invalides. Le délai de passage se raccourcit avec l'âge et les années

de survenance(cf. annexe B.0.3), cela s'explique par les différentes mesures mises en place par la Sécurité Sociale visant à assouplir les conditions de passage en invalidité.

	TRANCHE 1	TRANCHE 2	TRANCHE 3	TRANCHE 4	TRANCHE 5	GLOBAL
IT (MOIS)	26.7	25.6	24.7	22.2	16.7	23.4
IP (ANNÉES)	15.4	12.3	9.0	5.5	2.3	7.6

Table 3.7 Passage IT - IP par tranche d'âge

3.3.4 Indemnité journalière

Le montant de l'indemnité augmente avec l'âge, les arrêts des seniors sont donc plus coûteux que le reste de la population. Cela s'explique par le fait qu'en fin de carrière les salaires soient plus élevés. Les profils étant plus ou moins différents en fonction des années, aucune tendance particulière ne se distingue par rapport à l'inflation. Les seniors sont de plus en plus nombreux, leurs arrêts sont plus longs et les coûts associés sont plus élevés : leur charge sinistre est beaucoup plus importante.

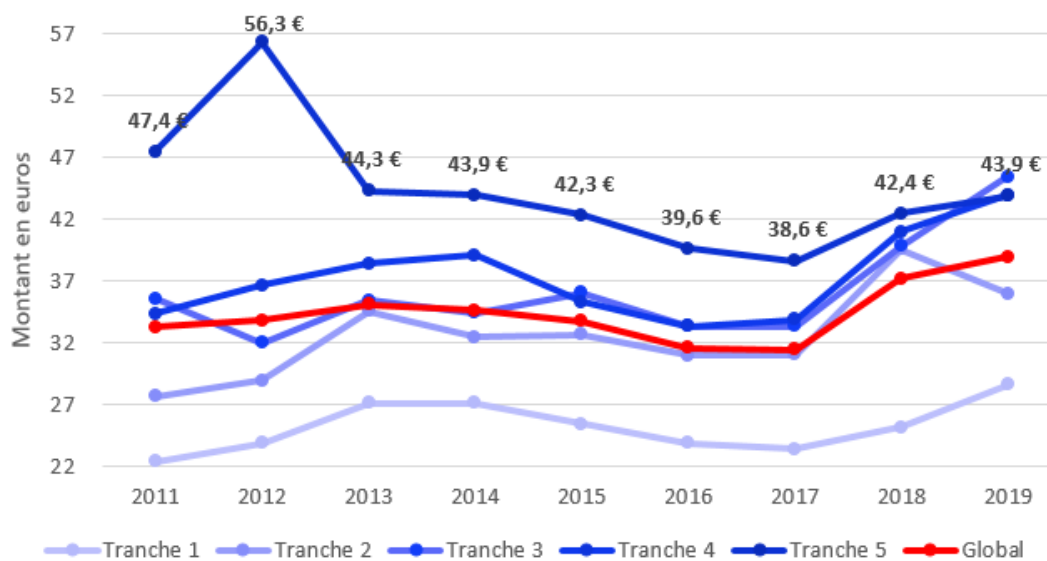


Figure 3.9 Indemnité journalière moyenne par classe d'âge et par survenance

Chapitre 4

Impacts de la réforme sur la sinistralité du portefeuille

La réforme des retraites de 2010 portant sur le recul générationnel de l'âge de départ à la retraite entraîne, mécaniquement, un vieillissement de la population active. Le portefeuille de sinistres en prévoyance collective se retrouve ainsi impacté. Dans cette partie, une méthode sera proposée pour mesurer cet impact sur la sinistralité en fonction de plusieurs indicateurs. En d'autres termes, il s'agit de répondre à la problématique suivante : en décalant l'âge de départ à la retraite d'un certain nombre de mois, combien de sinistres additionnels observe-t-on ? Quels sont la durée moyenne et le coût (prestations versées et provisions mathématiques) supplémentaires associés ?

4.1 Construction de la base d'analyse et hypothèses

L'identification des arrêts impactés par la réforme nécessite un traitement de la base avec des hypothèses définies propres aux risques étudiés (incapacité et invalidité). La construction de la base d'analyse comporte deux étapes importantes : la détermination de l'âge de départ à la retraite à la fois avant et après la réforme pour chaque assuré et l'identification de l'impact de la réforme.

4.1.1 Détermination de l'âge de départ à la retraite

La connaissance de l'âge de départ à la retraite pour chaque personne en arrêt est une information précieuse. Cependant, il est difficile de l'obtenir du fait du caractère personnel du choix de départ à la retraite et des différents dispositifs mis en place par la Sécurité Sociale et les organismes de retraite complémentaire.

Âge de départ à la retraite avant la réforme

Pour pallier à ce manque d'information, un âge de départ à la retraite théorique a été établi en fonction de l'âge à la survenance de l'arrêt et de l'âge d'ouverture des droits à la retraite avant la réforme (60 ans). En effet, **avant la réforme**, les assurés pouvaient bénéficier d'une retraite à taux plein soit à 60 ans (avec une durée de cotisation d'au moins 164 trimestres) soit à 65 ans. Ainsi, deux groupes se distinguent :

- Groupe de l'âge légal (60 ans) : assurés âgés de moins de 60 ans à la survenance de l'arrêt
- Groupe de l'âge du taux plein (65 ans) : assurés âgés de 60 ans et plus à la survenance de l'arrêt

Les assurés ayant 60 ans et plus sont donc considérés comme étant ceux n'ayant pas atteint un nombre de trimestres suffisant pour pouvoir liquider leur retraite au taux plein.

Âge de départ à la retraite après la réforme

ANNÉE DE NAISSANCE	ÂGE MINIMUM LÉGAL	ÂGE DU TAUX PLEIN
Avant 01/07/1951	60 ans	65 ans
1951	60 ans et 4 mois	65 ans et 4 mois
1952	60 ans et 9 mois	65 ans et 9 mois
1953	61 ans et 2 mois	66 ans et 2 mois
1954	61 ans et 7 mois	66 ans et 7 mois
À partir de 1955	62 ans	67 ans

Table 4.1 Âges de départ à la retraite par génération après la réforme de 2010**Exemples**

Les exemples ci-dessous aident à mieux comprendre les hypothèses de la construction de la base d'étude.

Exemple 1 : départ à l'âge d'ouverture des droits M. X, né en 1954, tombe en arrêt en 2011. Il est donc âgé de 57 ans à la survenance de son sinistre.

- Age de départ avant réforme : 60 ans
- Age de départ après réforme : 61 ans et 7 mois

Exemple 2 : départ à l'âge d'obtention du taux plein M. Y, né en 1954, tombe en arrêt en 2015. Il est donc âgé de 61 ans à la survenance de son sinistre.

- Age de départ avant réforme : 65 ans
- Age de départ après réforme : 66 ans et 7 mois

4.1.2 Identification des arrêts impactés par la réforme des retraites

La détermination des âges de départ à la retraite avant et après la réforme permet de mettre en évidence les sinistres additionnels dûs au recul de l'âge de départ à la retraite. Ces sinistres additionnels se déclinent en deux ensembles : les arrêts exclusivement dus à la réforme et les arrêts dont la durée de couverture s'est allongée.

Arrêts dus à la réforme

Il s'agit d'arrêts de travail survenus **après** l'âge de départ à la retraite avant la réforme. C'est, par exemple, une personne qui serait en incapacité à 66 ans.

- Age de départ avant réforme : 65 ans
- Age de départ après réforme : 67 ans
- Durée courue : 6 mois
- Age à la dernière vision de l'arrêt : 66.5 ans

La mise en arrêt de cet assuré fait alors suite à un allongement de la vie active induit par la réforme des retraites de 2010. Cependant, il existe d'autres profils en fonction du risque (incapacité ou invalidité) et de la génération concernée. La distance entre l'âge de l'assuré à la survenance et l'âge à la dernière vision (ou clôture) de l'arrêt permettent de les repérer dans la base. L'âge à la dernière vision de l'arrêt peut aller au delà de l'âge de départ à la retraite actuel, dans ce cas c'est l'age de retraite actuel qui est considéré. La formule utilisée est décrite ci-après et illustrée par la figure 4.1.

$$\delta = \min(a, b) - c$$

avec a = âge de départ à la retraite après la réforme, b = âge à la dernière vision de l'arrêt, c = âge à la survenance de l'arrêt.

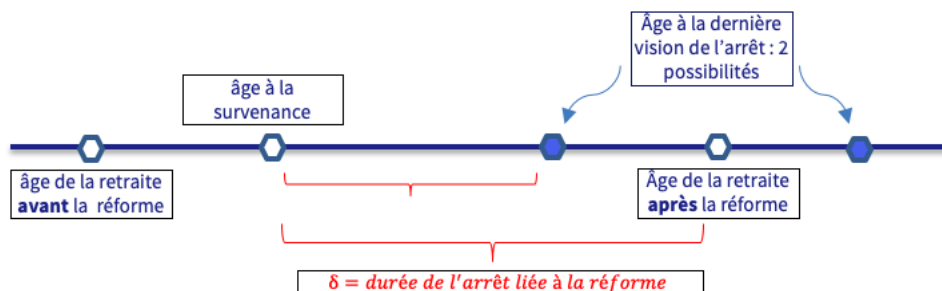


Figure 4.1 Illustration des arrêts dus à la réforme

Arrêts prolongés par la réforme

Il s'agit d'arrêts de travail survenus **avant** l'âge de départ à la retraite avant la réforme et qui se sont prolongés au delà de cet âge. C'est, par exemple, une personne qui serait en invalide à 59 ans.

- Age de départ avant réforme : 60 ans
- Age de départ après réforme : 62 ans
- Durée courue : 18 mois (1.5 ans)
- Age à la dernière vision de l'arrêt : 60.5 ans (= 59 ans + 1.5 ans)

Les 6 derniers mois (60.5 ans - 60 ans = 0.5 ans = 6 mois) de cet arrêt sont dus au recul de l'âge de départ à la retraite. Pour identifier les arrêts concernés, la distance entre l'âge de la retraite avant la réforme et l'âge de l'assuré à la dernière vision (ou clôture) de l'arrêt est calculée. L'âge à la dernière vision de l'arrêt peut aller au delà de l'âge de départ à la retraite actuel, dans ce cas c'est l'âge de retraite actuel qui est considéré. La formule utilisée est décrite ci-après et illustrée par la figure 4.2.

$$\delta = \min(a, b) - c$$

avec a = âge de départ à la retraite après la réforme, b = âge à la dernière vision de l'arrêt, c = âge de départ à la retraite avant la réforme .

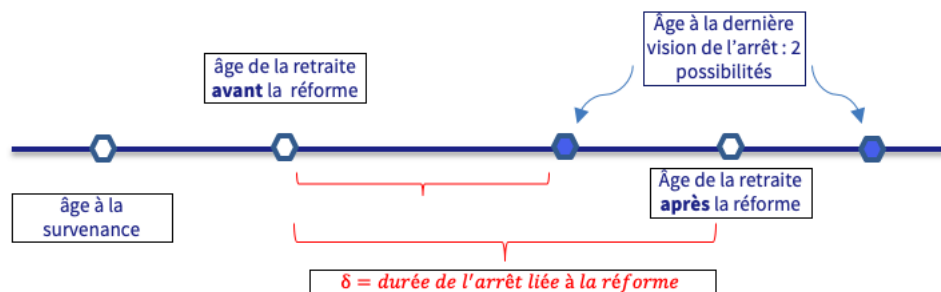


Figure 4.2 Illustration des arrêts prolongés par la réforme

Dans la base, certains arrêts sont survenus après l'âge de départ à la retraite actuel. Ceux-là sont considérés comme étant des anomalies de la base ou une volonté personnelle de certains assurés de rester en activité plus longtemps. Ils représentent moins de 1% de la base et ne sont donc pas pris en compte dans cette étude.

4.2 Présentation des résultats

La partie précédente décrit les traitements et hypothèses réalisés pour déterminer les arrêts qui sont touchés par la réforme. Il s'agit d'une part d'arrêts survenus après les âges réglementaires de départ à la retraite avant la réforme de 2010 et d'autre part d'arrêts qui se sont prolongés suite au recul de l'âge de la retraite. Les impacts sont analysés selon différents indicateurs, à savoir, le volume de sinistres, le maintien en arrêt et le coût (prestations versées et provisions mathématiques). Pour chacun de ces indicateurs il est question d'avoir une meilleure connaissance sur ce qui aurait été observé sans la réforme et ainsi pourvoir en estimer l'impact.

4.2.1 Impacts sur le volume

Le principal aspect mis en évidence dans le cadre de cette étude est le volume de sinistres. Il est important de savoir comment évolue le nombre d'arrêts travail suite à la mise en place de la réforme et ensuite pouvoir en déduire son impact. Pour chaque génération, une répartition du volume de sinistres a été faite à la fois pour l'ensemble du portefeuille de sinistres et pour les sinistres

additionnels. Pour des raisons de confidentialité, ces répartitions ne seront pas communiquées. Les proportions sont renseignées dans les tableaux 4.2 pour l'invalidité et 4.3 pour l'incapacité. La part de la réforme représente la proportion de sinistres additionnels dus à la réforme parmi l'ensemble des sinistres de la base. La part du décalage permet de mesurer l'impact de chaque décalage dans le volume de sinistres additionnels global.

GÉNÉRATION	DÉCALAGE	PART RÉFORME	PART DU DÉCALAGE
Avant 01/07/1951	Aucun	0%	0%
1951	4 mois	100%	0%
1952	9 mois	59%	0 %
1953	1 an et 2 mois	67%	3%
1954	1 an et 7 mois	68%	6%
À partir de 1955	2 ans	19%	91%
	TOTAL	20,79%	

Table 4.2 Invalidité - Proportions des sinistres additionnels par génération

Les invalides de notre base de sinistres qui sont impactés par la réforme sont âgés entre 51 et 61 ans à la survenance de leur sinistre. Parmi eux, seulement 0.41% ont 60 ans et plus. Au global, 20,79 % des invalides de la base de sinistres est due à la réforme des retraites de 2010. Le prolongement de 2 ans est celui qui impacte le plus le volume de sinistres, regroupant 91% des arrêts additionnels.

GÉNÉRATION	DÉCALAGE	PART RÉFORME	PART DU DÉCALAGE
Avant 01/07/1951	Aucun	0%	0%
1951	4 mois	100%	0%
1952	9 mois	95%	2 %
1953	1 an et 2 mois	77%	7%
1954	1 an et 7 mois	39%	10%
À partir de 1955	2 ans	1%	71%
	TOTAL	1,7%	

Table 4.3 Incapacité - Proportion de sinistres additionnels par génération

Le risque incapacité est très peu impacté par le décalage de l'âge de départ à la retraite. En effet, les arrêts de travail additionnels représentent seulement 1,7% de l'ensemble des sinistres en incapacité. Le décalage de 2 ans représente 71% de tous les arrêts additionnels dus à la réforme. Seulement 1% des incapables de la génération née à partir de 1955 sont dûs à la réforme des retraites.

4.2.2 Impacts sur les durées

Le recul de l'âge de départ à la retraite entraîne mécaniquement un prolongement de la durée d'indemnisation des arrêts de travail. C'est pourquoi la durée représente un indicateur de taille dans l'analyse des impacts de la réforme. Pour chaque décalage, les éléments suivants ont été établis :

- La durée moyenne observée : elle correspond à la différence entre la date de dernière vision (ou de clôture) et la date de survenance de l'arrêt. Elle est exprimée en nombre de mois ;
- La durée moyenne hors réforme : c'est la durée moyenne qui aurait été observée s'il n'y avait pas eu de réforme. Elle correspond alors à la différence entre la durée moyenne observée et le δ ;

- L'écart entre la durée moyenne observée et la durée moyenne théorique hors réforme ;
- La proportion du prolongement qui a été utilisée. C'est le ratio entre l'écart moyen et la durée du décalage

GÉNÉRATION	DÉCALAGE	DURÉE MOYENNE	DURÉE HORS RÉFORME	ÉCARTS	% PROLONGEMENT
Avant 01/07/1951	Aucun	-	-	-	-
1951	4 mois	21,9 m	18 m	3,9 m	99%
1952	9 mois	39,3 m	30,5 m	8,8 m	97%
1953	1 an et 2 mois	21,8 m	10,2 m	11,6 m	82%
1954	1 an et 7 mois	28,2 m	13,6 m	14,6 m	77%
À partir de 1955	2 ans	29,1 m	12,6 m	16,5 m	69%

Table 4.4 Invalidité - Résultats sur les durées moyennes par génération

L'écart entre les durées observées et la durée théorique hors réforme augmente en fonction du décalage. La consommation du prolongement décroît en fonction des générations. Une vision plus globale de la sinistralité des seniors (58 ans et plus) permet de mieux apprécier la portée de ces écarts. Les arrêts qui ne sont pas impactés par la réforme durent en moyenne 11 mois tandis que ceux qui sont impactés ou dus à la réforme durent en moyenne 28.5 mois. Leur durée moyenne au global étant de 21 mois, il est certain que la réforme contribue à l'augmentation du maintien dans l'état d'invalidité et, par extension, à celle de l'âge moyen de la population des seniors en invalidité.

GÉNÉRATION	DÉCALAGE	DURÉE MOYENNE	DUREE HORS RÉFORME	ECARTS	% PROLONGEMENT
Avant 01/07/1951	Aucun	-	-	-	-
1951	4 mois	19,0 m	16,1 m	2,9 m	72%
1952	9 mois	17,4 m	11,8 m	5,6 m	62%
1953	1 an et 2 mois	15,4 m	7,0 m	11,5 m	50%
1954	1 an et 7 mois	15,7 m	8,7 m	6,9 m	36%
À partir de 1955	2 ans	15,4 m	8,8 m	6,6 m	28%

Table 4.5 Incapacité - Résultats sur les durées moyennes par génération

L'écart entre les durées observées et la durée théorique hors réforme semble indépendant du décalage. L'utilisation du prolongement est faible pour les deux dernières générations. Pour les seniors, les arrêts qui ne sont pas impactés par la réforme durent en moyenne 5,4 mois tandis que ceux qui sont impactés ou dus à la réforme durent en moyenne 15,6 mois. La durée moyenne de l'ensemble étant de 7,3 mois, il est certain que la réforme contribue à l'augmentation du maintien dans l'état d'incapacité.

4.2.3 Impacts sur les prestations versées

L'allongement de la durée d'indemnisation implique une hausse des prestations. L'information du montant des pensions et des indemnités versées aux assurés étant disponible dans la base de données, les montants théoriques hors réforme en ont été déduits. Les tableaux 4.6 et 4.7 présentent l'évolution entre les sommes effectivement versées et celles qui auraient été payées s'il n'y avait pas eu de réforme.

GÉNÉRATION	DÉCALAGE	TAUX D'ÉVOLUTION
Avant 01/07/1951	Aucun	-
1951	4 mois	+117%
1952	9 mois	+125%
1953	1 an et 2 mois	+232%
1954	1 an et 7 mois	+201%
À partir de 1955	2 ans	+237%

Table 4.6 Invalidité - Résultats sur les prestations moyennes versées par génération

La réforme contribue à la hausse des pensions versées au titre de l'invalidité. L'étendue de ces taux d'évolutions est d'autant plus appréciable en fonction de la classe d'âge. La pension d'invalidité moyenne versée aux seniors en invalidité est de 20 968 €, essentiellement portée par les invalides impactés par la réforme dont la pension moyenne vaut 28 349 € tandis que ceux qui ne sont pas impactés ont une pension moyenne de 11 104 €. L'allongement de la durée de l'arrêt peut impliquer un changement de catégorie d'invalidité et donc une hausse du niveau des prestations. Cependant, l'assureur perçoit des primes pendant la période de cotisation pour compenser cette hausse des prestations. Le niveau de ces primes n'aurait pas été le même avant la réforme.

GÉNÉRATION	DÉCALAGE	TAUX D'ÉVOLUTION
Avant 01/07/1951	Aucun	-
1951	4 mois	+116%
1952	9 mois	+145%
1953	1 an et 2 mois	+174%
1954	1 an et 7 mois	+177%
À partir de 1955	2 ans	+182%

Table 4.7 Incapacité - Résultats sur les prestations moyennes versées par génération

La réforme contribue à l'augmentation des indemnités versées au titre de l'incapacité. L'indemnité moyenne versée aux seniors en incapacité est de 9 653 €, essentiellement portée par ceux qui sont impactés par la réforme dont la pension moyenne vaut 20 188 € tandis que ceux qui ne sont pas impactés ont une pension moyenne de 7 273 €.

4.3 Provisions mathématiques

Cette partie présente l'impact du changement réglementaire de l'âge de départ à la retraite sur les provisions mathématiques en incapacité et en invalidité. Pour ce faire, un provisionnement tête par tête est réalisé pour tous les sinistres en cours à la date du 30 mars 2020. Dans un premier temps les exigences réglementaires du provisionnement sont présentés de même que les différentes provisions calculées. Ensuite les résultats sont présentés pour l'invalidité. Pour finir, une projection d'impacts liés au projet de réforme de 2020 sera intégrée.

4.3.1 Exigences réglementaires

En France, la réglementation prudentielle du code des assurances impose aux organismes assureurs la constitution de provisions techniques permettant de faire face aux engagements pris vis à vis des assurés. Pour les garanties incapacité/invalidité, deux principales catégories de provisions doivent être considérées :

- Les provisions mathématiques (PM), qui représentent les engagements probables de l'assureur vis-à-vis des assurés. Elles concernent les sinistres survenus et déclarés à l'assureur. Cette provision est donc constituée au titre des sinistres en cours et connus.
- Les provisions pour sinistres à payer (PSAP) qui font référence aux prestations dues par l'assureur mais qui n'ont pas encore été payées.

A la date d'inventaire, ces deux catégories de provisions se différencient sur deux critères :

- La survenance du sinistre est connue ou non connue

- La provision est relative à des prestations passées non réglées à la date d'inventaire ou à des prestations probables futures.

Le cadre réglementaire

Il existe deux textes qui définissent le contexte juridique relatif au provisionnement pour les garanties traitant des risques incapacité et invalidité :

Loi Numéro 89-1009 du 31 Décembre 1989 Elle renforce les garanties offertes aux personnes assurées contre certains risques, dite loi Evin s'applique au contrat professionnel. Elle oblige les institutions de prévoyance, les mutuelles et les entreprises d'assurance à provisionner les sinistres en cours pour la couverture des risques décès, des risques portant sur l'intégrité physique de la personne, des risques incapacité de travail, invalidité, et chômage. Elle traite ainsi de l'évaluation des provisions techniques au niveau atteint, du maintien des garanties décès aux invalides et aux incapables (décret du 17 Juillet 2001) et des revalorisations futures des indemnités et des rentes d'invalidité.

Arrêté du 28 mars 1996 La loi Evin est complétée par l'arrêté du 28 mars 1996 qui fixe les règles de provisionnement des garanties incapacité/invalidité à travers différents articles référencés dans :

- Le code des assurances à l'article A331-22
- Le code de la sécurité sociale à l'article R731-4
- Le code de la mutualité dans l'article 1bis de l'arrêté du 28 juillet 1988

Les provisions mathématiques :

Selon l'article A331-22 du code des assurances, trois types de provisions mathématiques doivent être constituées au titre des incapacités en cours et des invalidités en cours :

- La provision pour maintien en incapacité : Provision correspondant aux prestations d'incapacité de travail à verser au titre des incapacités en cours à la date d'inventaire.

Elle doit permettre de couvrir l'engagement de l'assureur lié au maintien de l'assuré en incapacité.

- La provision pour rente en attente : Provision correspondant au passage d'un individu de l'état incapable à l'état invalide, elle est appelée provision pour passage ou provision pour rente en attente. Elle concerne les rentes d'invalidité pouvant intervenir au titre des individus en incapacité susceptibles de passer en invalidité.
- La provision pour maintien en invalidité : Provision correspondant aux prestations d'invalidité en cours à la date d'inventaire. Elle doit permettre de couvrir les engagements de l'assureur liés au maintien de l'individu en invalidité.

Ces provisions sont généralement évaluées « tête par tête » c'est-à-dire pour chaque assuré, mais ce mode de calcul n'est pas imposé par le code des assurances. Pour le calcul de ces provisions réglementaires, le code des assurances apporte cependant des précisions sur les paramètres à utiliser :

- Le taux d'actualisation doit être inférieur à 75 % du taux moyen des emprunts d'Etat français calculé sur les 24 derniers mois et ne pas dépasser 4,5 %.
- Les lois de maintien et de passage utilisées sont celles établies par le BCAC¹ ou certifiées par un actuaire indépendant de l'organisme assureur.

Les provisions pour sinistre à payer

La provision pour sinistres à payer concerne aussi bien les risques incapacité qu'invalidité et regroupe deux types de sinistres :

- Les sinistres connus à la date d'inventaire dont les prestations dues n'ont pas encore été réglées au moment de la clôture de l'exercice comptable.
- Les sinistres survenus mais non connus par l'assureur à la date d'inventaire car ils n'ont pas encore été déclarés. On les appelle les sinistres tardifs ou IBNR (Incurred But Not

1. BCAC : Bureau Commun d'Assurance Collective

Reported).

A la différence des provisions mathématiques présentées ci-dessus, la provision pour sinistre à payer est en pratique constituée au global pour un groupe homogène de contrats et non plus « tête par tête ». Cette provision peut se calculer à l'aide de méthodes basées sur l'estimation de la charge ultime telles que les triangles de liquidation. Dans le cadre de ce mémoire, seules les provisions mathématiques seront calculées.

4.3.2 Calcul des provisions mathématiques

Pour le calcul des provisions mathématiques correspondant au risque incapacité et invalidité, il est nécessaire de disposer de la loi de maintien en incapacité et en invalidité et de la loi de passage d'un état à l'autre. Ainsi, trois tables réglementaires ont été construites par le BCAC en 1993 à partir de données sur les portefeuilles des 4 plus grands assureurs français à cette période (AGF, AXA, GAN et UAP) en n'effectuant aucune distinction sur le sexe de l'assuré ou la catégorie professionnelle de celui-ci (cadre ou non cadre). Ces lois ont été rendues obligatoires en 1996 pour tous les organismes assureurs puis ont été modifiées par l'arrêté du 24 décembre 2010, suite à la réforme sur les retraites (allongeant ainsi la durée de versement des prestations). Cependant, au lieu d'utiliser ces tables réglementaires, l'entreprise d'assurances peut construire ses propres tables d'expérience, basées sur l'historique de ses portefeuilles et certifiées par un actuaire indépendant. Il existe trois types de tables : les tables de maintien en incapacité, les tables de passage de l'incapacité à l'invalidité et les tables de maintien en invalidité. Comme précisé précédemment, trois types de provisions mathématiques doivent être constituées. Ainsi, la provision à constituer au titre des individus incapables à la date d'inventaire qui correspond à la somme de la provision pour maintien en incapacité se distingue de la provision pour invalidité en attente, et la provision à constituer au titre des invalidités en cours.

La provision pour maintien en incapacité

Par définition, la provision pour maintien en incapacité est obtenue comme la somme des montants actualisés probables des indemnités à verser au titre de l'incapacité.

$$PM^{\text{incap}}(x, \text{anc}) = \sum_{k=\text{anc}}^{36-1} \frac{1}{2 * l_{x,\text{anc}}^{\text{incap}}} * \left[l_{x,k}^{\text{incap}} * \left(\frac{1}{1+i_m} \right)^{k-\text{anc}} + l_{x,k+1}^{\text{incap}} * \left(\frac{1}{1+i_m} \right)^{k+1-\text{anc}} \right]$$

Avec : x , l'âge de l'assuré à l'entrée en incapacité

anc , l'ancienneté (en mois) de l'assuré au moment de l'évaluation de la provision

i_m , le taux d'actualisation mensuel et $\frac{1}{1+i_m}$ est appelé facteur d'actualisation.

$l_{x,k}^{\text{incap}}$, le nombre d'individus en incapacité d'ancienneté k mois et entré à l'âge x dans cet état.

On note que le passage d'un facteur d'actualisation annuel à mensuel est obtenu par :

$$\frac{1}{(1+i_m)^{(k-\text{anc})}} = \frac{1}{(1+i_a)^{(k-\text{anc})/12}}$$

,

Avec i_a le taux technique annuel et i_m le taux technique mensuel.

Dans l'hypothèse où aucune revalorisation de l'indemnité mensuelle n'est effectuée, le montant de PM obtenu par la formule ci-dessus est à multiplier par le montant de l'indemnité versée à l'assuré. Sinon, chaque flux devra être multiplié par le montant de l'indemnité revalorisée.

La provision pour invalidité en attente

Cette provision vient s'ajouter à la provision pour maintien en incapacité à constituer au titre des incapacités en cours et a pour but d'évaluer les engagements de l'assureur en cas de passage de l'assuré d'un état d'incapacité de travail à l'invalidité. Son calcul se base sur les probabilités de passage et sur le montant des rentes d'invalidité à verser en cas de passage.

$$\begin{aligned}
 & PM_{\text{incap}}^{\text{inval}}(x, \text{anc}) \\
 &= \sum_{k=1}^{36-\text{anc}-1} \frac{1}{2 \cdot l_{x,atjc}^{\text{incap}}} \cdot \left[s_{x,\text{anc}+k} + \left(\frac{1}{1+i_a} \right)^{k/12} + s_{x,\text{anc}+k+1} \cdot \left(\frac{1}{1+i_a} \right)^{(k+1)/12} \right] \\
 & \cdot P \left(x + \frac{\text{anc} + 6 + k}{12}, 0 \right)
 \end{aligned}$$

Où : $P \left(x + \frac{\text{anc} + 6 + k}{12}, 0 \right)$ est déterminé par interpolation linéaire entre $P^{\text{inv}} \left(\text{Ent} \left(x + \frac{\text{anc} + 6 + k}{12} \right), 0 \right)$ et $P^{\text{inv}} \left(\text{Ent} \left(x + \frac{\text{anc} + 6 + k}{12} \right) + 1, 0 \right)$ avec Ent la fonction partie entière. $P^{\text{inv}}(x, 0)$ désigne la provision pour une rente d'invalidité d'un euro correspondant à un individu entré à l'âge x en invalidité et aucune ancienneté. $s_{x,j}$ est le nombre de personnes entrées en incapacité à l'âge x et passé en incapacité le $j^{\text{ième}}$ mois.

La provision pour maintien en invalidité

La provision pour maintien en invalidité se calcule sur le même principe que la provision pour maintien en incapacité. Cependant le pas de temps utilisé est annuel.

$$PM^{\text{inval}}(x, \text{anc}) = \sum_{k=\text{anc}}^{62-x-1} \frac{1}{2 * l_{x,\text{anc}}^{\text{inval}}} * \left[l_{x,k}^{\text{inval}} * \left(\frac{1}{1+i_a} \right)^{k-\text{anc}} + l_{x,k+1}^{\text{inval}} * \left(\frac{1}{1+i_a} \right)^{k+1-\text{anc}} \right]$$

Les mêmes notations sont utilisées que précédemment. $l_{x,k}^{\text{inval}}$ représente le nombre d'individus en invalidité, d'ancienneté k années et entré à l'âge x dans cet état. Le taux technique i_a utilisé pour l'actualisation des flux est ici annuel. Pour un pas de temps mensuel, la formule sera adaptée et les flux seront actualisés à l'aide du taux mensuel équivalent.

4.3.3 Effets du décalage d'âge de départ à la retraite sur la PM en invalidité

Cette partie expose l'impact de l'augmentation de l'âge de départ à la retraite sur les provisions mathématiques pour l'ensemble du portefeuille sinistré. Les PM ont été calculées pour tous les

sinistres en cours à la date du 31 mars 2020 regroupant ainsi uniquement des assurés nés à partir de 1955 qui représentent donc la génération concernée par le décalage de 2 ans. Deux groupes se distinguent :

- Assurés âgés de moins de 60 ans : leurs arrêts se sont prolongés suite à la réforme
- Assurés ayant 60 ans et plus : leurs arrêts sont dus à la réforme.

Assurés âgés de moins de 60 ans

La figure 4.3 présente, pour chaque survenance, le montant de la PM avec un âge de départ à la retraite à 60 ans (bleu foncé) et la PM avec un âge de départ à 62 ans (bleu clair). Le tableau 4.8 fournit le nombre de dossiers ainsi que les taux d'évolutions des PM entre le départ à 60 ans et celui à 62 ans pour chaque survenance. Le nombre de dossiers exact et les montants associés ne seront pas affichés pour des raisons de confidentialité.

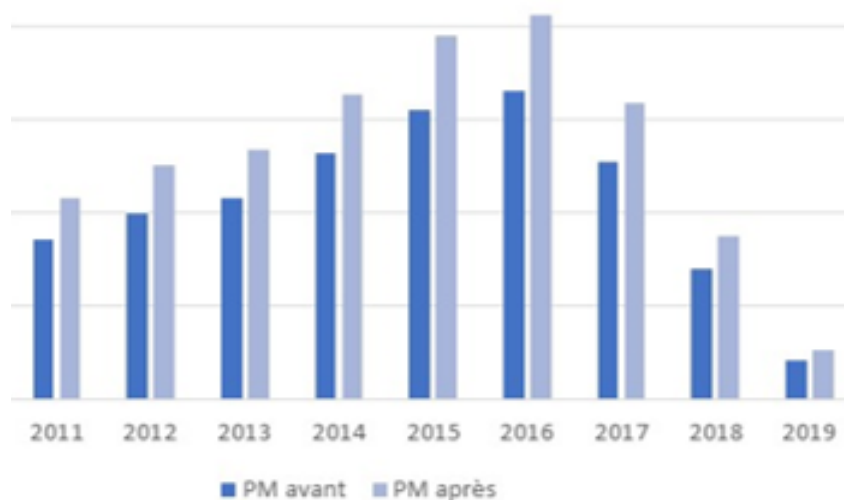


Figure 4.3 PM avant/après réforme par survenance

Impacts de la réforme sur la sinistralité du portefeuille

SURVENANCE	2011	2012	2013	2014	2015	2016	2017	2018	2019	TOTAL
STOCK (EN MILLIERS)	1,2	1,3	1,5	1,8	2,1	2,3	1,8	1,0	0,2	13
% D'ÉVOLUTION DE LA PM	+26%	+25%	+24%	+24%	+25%	+24%	+25%	+25%	+23%	+25%

Table 4.8 Evolution de la PM invalidité par année de survenance

Les provisions mathématiques se sont alourdies suite au passage de l'âge de départ à la retraite de 60 à 62 ans. Sur l'ensemble du portefeuille de sinistres, une augmentation de +25% a été observée. Un regard plus approfondi permet de mettre en évidence les tranches d'âges pour lesquelles l'augmentation de la PM est la plus significative. En effet la tranche d'âge composée d'invalides âgés de plus de 58 ans (tranche 5) est la plus touchée par le décalage de l'âge de la retraite, avec une hausse de +319,6% comme indiqué sur la figure 4.4. Sur les tranches 1 et 2 composées de personnes plus jeunes (14 - 41 ans), l'écart est d'un peu moins de 10%. En définitive, les personnes âgées de 50 ans et plus (tranches 4 et 5) absorbent la hausse de 25% signalée par le tableau 4.8 de manière intrinsèque. Le taux d'évolution des PM avant/après réforme est croissant avec l'âge.

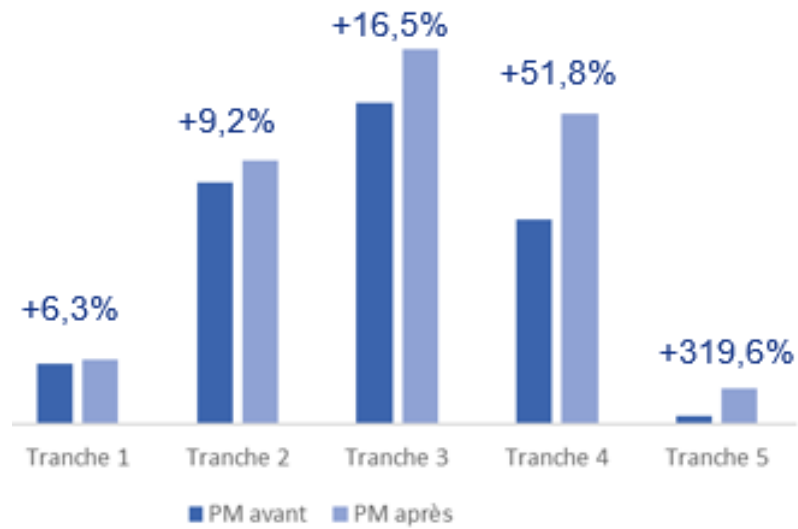


Figure 4.4 PM avant/après réforme par tranches d'âge à la survenance

Assurés âgés de 60 ans et plus

Les invalides âgés entre 60 et 62 ans à la survenance de leur arrêt représentent 13% de l'ensemble des invalides de la base de provisionnement. La PM associée représente, au global, 0,45% de l'ensemble de la PM calculée avec un âge de départ à la retraite à 62 ans. L'impact est donc faible et cela s'explique en partie par le fait que les survenances 2018 et 2019 ne soient pas encore suffisamment consolidées. En effet, certains invalides sont encore inconnus de la base d'étude du fait du délai de passage de l'incapacité vers l'invalidité.

SURVENANCE	NOMBRE DE DOSSIERS	% DE LA PM GLOBALE
2015	9	0,04%
2016	64	0,68%
2017	106	1,41%
2018	61	1,55%
2019	12	1,00%
Total	252	0,45%

Table 4.9 Impact de la PM des 60 ans et plus dans la PM globale par survenance

En définitive, la réforme des retraites de 2010 à travers le relèvement de l'âge légal de départ à la retraite et du taux plein a eu des conséquences sur la sinistralité de notre portefeuille de prévoyance collective. Les principales garanties étudiées, l'invalidité et l'incapacité, sont impactées en terme de volume, de maintien et de coût. Il a été constaté un volume de sinistres supplémentaires représentant +20,79% en invalidité et seulement +1.7% en incapacité. Le risque invalidité est donc très impacté en terme d'incidence mais aussi en terme de maintien. En effet, les provisions mathématiques calculées avec le nouvel âge réglementaire subissent une hausse de +25% en invalidité tandis qu'un surplus de 1.3% sur la provision d'invalidité en attente est observé en incapacité. L'invalidité étant un risque court (3 ans maximum), l'impact est de moindre envergure.

4.3.4 Projet de réforme 2020

Préconisation du rapport Delevoye

Après 18 mois de concertations, Jean-Paul DELEVOYE [8], Haut-commissaire à la réforme des retraites, a remis son rapport sur la refonte du système des retraites le 18 juillet 2019. En 2018, l'âge de départ effectif à la retraite était de 63,4 ans en moyenne (hors départs anticipés). Cela s'explique par l'augmentation progressive de la durée d'assurance requise pour bénéficier d'une

retraite à taux plein et par les carrières qui débutent de plus en plus tard (en moyenne 22 ans). Le système actuel incite donc à différer sa date de départ en retraite pour obtenir le taux plein. Le futur système de retraite prévoit deux âges : l'âge légal toujours fixé à 62 ans et la création d'un âge « pivot » à 64 ans en dessous duquel le montant des pensions de retraite ferait l'objet d'une « décote » de 5% par an (équivalent du dispositif de taux plein actuel qui est alors supprimé). En cas de départ après l'âge pivot le montant de la pension sera majoré d'une surcote de 5% par an. Selon le haut commissaire à la réforme des retraites, le palier de 64 ans correspond à l'âge moyen projeté de départ à taux plein en 2025. À titre de comparaison, il était de 63 ans en 2018. Cet âge pourrait ensuite évoluer en fonction de l'espérance de vie des Français. La mise en place de ce nouvel âge de départ à la retraite se ferait entre 2022 et 2027 selon le calendrier suivant :

Année de naissance	ÂGE PIVOT	ANNÉE
Avant 01/07/1961	62 ans	2022
Après 01/07/1961	62 ans et 4 mois	2023
1962	62 ans et 9 mois	2024
1963	63 ans et 2 mois	2025
1964	63 ans et 7 mois	2026
À partir de 1965	64 ans	2027

Table 4.10 Proposition de calendrier de mise en place de l'âge pivot

Projection d'impacts

Dans cette partie on souhaite estimer l'effet d'une future réforme sur la sinistralité du portefeuille. Le calendrier proposé pour la mise en place de l'âge pivot se décline sous le même schéma que le décalage générationnel de l'âge de départ à la retraite induit par la réforme de 2010. L'estimation de l'impact de la nouvelle réforme se fera selon deux axes :

1. Évaluer l'impact sur le stock de sinistres encore ouverts pendant la période de mise en place du système (entre 2022 et 2027). En effet, pour chaque arrêt encore en cours, l'âge pivot

est déterminé selon la date de naissance de l'assuré et le calendrier 4.10. La PM est ensuite calculée avec les âges de départ suivants : 60 ans (hors réformes), 62 ans (suite à la réforme de 2010), entre 62 et 64 ans selon la date de naissance (projet de réforme de 2020). Cette façon de procéder permet d'isoler l'impact associé à chaque nouvelle réglementation. Le graphique 4.5 représente l'évolution, pour chaque année, de la proportion que représente chaque réforme dans la provision mathématique calculée suite à la mise en place de l'âge pivot. La courbe grise représente le niveau de PM qui aurait été calculée si aucune réforme n'avait eu lieu. La courbe de couleur bleu clair montre que l'effet de la réforme des retraites sur la provision décroît au fil du temps, absorbée par l'impact de l'âge pivot (courbe de couleur bleu foncé).

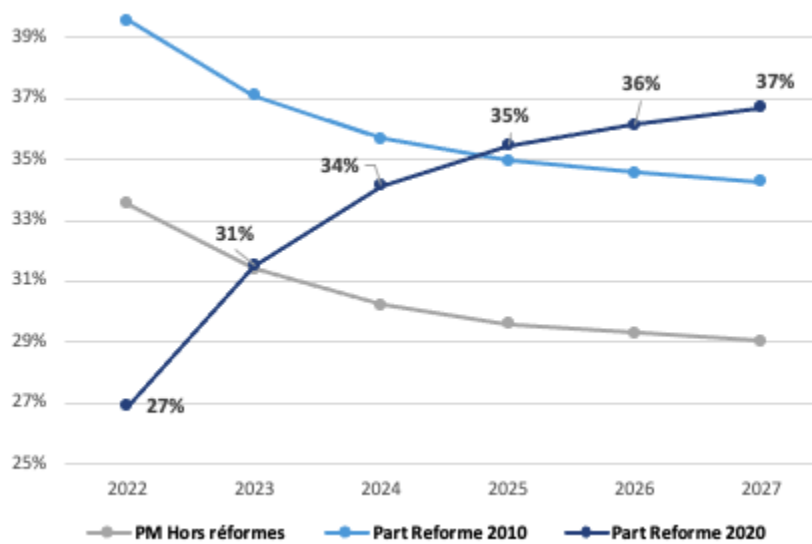


Figure 4.5 Évolution de l'impact de chaque réforme sur la PM du stock de sinistres en invalidité

2. Estimation du volume de sinistrés de la nouvelle population à risque. Il s'agit de simuler le volume de sinistres additionnels qui serait observé suite à la mise en place de l'âge pivot. Le volume de sinistres de la survenance 2019 pour les assurés âgés entre 57 et 64 ans de la base DSN est utilisé en appliquant les correctifs suivants :

- Le taux de passage de la DSN à MRM
- Les évolutions de la population active projetées par l'Insee

- La probabilité moyenne de passage de l'incapacité vers l'invalidité

C'est ainsi, en se basant sur les impacts connus de la réforme de 2010 corrigés des facteurs démographiques et , que les effets de l'instauration de l'âge pivot sur la sinistralité a été mis en exergue. Pour des raisons de confidentialité, les résultats détaillés ne seront pas communiqués. Au global, la proposition de loi du rapport Delevoye qui prévoit la création d'un âge d'équilibre occasionnera +20,85% d'invalides supplémentaires et seulement 1,84% pour l'incapacité. En ce qui concerne la durée, la prestation moyenne et le montant des provisions mathématiques, une référence aux conclusions retenues dans le cadre de la réforme de 2010 reste valable. A titre d'exemple, il sera raisonnable de considérer qu'un décalage futur de l'âge de départ à la retraite de 4 mois entraînerait un allongement de la durée moyenne de 3,9 mois pour l'invalidité et 2,9 mois pour l'incapacité (cf. tableaux 4.4 et 4.5).

RISQUE	RÉFORME 2010	RÉFORME 2020
INVALIDITÉ	+20,79%	+20,85%
INCAPACITÉ	+1,7%	+1,84%

Table 4.11 Proportion de sinistres additionnels liés à chaque réforme

	2022	2023	2024	2025	2026	2027
PM Hors réforme	34%	31%	30%	30%	29%	29%
Réforme 2010	40%	37%	36%	35%	35%	34%
Réforme 2020	27%	31%	34%	35%	36%	37%

Table 4.12 Décomposition de la PM des arrêts en cours

Chapitre 5

Modélisation des taux d'incidence

La partie précédente a permis de mettre en exergue les impacts du décalage d'âge de départ à la retraite sur la sinistralité pour les garanties incapacité et invalidité. Au regard des différentes conclusions retenues et du contexte socio-économique actuel il est nécessaire de pouvoir approfondir cette étude dans le but de prédire les effets d'une nouvelle réforme. Pour ce faire, dans un premier temps il faut comprendre dans quelles mesures les effets de la réforme des retraites sur la sinistralité actuelle des seniors peuvent être mis en évidence. L'objectif de cette partie est de décrire, expliquer et prédire un phénomène qui dépend du temps : les taux d'incidence en incapacité. Nous souhaitons construire un modèle déduit de nos observations et qui soit pertinent. Dans un premier temps, les motivations au tour du choix des taux d'incidence en incapacité ainsi que la méthode et les hypothèses utilisées pour leur calcul seront exposés. Ensuite certains éléments mathématiques qui seront utilisés dans la suite seront décrits.

5.1 Choix de la modélisation

5.1.1 Quantité modélisée

Dans le cadre de cette modélisation, deux quantités ont été retenues pour répondre à notre problématique. L'idée initiale était d'essayer d'expliquer et de prédire le volume de sinistres suite à

la mise en place de la réforme. Cependant, le volume n'est pas suffisamment représentatif de l'évolution du risque tandis que l'incidence l'est parce qu'elle prend en compte l'exposition au risque. La table 5.1 montre renseigne les différents avantages et inconvénients inhérents à l'utilisation de chacune de ces quantités.

	INCIDENCE DE SINISTRES	VOLUME DE SINISTRES
AVANTAGES	Historique limité (2018 - 2019)	Historique complet (2011 - 2019)
INCONVÉNIENTS	<ul style="list-style-type: none"> • Exposition prise en compte • Transposable à un autre périmètre 	Limité au périmètre MRM

Table 5.1 Choix de la quantité à modéliser

L'incidence de sinistres est donc la quantité à modéliser. Cependant, une difficulté subsiste quant à la modélisation des taux d'incidence. L'exposition au risque est une donnée désormais disponible grâce aux déclarations sociales et nominatives (DSN). Ce dispositif étant récent, seules les années 2018 et 2019 sont disponibles. Dans le cadre de notre problématique, il est nécessaire de disposer d'un historique d'au moins deux années avant l'entrée en vigueur du décalage d'âge (2016). Les survenances dont l'exposition est inconnue, 2014 à 2017, seront donc estimées sur la base de celle de 2018 à laquelle les correctifs suivants seront appliqués :

- Les taux d'évolution des cotisations entre l'année N et l'année 2018 calculés sur les contrats en effet à la fois en 2018 et en l'année N.
- L'évolution de la proportion de seniors dans la population active entre 2018 et l'année N ;
- L'évolution de l'inflation pour tenir compte de l'évolution des salaires entre les différentes années

La formule suivante a été utilisée pour calculer l'exposition mensuelle pour la survenance 2018.

$$\text{EXPOSITION}_{2018,T} = \frac{\min(\alpha, \delta/T/2018) - \max(\beta, 01/T/2018) + 1}{\delta}$$

avec T un mois donné, α =date de fin d'exposition hors arrêt travail, β =date de début d'exposition hors arrêt travail, δ =nombre de jours du mois.

L'exposition est obtenue avec le calcul ci-dessous :

$$\text{INCIDENCE}_T = \frac{\text{NOMBRE D'ARRÊTS DE TRAVAIL SURVENUS AU MOIS T}}{\text{EXPOSITION}_{N,T}}$$

Le choix du risque étudié a porté sur l'incapacité bien que l'invalidité soit le plus impactée. Les sinistres en invalidité sont quasiment tous une aggravation de l'incapacité (93%). Ainsi, à partir de l'estimation qui sera faite, des facteurs d'aggravation du risque à estimer pourront permettre d'en déduire une modélisation de l'invalidité. Pour la modélisation, seule la génération des personnes nées à partir de 1955 a été retenue, celles concernées par le passage de l'âge de départ à la retraite de 60 à 62 ans. En effet, elles constituent 90% des sinistres additionnels dus à la réforme.

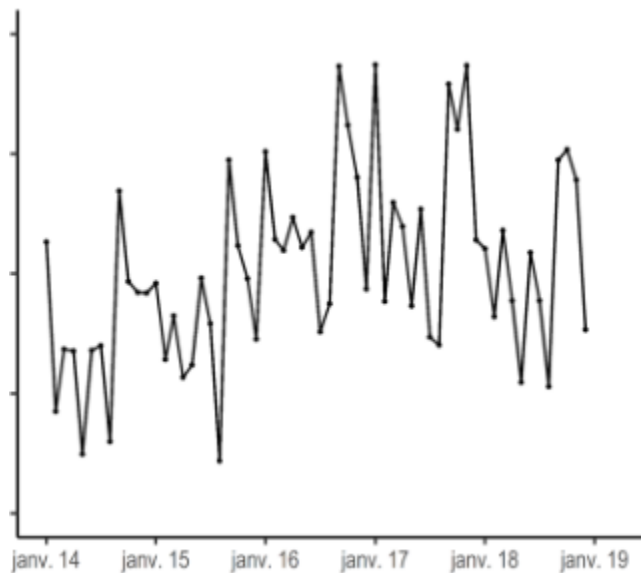


Figure 5.1 Présentation de la série d'incidences

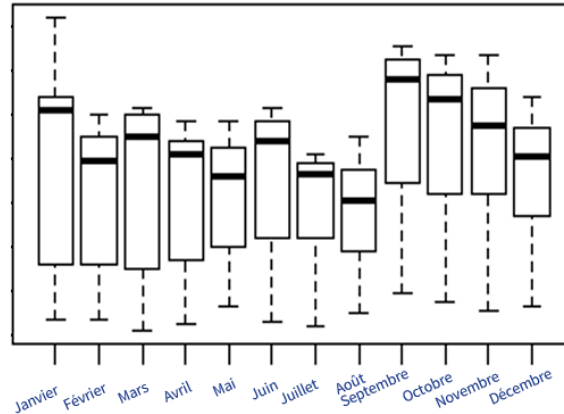


Figure 5.2 Boxplot des incidences mensuelles

5.1.2 Modèles utilisés

Pour modéliser l'évolution des séries d'incidence, nous allons utiliser les séries temporelles. Une série temporelle (ou encore une série chronologique) est une suite finie (x_1, \dots, x_t) de données représentant l'évolution d'une quantité spécifique au cours du temps. L'intérêt porte sur la modélisation et à la prédiction des séries temporelles perturbées par des interventions connues telles que des modifications tarifaires, l'introduction d'un nouveau produit sur un marché, une grève, etc.

La modélisation d'une série temporelle à partir de son historique en utilisant la classe des modèles linéaires ARIMA suppose la stabilité structurelle au cours du temps du modèle retenu et de ses coefficients. Dès lors, si l'on est en présence d'une structure évolutive et/ou de paramètres variant au cours du temps, les prédictions issues du modèle univarié seront sérieusement affectées. Lorsqu'on essaie de modéliser des séries chronologiques à caractère économique, on est amené à tenir compte d'événements de nature diverse, extérieurs au modèle, qui viennent perturber les séries. L'effet de ces événements se fait sentir soit par la présence d'un ou plusieurs points dits aberrants ("outliers"), qui occasionnent une rupture ponctuelle dans la série, soit par un changement sensible dans l'évolution de la série.

Prendre en compte l'effet d'un événement extérieur sur une série permet d'améliorer la modélisation de cette série, mais surtout de **mesurer l'impact de cet événement, ce qui peut être**

intéressant pour quantifier l'effet d'une mesure. Cette mesure de l'impact pourra également être utilisée pour établir des séries corrigées des valeurs aberrantes, avec lesquelles il est souvent préférable de travailler. Le but étant bien sûr d'améliorer la modélisation d'une série pour pouvoir fournir de meilleures prévisions sur cette série.

BOX et TIAO [5] sont à l'origine de la théorie dite de "l'analyse d'intervention" qui permet de prendre en compte différentes interventions extérieures au modèle lors de la modélisation d'une série chronologique. L'apport de l'analyse d'intervention à une modélisation de série de type, par exemple, SARIMA (BOX et al. [6]), se situe au niveau de l'information disponible au praticien pour modéliser cette série. En effet, l'approche de BOX et al. utilise simplement l'information quantitative contenue dans les données, alors que l'analyse d'intervention permet d'ajouter de manière additive une information de type qualitatif, par le biais de variables binaires exogènes. De nombreux auteurs se sont intéressés par la suite à l'analyse d'intervention et ses applications et ont complété les travaux de BOX et TIAO. Le but de notre étude est de mettre en évidence l'intérêt de la théorie de l'analyse d'intervention pour la modélisation des séries d'incidence en incapacité pour les assurés de 58 ans et plus.

5.2 Les processus ARMA

Le modèle ARMA est une des représentations les plus utilisées pour les processus temporels stationnaires. Cette représentation consiste en l'adjonction d'une composante auto-régressive d'ordre fini (AR) et d'une composante moyenne mobile d'ordre fini (MA). Quelques définitions utiles pour étudier les séries temporelles sont données avant de décliner les différentes étapes de la modélisation.

5.2.1 Rappels mathématiques

Soit $(\Omega, \mathcal{A}, \mathbb{P})$, un espace probabilisé avec Ω l'espace des évènements, \mathcal{A} une tribu adaptée à Ω et \mathbb{P} une mesure de probabilité définie sur \mathcal{A} . Soient (F, \mathcal{F}) et (Ω', \mathcal{A}') deux espaces mesurables.

Un processus stochastique est une application Y

$$Y: \Omega \times F \rightarrow \Omega'$$

$$(\omega, t) \mapsto Y(\omega, t).$$

$Y(\omega, t)$ est encore notée $Y_t(\omega)$ et est telle que pour t fixé appartenant à F , Y est une variable aléatoire sur (Ω, \mathcal{A}) . Par extension, on écrira un processus sous la forme d'une suite de variables aléatoires indicées par t , notée $(Y_t, t \in F)$ ou plus simplement (Y_t) . La loi du processus est l'image \mathbb{P}^Y de \mathbb{P} par Y . Lorsque $(\Omega'$ est \mathbb{R} (respectivement \mathbb{R}^m), le processus est dit réel unidimensionnel ou univarié (respectivement multidimensionnel de dimension m ou m -varié); si $(\Omega'$ est fini ou dénombrable, on parle de processus à valeurs discrètes. Une série temporelle $(Y_t)_{t=1, \dots, T}$ est la réalisation particulière d'un processus stochastique $(Y_t, t \in F)$. Dans cette étude, les séries que étudiées seront toutes réelles et univariées. Les processus considérés sont discrets, c'est-à-dire où T est de la forme $1, \dots, n$. Les différents types de modélisation appliqués aux séries temporelles sont celles expliquées ci-dessous.

Définition 1 *Un processus temporel aléatoire $(x_i; t \in \mathbb{Z})$ est dit strictement ou fortement stationnaire si $\forall t_1 < t_2 < \dots < t_n$, tel que $t_i \in \mathbb{Z}$ et pour tout temps $h \in \mathbb{Z}$ avec $t_i + h \in \mathbb{Z}, \forall i = 1, \dots, n$ la suite $(x_{t_1+h}, \dots, x_{t_n+h})$ a la même loi de probabilité que la suite $(x_{t_1}, \dots, x_{t_n})$*

Cette propriété très forte est très difficile à vérifier. C'est pourquoi, dans la pratique, on se limite généralement à requérir la stationnarité faible du processus.

Définition 2 *Un processus $(x_i; t \in \mathbb{Z})$ est dit stationnaire au sens faible, ou stationnaire d'ordre deux si les trois conditions suivantes sont satisfaites :*

1. $\forall t \in \mathbb{Z}, \mathbb{E}[x_t^2] < \infty$
2. $\forall t \in \mathbb{Z}, \mathbb{E}[x_t] = \mu$, indépendant de t
3. $\forall (t, h) \in \mathbb{Z}^2, \text{Cov}(x_t, x_{t+h}) = \mathbb{E}[(x_{t+h} - \mu)(x_t - \mu)] = \delta(h)$, indépendant de t

Parmi la classe des processus stationnaires, il existe des processus particuliers que sont les processus bruit blanc. Un bruit blanc est un processus aléatoire de moyenne nulle, de variance constante et dont la covariance de ses termes pris deux à deux est nulle. Ces processus constituent la base élémentaire des processus ARMA comme expliqué par la suite.

Définition 3 *Un processus $(\epsilon_t; t \in \mathbb{Z})$ est un bruit blanc faible s'il satisfait les deux conditions suivantes :*

1. $\forall t \in \mathbb{Z}, \mathbb{E}[\epsilon_t] = 0$

2. $\forall t \in \mathbb{Z}, \delta(h) = \mathbb{E}[\epsilon_t \epsilon_{t-h}] = \begin{cases} \sigma^2 & \text{si } h = 0 \\ 0 & \forall h \neq 0 \end{cases}$

On parle d'un bruit blanc fort si en plus $(\epsilon_t; t \in \mathbb{Z})$ est une suite de variables aléatoires réelles indépendantes. On parle de bruit blanc gaussien lorsque les résidus respectent les deux conditions et suivent une loi normale.

Définition 4 *Soit ϵ_t un bruit blanc faible de variance σ^2 . On dit qu'un processus $(X_t)_{t \in \mathbb{Z}}$ est un processus autorégressif ou encore processus AR (AutoRegressive) d'ordre p , noté $AR(p)$ si :*

$(X_t)_{t \in \mathbb{Z}}$ est stationnaire

$$\forall t \in \mathbb{Z} : X_t = \sum_{i=1}^p \varphi_i X_{t-i} + \epsilon_t, \text{ où } (\varphi_1, \dots, \varphi_p) \in \mathbb{R}^q \text{ et } \varphi_p \neq 0$$

Définition 5 *Soit ϵ_t un bruit blanc faible de variance σ^2 . On dit qu'un processus $(X_t)_{t \in \mathbb{Z}}$ est un processus moyenne mobile ou encore processus MA (Moving Average) d'ordre q , noté $MA(q)$ si :*

$$\forall t \in \mathbb{Z} : X_t = \epsilon_t + \sum_{i=1}^q \theta_i \epsilon_{t-i}$$

où $(\theta_1, \dots, \theta_q) \in \mathbb{R}^q$ et $\theta_q \neq 0$

Définition 6 *Soit ϵ_t un bruit blanc faible de variance σ^2 . On dit qu'un processus $(X_t)_{t \in \mathbb{Z}}$ est un processus ARMA d'ordre (p, q) , noté $ARMA(p, q)$ si :*

1. $(X_t)_{t \in \mathbb{Z}}$ est stationnaire

$$2. \quad X_t = \sum_{i=1}^p \varphi_i X_{t-i} + \sum_{i=1}^q \theta_i \epsilon_{t-i} + \epsilon_t$$

Le modèle $ARMA(p, q)$ se décompose alors en trois parties : un processus AR d'ordre p , fonction de ses p dernières valeurs, un processus MA d'ordre q , fonction des q derniers termes d'erreurs et, enfin, un bruit blanc ϵ_t , qui correspond à l'erreur du pas t .

Définition 7 La fonction d'autocovariance notée δ est définie par :

$$\begin{aligned} \delta: F \times F &\rightarrow \Omega' \\ (s, t) &\mapsto \delta(s, s+t) = \text{cov}(Y_s, Y_{s+t}) = \delta(t) \end{aligned}$$

Pour Y un processus réel, δ est une fonction paire. La fonction d'autocorrélation $\rho(\cdot)$ est une normalisation de la fonction d'auto-covariance et satisfait la condition supplémentaire $\rho(0) = 1$

L'estimateur de la fonction d'autocorrélation et l'estimateur de la fonction d'autocorrélation partielle permettront de faire une première sélection parmi les nombreux modèles stationnaires susceptibles de représenter la dépendance des données. Par exemple, une ACF (Autocorrelation Function) empirique très proche de 0 suggère qu'un modèle adapté aux données pourrait être un bruit blanc fort. On considère le modèle sous sa forme additive : $Y_t = f(t) + s(t) + \epsilon_t; t \in T$ où f = tendance : fonction déterministe s = saisonnalité : fonction déterministe et périodique ϵ_t = résidu : processus stochastique centré stationnaire.

On considère le modèle sous sa forme additive :

$$Y_t = f(t) + s(t) + \epsilon_t; t \in T$$

où :

f = est une fonction déterministe représentant la tendance

s = est une fonction déterministe et périodique représentant la saisonnalité

ϵ_t = est un processus stochastique centré stationnaire représentant le résidu

On veut estimer et éliminer f et s , de façon à ne garder que l'observation de la partie stationnaire (ϵ_t)

Dans tout ce qui suit, on supposera le processus Y centré. Il suffira dans le cas où Y est non centré de moyenne μ ($E(Y_t) = \mu; \forall t$), de poser $Z_t = Y_t - \mu$ et d'appliquer nos résultats à $Z = (Z_t; t \in T)$

Deux types de méthodes permettent de régler ce problème :

1. Estimer la tendance et la saisonnalité puis les soustraire de la série considérée
2. Différencier la série : c'est la méthode de Box-Jenkins.

La deuxième approche est celle retenue pour l'étude. La démarche adoptée pour mettre en oeuvre les modèles ARMA se décline en 5 étapes :

1. Stationnarisation et suppression de la saisonnalité
2. Identification a priori de modèles potentiels
3. Estimation et vérification des modèles potentiels
4. Validation et choix du modèle
5. Prévisions

5.2.2 Mise en oeuvre de la modélisation

Stationnarisation et suppression de la saisonnalité

Les processus temporels présentent généralement une tendance et/ou une saisonnalité et ne sont donc pas stationnaires. C'est le cas de la série d'intérêt. Le graphique 5.1 indique la présence de ces phénomènes dans notre série, celle-ci ne peut donc pas être modélisée par un processus ARMA. Pour stationnariser le processus, une différence a été appliquée. La série brute n'est pas modélisée mais plutôt la série différenciée en tendance et/ou en saisonnalité.

L'opérateur retard (noté L pour Lag ou B pour Backward) est défini comme suit :

Soit un processus stochastique $(Y_t, t \in Z)$. Alors $BY_t = Y_{t-1}, \forall t \in Z$. Cet opérateur définit une application qui, à toute variable Y_t associe la variable retardée Y_{t-1} .

On introduit l'opérateur $\nabla = 1 - B$ et de façon générale $\nabla^j = \nabla(\nabla^{j-1})$, avec $\nabla^0 = I$. Si on applique ∇^k à une fonction à tendance polynômiale d'ordre k , on la réduit à une constante. En pratique, k est petit.

Pour les séries avec une saisonnalité de période s , nous utilisons l'opérateur $\nabla_s = 1 - B^s : \nabla_s Y_t = Y_t - Y_{t-s}$. En appliquant ∇_s à notre modèle (*), on obtient :

$$\nabla_s (f_t + s_t + \varepsilon_t) = f_t - f_{t-s} + \varepsilon_t - \varepsilon_{t-s}$$

Ce qui ramène donc à une estimation de la tendance. Ces transformations permettent ainsi d'obtenir une série stationnaire, en particulier sans tendance ni saisonnalité.

Identification a priori de modèles potentiels

L'identification conduit à déterminer les ordres p et q des parties autorégressives et moyenne mobile du modèle et les ordres de différenciation éventuelles. Un processus stationnaire peut être modélisé par la classe des $ARMA(p, q)$. La meilleure façon de procéder consiste à étudier l'autocorrélogramme partiel (PACF) afin de déterminer la partie autorégressive, puis l'autocorrélogramme simple (ACF) pour identifier la partie moyenne mobile. Pour cela, il suffit de déterminer les coefficients de corrélation qui sont significativement différents de zéro.

MODÈLE \ CRITÈRE	ACF	PACF
AR(p)	Décroissance rapide vers 0	Significative jusqu'au p^{ieme} retard
MA(q)	Significative jusqu'au q^{ieme} retard	Décroissance rapide vers 0
ARMA(p,q)	Décroissance rapide vers 0	Décroissance rapide vers 0

Table 5.2 Critères de détermination des modèles en fonction de l'ACF et du PACF

Estimation et vérification des modèles potentiels

L'estimation des paramètres des modèles potentiels $ARMA(p, q)$ (p et q supposés connus) peut se faire à l'aide des méthodes classiques telles que la méthode du maximum de vraisemblance et celle des moindres carrés. Avant de passer à la modélisation, il convient de s'assurer de la stationnarité de la série. Pour ce faire, on pourra effectuer des tests de racine unité qui permettront de confirmer que la série est suffisamment différenciée. Les modèles potentiels sont ensuite vérifiés en deux étapes. Dans un premier temps, la significativité des paramètres est contrôlée à l'aide de tests d'hypothèse. Par exemple, pour le terme d'ordre p de la partie AR, on peut effectuer le test d'hypothèse suivant :

$$\begin{cases} H_0 : \text{le processus est un } ARMA(p, q) \\ H_1 : \text{le processus est un } ARMA(p-1, q) \end{cases}$$

En utilisant la statistique de test

$$t = \frac{\hat{\varphi}_p}{\sqrt{\text{V}(\hat{\varphi}_p)}}$$

le test de Student permet de rejeter H_0 au niveau 5% si $|t|$ est supérieur à 1,96 .

Ensuite, on vérifie la blancheur du résidu à l'aide du test de Ljung-Box : Soit $(\epsilon_t)_{t \in \mathbb{Z}}$ un processus stationnaire.

$$\begin{cases} H_0 : (\epsilon_t)_{t \in \mathbb{Z}} \text{ est un bruit blanc} \\ H_1 : (\epsilon_t)_{t \in \mathbb{Z}} \text{ n'est pas un bruit blanc} \end{cases}$$

Si on dispose de $(\epsilon_1, \dots, \epsilon_T)$, on considère la statistique

$$Q_k = T(T+2) \sum_{h=1}^k \frac{\hat{\rho}^2(h)}{T-h}$$

Sous H_0 , elle suit une loi Khi-deux à k degrés de liberté. L'hypothèse H_0 est alors rejetée au niveau $\alpha = 5\%$ si $Q_k > X_k^2(1-0,05)$, où $X_k^2(1-0,05)$ désigne le quantile d'ordre $1-\alpha$ d'une loi Khi-deux à k degrés de liberté.

Choix définitif du modèle

Pour le choix définitif parmi les modèles ayant passé le cap de la vérification, il existe deux types de critères, qui ne sont pas très bien fondés statistiquement (au sens mathématique) mais qui

donnent de bons résultats empiriquement. Il s'agit des critères d'information. La méthode consiste à choisir le modèle en se basant sur une mesure de l'écart entre la vraie loi (inconnue) et celle du modèle proposé. La mesure habituellement utilisée est l'information de Kullback. On choisit le modèle donnant la plus petite valeur de l'estimation de l'information de Kullback. Les critères d'estimation basés sur l'information de Kullback les plus répandus sont :

- Le critère d'information d'Akaike (AIC), basé sur la statistique

$$\text{AIC}(p, q) = 2(p + q + 1) - 2 \ln \left(L_X \left(\hat{p}, \hat{q}, \hat{\sigma}^2 \right) \right),$$

où L_X est la vraisemblance de Y lorsque le processus est gaussien.

- Le critère bayésien de Schwarz (BIC), utilisé dans le cas d'un processus ARMA(p,q) centré causal et inversible. Il est noté SBC (Schwarz Bayesian Criterion) ou BIC (Bayesian modification of the AIC criterion). Il a été introduit par Akaike pour corriger la tendance de l'AIC à surestimer le nombre de paramètres. Il est basé sur la statistique :

$$\text{SBC}(p, q) = \text{BIC}(p, q) = 2(p + q + 1) \ln n - 2 \ln \left(L_X \left(\hat{p}, \hat{q}, \hat{\sigma}^2 \right) \right),$$

où n est le nombre d'observations.

Le choix se portera sur le couple (p, q) qui minimise ces statistiques. Pour des modèles ARIMA saisonniers, on obtient des modèles multiplicatifs du type :

$$(1 - B)^d (1 - B^s)^D \phi(B) \Phi(B^s) Y_t = C + \theta(B) \Theta(B^s) \varepsilon_t \quad (5.1)$$

où :

B est l'opérateur défini sur les variables aléatoires de la manière suivante : $B(X_t) = X_{t-1}$ et $B^b(X_t) = X_{t-b}$ pour $b > 1$,

$\phi(B)$ est un polynôme en B de degré p tel que $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$,

$\Phi(B)$ est un polynôme en B de degré P tel que $\Phi(B) = 1 - \Phi_1 B - \Phi_2 B^2 - \dots - \Phi_P B^P$,

$\theta(B)$ est un polynôme en B de degré q tel que $\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q$,

$\Theta(B)$ est un polynôme en B de degré Q tel que $\Theta(B) = 1 + \Theta_1 B + \Theta_2 B^2 + \dots + \Theta_Q B^Q$,

$(\varepsilon_D)_{v \in Z}$ est un processus bruit blanc gaussien de variance σ^2 ,

S représente la saisonnalité de la série,

C est une constante et $(d, D) \in \mathbb{N}^2$

Les différences d'ordre d et les différences saisonnières d'ordre D doivent permettre de rendre la suite de variables aléatoires faiblement stationnaire.

Un processus (Y_t) satisfaisant cette équation est appelé processus $SARIMA[(p, d, q)(P, D, Q)]_s$.

Validation

Une fois qu'un modèle a été ajusté, la dernière étape est le diagnostic du modèle. Cette vérification est effectuée en étudiant les graphiques d'autocorrélation des résidus pour voir si une structure supplémentaire (grandes valeurs de corrélation) peut être trouvée. Si toutes les autocorrélations et les autocorrélations partielles sont faibles, le modèle est considéré comme adéquat et des prévisions sont générées. Si certaines des autocorrélations sont importantes, les valeurs de p et/ou q sont ajustées et le modèle est ré-estimé. Ce processus de vérification des résidus et d'ajustement des valeurs de p et q se poursuit jusqu'à ce que les résidus résultants ne contiennent plus de structure supplémentaire. Une fois qu'un modèle approprié est sélectionné, le processus peut être utilisé pour générer des prévisions. La validation du modèle se fait selon deux aspects :

1. L'analyse de la qualité statistique du modèle estimé (tests sur les estimateurs, qualité globale de l'ajustement)
2. L'analyse des résidus d'estimation : les résidus doivent être des réalisations de bruits blancs pour que le modèle soit correcte La modélisation de la série du bruit estimé permet de détecter :

Indépendance des résidus : Ils sont donc considérés comme comme les observations de variables aléatoires indépendantes, la moyenne et la variance peuvent donc être estimées

Dépendance entre les résidus : Les bruits sont modélisés par une série stationnaire plus complexe (grâce à la théorie des processus stationnaires), qui tiennent compte de ce fac-

teur. Cette dépendance permet en particulier d'utiliser les observations passées pour prédire des valeurs futures.

Des tests paramétriques et non paramétriques permettent de vérifier la validité du modèle à travers le contrôle des résidus. Ils doivent être des valeurs observées de variables aléatoires indépendantes et identiquement distribuées. Le détail de ces tests est donné en annexe 4.

5.3 Analyse d'intervention

Cette partie permettra de montrer comment l'analyse des interventions essaie de répondre à la question : "étant donné une ou plusieurs interventions, y a-t-il un changement dans l'évolution de la série temporelle depuis la date connue d'occurrence de cette (ces) intervention(s) et si oui, quelles sont les caractéristiques et l'amplitude de ce(s) changement(s) ?". Dans un premier les filtres linéaires seront décrits avant de présenter les modèles d'intervention.

5.3.1 Filtres linéaires

Pour définir les modèles d'intervention de BOX et TIAO, nous présentons préalablement quelques notions fondamentales sur les filtres linéaires réalisables dissipatifs.

Dans la suite nous appellerons filtre H une transformation linéaire invariante au cours du temps transformant une série temporelle $\xi = (\xi_t, t \in Z)$ à temps discret, appelée entrée, en une sortie $y = (y_t, t \in Z)$ à temps discret. Lorsque l'ensemble des entrées admissibles du filtre H est un espace vectoriel, on a classiquement (cf. KOOPMANS) :

$$y_t = \sum_{-\infty}^{\infty} h_j \xi_{t-j}, \quad t \in Z$$

où la suite $(h_j, j \in Z)$ représente la réponse impulsionnelle du filtre. La série formelle

$$h(z) = \sum_{-\infty}^{\infty} h_j z^j$$

s'appelle *fonction de transfert du filtre*.

Si $h_j = 0$ pour $j < 0$, la sortie au temps t dépend de l'entrée seulement pour les temps $s \leq t$ i.e., du présent et du passé de ξ . Un filtre présentant cette caractéristique est dit réalisable (ou causal).

On dit que le filtre est dissipatif lorsque la réponse impulsionnelle est de carré sommable. Enfin un filtre est stable si pour toute entrée bornée la sortie est bornée. Un filtre est stable si et seulement si

$$\sum_{j=-\infty}^{+\infty} |h_j| < \infty$$

Compte tenu des définitions données ci-dessus, la fonction de transfert d'un filtre réalisable dissipatif s'écrit :

$$h(z) = \sum_{j=0}^{\infty} h_j z^j$$

avec

$$\sum_{j=0}^{\infty} |h_j|^2 < \infty \quad (5.2)$$

La condition 5.2 implique que la fonction de transfert se définit comme une fonction analytique dans tout le disque ouvert $D = \{z \in C : |z| < 1\}$.

La réponse fréquentielle h d'un filtre réalisable dissipatif est la fonction de $L_c^2[-\pi, \pi]$ définie comme la transformée de Fourier discrète de la réponse impulsionnelle $(h_j, j \in N)$

$$h(\lambda) = \sum_0^{\infty} h_j e^{-i\lambda j}, \quad -\pi < \lambda \leq \pi \quad (5.3)$$

En considérant (5.3) comme un développement en série de Fourier on peut alors chercher la condition pour qu'une fonction h ait un développement en série de Fourier de coefficients de Fourier h_n tels que :

(i) $h_n = 0$ pour tout $n < 0$

(ii) $\sum_0^{\infty} |h_n|^2 < \infty$ H^2

Une telle condition peut se trouver en utilisant l'espace de Hardy. L'espace de Hardy H^2 se définit comme l'ensemble des classes de fonctions analytiques dans le disque ouvert $D = \{z \in C : |z| < 1\}$,

de la forme

$$h(z) = \sum_0^{\infty} h_j z^j$$

satisfaisant à

$$\sup_{0 \leq r < 1} \int_{-\pi}^{\pi} |h(re^{-i\lambda})|^2 d\lambda < \infty \quad (5.4)$$

La relation de Parseval permet de montrer que la condition 5.4) est équivalente à la condition 5.2 Ce qui implique que tout filtre réalisable dissipatif appartient à H^2 . On remarquera qu'une fraction rationnelle appartient A^2 si et seulement si ses poles sont à l'extérieur du disque unité fermé,

$$\bar{D} = \{z \in \mathbf{C} : |z| \leq 1\}$$

De plus, on peut vérifier que si $f \in H^2$ alors $\hat{f}(\lambda) = f(e^{-i\lambda})$ est limité au sens de L^2 et p.p. de $f(re^{-i\lambda})$ quand r tend vers 1 par valeurs inférieures et que $\sum_0^{\infty} h_j e^{-i\lambda j}$ est le développement en série de Fourier de $\hat{f}(\lambda)$.

Compte tenu des résultats précédents, il est clair que la condition recherchée est que h appartienne à l'espace H^2 .

En outre on peut montrer que l'application qui associe à toute fonction h de H^2 sa valeur au bord ∂D du disque unité D (i.e., le cercle unité $T = \{z \in \mathbf{C} : |z| = 1\}$) c'est-à-dire \hat{h} de $L^2(T)^{1)}$ est un isomorphisme d'espace de Hilbert de H^2 sur le sous espace hilbertien $L_+^2(T)$ de $L^2(T)$ des fonctions g de $L^2(T)$ tels que les coefficients de Fourier $\hat{g}_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} g(e^{-i\lambda}) e^{ni\lambda} d\lambda = 0$ pour tout $n < 0$.

Cette isométrie permet donc d'identifier H^2 à $L_+^2(T)$ en munissant H^2 du produit scalaire

$$\begin{aligned} (h_1, h_2)_{H^2} &= (\hat{h}_1, \hat{h}_2)_{L_+^2} = \int_{-\pi}^{\pi} h_1(e^{-i\lambda}) \bar{h}_2(e^{-i\lambda}) d\lambda \\ &= \sum_0^{\infty} h_j^{(1)} \bar{h}_j^{(2)} \end{aligned}$$

L'intérêt d'une telle identification réside dans le fait que toute fonction f non identiquement nulle appartenant à H^2 admet la factorisation unique :

$$f(z) = \rho B(z)F(z)S(z),$$

où $\rho = e^{ia}$ avec $a = \arg(f/B)(0)$

$$B(z) = z^p \prod_{n=1}^{\infty} \left[\frac{a_n - z}{1 - \bar{a}_n z} \cdot \frac{\bar{a}_n}{|a_n|} \right]^{p_n}, \quad 0 < |a_n| < 1, \quad \sum (1 - |a_n|) < \infty$$

,

$$F(z) = \exp \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{e^{-i\lambda} + z}{e^{-i\lambda} - z} \log |f(e^{-i\lambda})| d\lambda \right\}$$

,

$$S(z) = \exp \left\{ - \int_{-\pi}^{\pi} \frac{e^{-i\lambda} + z}{e^{-i\lambda} - z} d\mu(\lambda) \right\}$$

où $\{a_n\}$ est l'ensemble des zéros de f dans D , d'ordre respectivement p_n , avec $a_n \neq 0$ p l'ordre du zéro à l'origine, B le produit de BLASCHKE associé à ces zéros et où μ est une mesure positive bornée sur $[-\pi, \pi]$ singulière par rapport a la mesure de Lebesgue (i.e., $d\mu/d\lambda = 0 \forall \lambda$ p.p. sur $[-\pi, \pi]$)

Ainsi, tout filtre réalisable et dissipatif pourra être factorisé de façon unique. De plus, si le filtre f est une fraction rationnelle dont les pôles sont l'extérieur de \bar{D} donc appartenant à H^2 , la composante singulière disparaît.

Lorsque l'on cherche à approximer un filtre discret réalisable dont la réponse impulsionnelle est une suite infinie, il est souvent intéressant d'utiliser la sous classe des filtres récurrents auto régressif-moyenne mobiles (ARMA). Un filtre de cette sous classe transforme l'entrée ξ en une sortie y de la forme

$$y_t = \sum_{j=1}^t (-\delta_j) y_{t-j} + \sum_{k=0}^s \omega_k \xi_{t-k} \quad (5.5)$$

L'équation différentielle (5.5) peut s'écrire sous forme plus compacte

$$\delta(B)y_t = \omega(B)\xi_t$$

où

$$\delta(B) = \sum_{j=0}^r \delta_j B^j, \delta_0 = 1$$

et

$$\omega(B) = \sum_{k=0}^s \omega_k B^k$$

où B est l'opérateur de retard

Le filtre ARMA (r, s) a pour fonction de transfert la fraction rationnelle

$$h(z) = \frac{\omega(z)}{\delta(z)}$$

et pour réponse impulsionnelle la fraction

$$\hat{h}(\lambda) = \frac{\omega(e^{-i\lambda})}{\delta(e^{-i\lambda})} = \frac{\sum_{k=0}^s \omega_k e^{-i\lambda k}}{\sum_{j=0}^r \delta_j e^{-i\lambda j}}$$

On peut montrer (cf. KOOPMANS) qu'une condition suffisante pour que le filtre ARMA soit réalisable et dissipatif est que le polynôme $\delta(z) = 0$ n'ait pas de racines dans le disque fermé $\bar{D} = \{z : |z| \leq 1\}$ ou $\delta(z) = \sum_{j=0}^r \delta_j z^j$; $\delta_0 = 1$

5.3.2 Présentation du modèle d'intervention

Nous allons maintenant faire une présentation théorique du modèle d'intervention de BOX et TIAO. On note $(Y_t)_{t \in \mathbb{Z}}$ la suite de variables aléatoires à modéliser, perturbée par une intervention extérieure. Le modèle d'intervention proposé par BOX et al. se présente alors sous la forme suivante :

$$Y_t = C + \frac{\omega(B)B^b}{\delta(B)} \xi_t + N_t \quad (5.6)$$

où $(N_t)_{t \in \mathbb{Z}}$ est supposé suivre un modèle SARIMA sans constante de la forme (??)

$\omega(B)$ est un polynôme en B de degré s tel que $\omega(B) = 1 - \omega_1 B - \omega_2 B^2 - \dots - \omega_s B^s$,

$\delta(B)$ est un polynôme en B de degré r tel que $\delta(B) = 1 - \delta_1 B - \delta_2 B^2 - \dots - \delta_r B^r$,

b est un entier qui représente un retard à déterminer.

La fonction déterministe $\delta^{-1}(B)\omega(B)B^b\xi_t$ représente l'effet de l'intervention qui vient s'ajouter de manière additive au bruit $(N_t)_{t \in \mathbb{Z}}$, elle est appelée *fonction d'intervention*.

Dans l'équation 5.6 la suite de variables aléatoires $(\xi_t)_{t \in \mathbb{Z}}$ représente l'effet d'une intervention extérieure à la date T, mis sous la forme d'une variable déterministe qui prend pour valeur 1 ou 0 suivant la présence ou l'absence de l'intervention. Cette variable est en général modélisée par deux classes de fonctions :

- une fonction en forme de saut :

$$\xi_t = S_t^{(T)} = \begin{cases} 0 & \text{si } t < T \\ 1 & \text{si } t \geq T \end{cases}$$

- une fonction en forme d'impulsion :

$$\xi_t = P_t^{(T)} = \begin{cases} 0 & \text{si } t \neq T \\ 1 & \text{si } t = T \end{cases}$$

On remarque cependant que grâce à l'égalité suivante : $(1 - B)S_t^{(T)} = P_t^{(T)}$, on peut toujours passer d'un saut à une impulsion.

Plus généralement, la série chronologique peut être perturbée par k interventions de natures différentes. Avec les notations précédentes, le modèle d'intervention (5.6) a alors une représentation plus générale donnée par :

$$Y_t = C + \sum_{j=1}^k \frac{\omega_j(B)B^{b_j}}{\delta_j(B)} \xi_t^{(T_j)} + N_t \quad (5.7)$$

où, pour $j = 1, \dots, k$:

$\omega_j(B)$ est un polynôme en B de degré l_j ,

$\delta_j(B)$ est un polynôme en B de degré r_j et b_j est un entier qui représente un retard à déterminer.

La classe des modèles d'intervention considérés par BOX et TIAO est donc constituée des filtres dissipatifs et réalisables telle que la réponse impulsionnelle associée à chacun de ces filtres est une fraction rationnelle en $e^{-i\lambda}$. On peut remarquer aussi que le bruit N_t éventuellement transformé

par le filtre $(1 - B)^d$ et plus généralement par le filtre $(1 - B)^d(1 - B^s)^D$ (où s est la saisonnalité du processus et D un entier positif ou nul) est la sortie d'un filtre réalisable et dissipatif dont la réponse impulsionnelle est une fraction rationnelle en $e^{-i\lambda}$.

Une hypothèse fondamentale lors de l'utilisation de l'analyse d'intervention est que la structure du modèle, par exemple SARIMA, soit la même avant et après l'intervention. Ainsi, après avoir déterminé la date d'intervention, on fixe alors les deux sous-ensembles de données correspondant à l'évolution du processus avant et après l'intervention. On ajuste ensuite le même modèle sur chacun de ces deux sous-ensembles. Dans notre cadre, puisque l'intérêt porte sur les processus linéaires, il faut ajuster un processus SARIMA à l'aide des outils classiques que sont les fonctions d'autocorrélation et d'autocorrélation partielle. En ce qui concerne la forme de la fonction d'intervention, il n'existe pas de méthode automatique fiable permettant de la déterminer. Cependant BOX et TIAO (1975) ont proposé différents types de fonctions permettant de s'adapter à la forme graphique que prend la série, suite à l'effet de l'intervention extérieure, d'où l'importance d'une analyse graphique ou géométrique de la série à étudier.

Cette analyse graphique nécessite donc une approche locale de la série qui s'éloigne de l'analyse souvent globale utilisée quand on fait une modélisation paramétrique d'un processus. Dans le paragraphe suivant, les types de fonctions d'intervention utilisés en pratique seront exposés.

5.3.3 Fonctions d'intervention type

On se propose de présenter dans ce paragraphe trois types classiques de fonctions d'intervention et leurs illustrations.

Modèle avec intervention ponctuelle

Supposons que l'on observe Y_1, \dots, Y_t, \dots une trajectoire. On suppose qu'elle est la réalisation d'un processus SARIMA $(Y_t)_{t \in \mathbb{Z}}$. Supposons d'autre part qu'à la date unique $t = T$ l'évolution de cette trajectoire soit perturbée par une intervention extérieure, c'est l'exemple de la présence d'un point aberrant dans la série, que l'on a schématisé par le graphique 5.3.

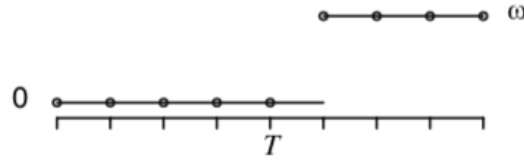


Figure 5.3 Illustration de l'intervention ponctuelle

Le modèle d'intervention s'écrit alors sous la forme suivante :

$$Y_t = C + \omega B S_t^{(T)} + N_t \quad (5.8)$$

où $(N_t)_{t \in \mathbb{Z}}$ suit un processus SARIMA défini par (5.1) sans constante et où $S_t^{(T)}$ est un saut au temps $t=T$. Le paramètre ω représente l'impact de l'intervention extérieure sur la série. Son estimation permet de donner une mesure de cet impact. Dans la littérature anglo-saxonne ce modèle est appelé de type AO (additive outlier).

Modèle avec intervention à effet rémanent

Nous supposons maintenant que l'intervention au temps $t = T$ vient perturber l'évolution de la série avec un effet qui décroît progressivement dans le temps, de façon exponentielle. Cette intervention est schématisée par le graphique 5.4.

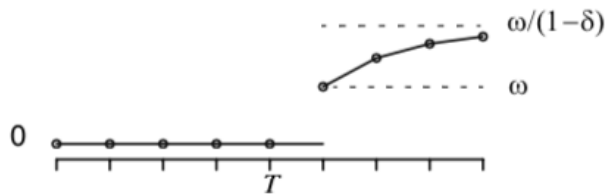


Figure 5.4 Illustration de l'Intervention à effet rémanent

Le modèle d'intervention s'écrit alors sous la forme suivante :

$$Y_t = C + \frac{\omega B}{1 - \delta B} S_t^{(T)} + N_t \quad (5.9)$$

où $(N_t)_{t \in \mathbb{Z}}$ suit un processus SARIMA défini par (5.1) sans constante et où $S_t^{(T)}$ est un saut au temps $t=T$. Le paramètre ω représente l'impact de l'intervention extérieure sur la série (son

estimation fournira une mesure de cet impact) et δ est un paramètre strictement compris entre 0 et 1 qui mesure la vitesse de croissance de l'effet de l'intervention. En clair, si δ est proche de 0, cela signifie que l'impact est quasiment ponctuel (on se ramène alors au cas précédent) et si δ est proche de 1, cela signifie que l'impact se prolonge dans le temps. Dans la littérature anglo-saxonne ce modèle est appelé de type TC (temporary or transient change).

Modèle avec intervention à effet rémanent et changement de niveau

Nous supposons maintenant que l'intervention au temps $t = T$ vient perturber l'évolution de la série avec un effet rémanent suivi d'un changement de niveau. Cette intervention est schématisée par le graphique 5.5.

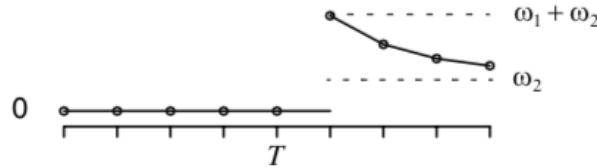


Figure 5.5 Illustration de l'intervention à effet rémanent et changement de niveau

Le modèle d'intervention s'écrit alors sous la forme suivante :

$$Y_t = C + \left(\frac{\omega_1}{1 - \delta B} + \frac{\omega_2}{1 - B} \right) BS_t^{(T)} + N_t \quad (5.10)$$

où $(N_t)_{t \in \mathbb{Z}}$ suit un processus SARIMA défini par (5.1) sans constante et où $S_t^{(T)}$ est un saut au temps $t=T$. Le paramètre ω_1 représente la différence de niveau consécutive au choc, ω_2 sera donc positif si le niveau observé après le choc est supérieur au niveau observé avant le choc, et négatif sinon. La somme de ω_1 et ω_2 représente l'impact de l'intervention extérieure sur la série, son estimation fournira une mesure de l'impact de l'intervention. De même que précédemment, δ mesure la vitesse de décroissance de l'effet de l'intervention. Dans la littérature anglo-saxonne ce modèle est appelé de type LS (level shift).

5.4 Application aux séries d'incidence

5.4.1 Résultats de la modélisation ARMA

Les modèles ARMA s'appliquent à des séries stationnaires. Les figures 5.1 et 5.2 montrent que la série d'incidences n'est pas stationnaire. En effet, les mouvements de baisse et de hausse des valeurs indiquent une tendance et les changements cycliques au cours du temps témoignent d'une saisonnalité annuelle (de période 12).

Stationnarisation et identification du modèle

La série a été différenciée pour la rendre stationnaire. L'ACF et le PACF permettent d'identifier des modèles potentiels comme indiqué par le tableau 5.2). Les graphiques 5.6 suggèrent un modèle ARMA(p,q) car il y a une décroissance rapide vers 0 pour chacun des deux critères . La présence de saisonnalité et de tendance justifient le choix d'un modèle SARIMA. Différentes combinaisons de paramètres sont testées et celle qui minimise les critères d'ajustement (BIC, AIC) est conservée. La fonction *auto.arima* du logiciel **R** permet d'obtenir la meilleure combinaison. Dans notre cas, le modèle retenu est un modèle $SARIMA[(2, 1, 0)(0, 1, 0)]_{12}$.

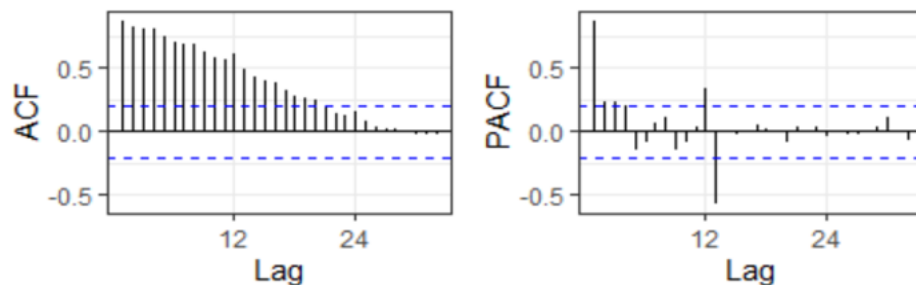


Figure 5.6 ACF et PACF du processus résiduel

Estimation et vérification du modèle Les paramètres des modèles sont estimés à l'aide de la méthode du maximum de vraisemblance. Pour vérifier la validité des modèles estimés, il faut s'assurer de la significativité des paramètres et de la blancheur des résidus.

PARAMÈTRE	ESTIMATION	P-VALEUR
AR(1)	-0,642	0.000
AR(2)	-0,353	0.000

Table 5.3 Significativité des paramètres du processus ARMA

Les p-valeurs du test étant toutes inférieures à 5%, aucun des paramètres ne peut alors être rejeté. Les résidus sont un bruit blanc gaussien au vu du QQ-plot ci-dessous. Le test de normalité de Shapiro WILK permet de confirmer ce constat graphique. L'ACF des résidus prouve que les résidus ne sont pas corrélés.

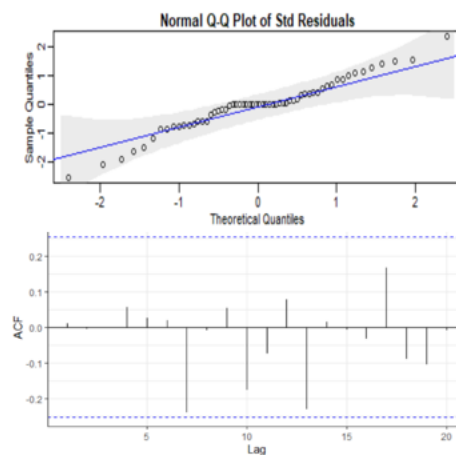


Figure 5.7 QQ-plot et ACF des résidus du processus ARMA

Ajustement du modèle et prévisions

La finalité de notre modélisation est de pouvoir réaliser des prévisions fiables des séries d'incidence. Suite à vérification de la validité des paramètres du modèle et de la blancheur des résidus, il faut analyser la qualité d'ajustement de notre modèle. Les prévisions du modèle parfait se superposent aux données observées. La proximité entre les données estimées et les données observées permet de contrôler la qualité du modèle. D'après la figure 5.8, le modèle SARIMA s'ajuste bien aux données sur la période de janvier 2014 à décembre 2015. On constate un détachement assez

important à partir de 2016.

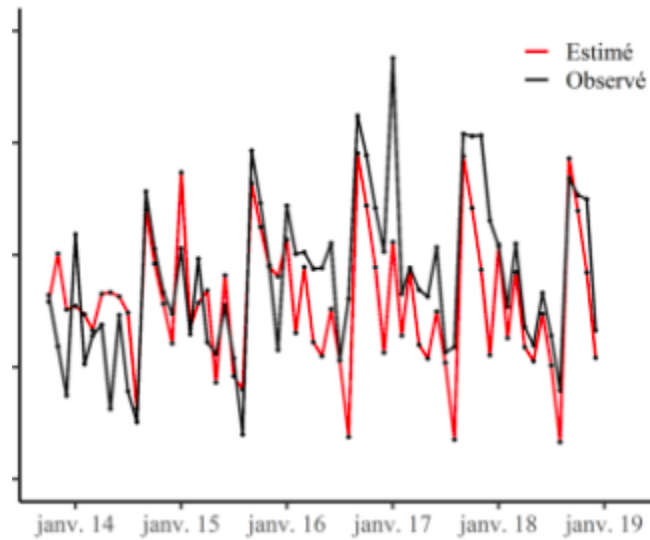


Figure 5.8 Ajustement du modèle SARIMA

5.4.2 Résultats de la modélisation ARMA avec intervention

Pour la modélisation ARMA avec intervention, il est important de bien choisir la période à laquelle intervient la perturbation. Ce choix ne se fait pas de manière arbitraire. En effet, il existe des méthodes statistiques qui permettent de détecter les moments auxquels la série subit une ou plusieurs perturbations. Ces méthodes sont l'objet de plusieurs recherches notamment dans la littérature anglo-saxonne où elles sont connues sous le terme "Change point detection". Parmi les publications à ce sujet on peut citer AMINIKHANGHAHI et COOK [1] qui énumèrent, catégorisent et comparent plusieurs des méthodes proposées pour détecter les points de changement dans les séries chronologiques. La connaissance métier permet aussi de pouvoir déterminer le ou les moment(s) de l'intervention. Par exemple, lorsqu'une grève des transports en commun a lieu, le nombre de véhicules qui passent par une autoroute est plus élevé que ce qui est observé habituellement. La perturbation pour cette série d'observations intervient donc à la date de la grève. Dans notre cas, les méthodes de détection de changement de niveau ne sont pas utilisées. Les statistiques descriptives montrent qu'en 2016 il y a eu un pic de sinistralité qui est expliqué par l'entrée en vigueur du

décalage de 2 ans introduit par la réforme des retraites de 2010. Le modèle ARMA ajusté (cf. 5.8) présente un détachement à partir de janvier 2016. Les analyses graphiques permettent donc valablement de retenir janvier 2016 comme étant la période d'intervention de la perturbation.

Trois types de fonctions d'intervention ont été présentés plus haut. Dans le cas de la série d'observations du nombre de véhicules sur l'autoroute, on imagine bien que selon l'évolution de la grève on peut être face à une intervention ponctuelle ou un effet rémanent avec changement de niveau. En effet, si la grève se poursuit, le nombre de véhicules sera tout aussi important. Par contre si le trafic du réseau de transports en commun reprend de manière progressive, le flux de véhicules diminuera. Dans le cas de notre étude, l'évolution du volume de sinistre présenté sur le graphique 3.1 suggère une intervention à effet rémanent (cf. 5.4).

Pour mieux analyser les effets de la réforme, il est nécessaire d'isoler tout autre effet externe susceptible d'influer sur le pic de sinistralité observé à partir de 2016. Cela permet de tester la cohérence de l'hypothèse faite mais aussi d'assurer la fiabilité des résultats du modèle. Pour cela des variables exogènes sont intégrés à notre processus ARMA. Il sera question de neutraliser l'effet de la composition du portefeuille. cet aspect de la modélisation ne sera pas développé ici. Nous nous contenterons de présenter les résultats de cette modélisation sans entrer dans les détails. Le lecteur intéressé par le sujet pourra se référer à la thèse de BERGEL-HAYAT publiée en 2008 [3] qui développe le thème de la prise en compte de variables explicatives dans les modèles de séries temporelles à travers des applications. Les variables exogènes utilisées sont la proportion de sinistres par sexe et par CSP. Le modèle s'écrit :

$$Y_t = C + \sum_{j=1}^k b_j X_{kj} + \frac{\Theta(B)}{\theta(B)} \epsilon_t \quad (5.11)$$

où :

Y_t est la série d'incidences,

$X_{1t}, X_{2t}, \dots, X_{5t}, \dots$ sont les variables aléatoires explicatives à chaque mois,

$b_1, b_2, \dots, b_5, \dots$ les coefficients associés à ces variables

Le ratio de polynômes $\frac{\Theta(B)}{\theta(B)}$ présente la partie ARMA du modèle

Les coefficients estimés des variables exogènes ne sont pas significatifs au seuil de 5% (cf. 5.4). L'écart type résiduel augmente à chaque rajout de variable exogène, ce qui n'améliore pas la qualité du modèle. Il est donc raisonnable d'affirmer que la composition du portefeuille n'affecte pas l'évolution des taux d'incidence des seniors, dans notre cas. Dans le chapitre 2, nous avons montré qu'aucune autre réforme des retraites n'interagit avec celle de 2010. L'analyse d'intervention peut alors être mise en place.

RÉGRESSEUR	ESTIMATEUR	P-VALEUR	PERTE D'ÉCART- TYPE RÉSIDUEL
Proportion homme	-0,0024	0,857	2,22%
Proportion femme	-0,0027	0,987	2,15%
Proportion cadre	-0,0016	0,333	0,02%
Proportion non cadre	-0,0093	0,527	1,48%
Proportion EDP ¹	-0,0023	0,862	2,21%

Table 5.4 Significativité des paramètres du processus ARMA avec variables explicatives

Avant d'essayer de développer un modèle univarié capturant les effets de l'intervention, il est nécessaire d'énumérer les hypothèses sous lesquelles une telle tentative est faite. Ces hypothèses sont les suivantes :

- Les taux d'incidence des seniors est restée inchangée pendant la période de l'étude.
- Le vieillissement de la population active entraîne une augmentation de la sinistralité
- Aucune autre intervention majeure n'a eu lieu entre 2016 et 2018.
- La structure du bruit du modèle univarié est restée inchangée avant et après l'intervention.
- Les taux d'incidences augmentent de façon exponentielle après l'intervention.

Estimation et vérification du modèle ARMA avec intervention Les paramètres des modèles sont estimés à l'aide de la méthode du maximum de vraisemblance. Pour vérifier la validité des modèles estimés, il faut s'assurer de la significativité des paramètres et de la blancheur des résidus.

PARAMÈTRE	ESTIMATION	P-VALEUR
AR(1)	-0,642	0,000
AR(2)	-0,353	0,000
ω	0,0083	0,000
δ	0,6203	0,000

Table 5.5 Significativité des paramètres du modèle ARMA avec intervention

Les p-valeurs du test étant toutes inférieures à 5%, aucun des paramètres ne peut alors être rejeté. Les résidus sont un bruit blanc gaussien au vu du QQ-plot ci-dessous. Le test de normalité de Shapiro WILK permet de confirmer ce constat graphique. L'ACF des résidus prouve que les résidus ne sont pas corrélés.

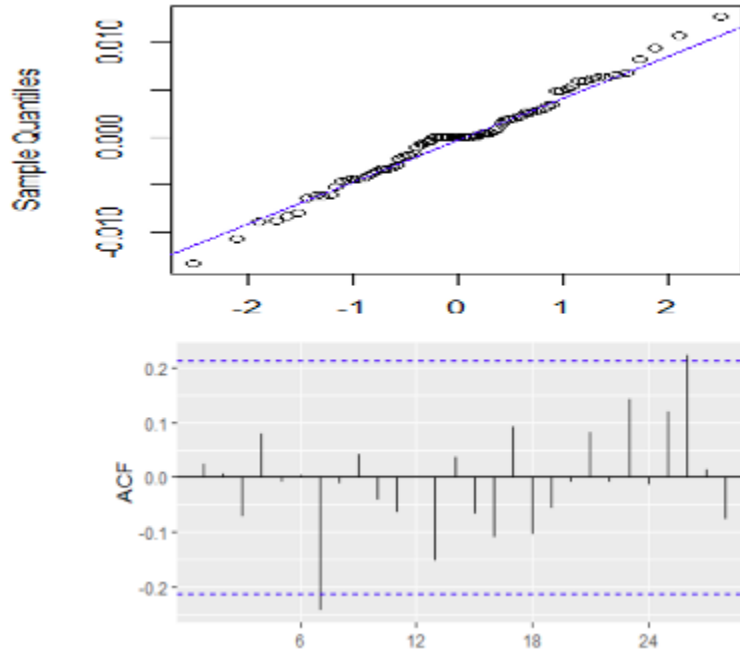


Figure 5.9 QQ-plot et ACF des résidus du modèle ARMA avec intervention

Les p-valeurs du test étant toutes inférieures à 5%, aucun des paramètres ne peut alors être rejeté. Les résidus sont un bruit blanc gaussien au vu du QQ-plot ci-dessous. Le test de normalité de Shapiro WILK permet de confirmer ce constat graphique. L'ACF des résidus prouve que les résidus ne sont pas corrélés.

Ajustement du modèle et prévisions Suite à vérification de la validité des paramètres du modèle et de la blancheur des résidus, il faut analyser la qualité d'ajustement de notre modèle. Le modèle parfait se superpose aux données observées. La proximité entre les données estimées et les données observées permet de contrôler la qualité du modèle. D'après la figure 5.10, sur l'étendue de la période de l'étude, le modèle avec intervention s'ajuste mieux aux données que le modèle SARIMA classique.

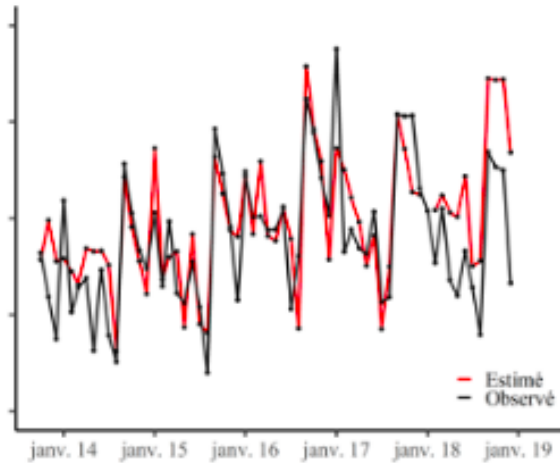


Figure 5.10 Ajustement du modèle ARMA avec intervention

5.4.3 Conclusions et limites de la modélisation

Avant de faire des prévisions sur le long terme, il est important de s'assurer, au delà de l'analyse graphique, que la prise en compte de l'intervention améliore notre modèle. Pour ce faire, nous allons comparer, pour les deux modèles, les critères suivants :

- L'erreur moyenne absolue (MAE²) est un très bon indicateur de performance pour mesurer la précision des prévisions. C'est la moyenne, sur l'échantillon test, des différences absolues entre la prédiction et l'observation réelle, où toutes les différences individuelles ont le même poids.

$$\text{MAE} = \frac{1}{n} \sum_{j=1}^n |y_j - \hat{y}_j|$$

- L'erreur quadratique moyenne (RMSE³) est la racine carrée de la moyenne des différences quadratiques entre la prédiction et l'observation réelle.

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{j=1}^n (y_j - \hat{y}_j)^2}$$

Les critères MAE et RMSE expriment tous deux l'erreur moyenne de prédiction du modèle en unités de la variable d'intérêt. Les deux mesures peuvent aller de 0 à ∞ et sont indifférentes aux

2. Mean Absolute Error

3. Root Mean Squared Error

signes des erreurs. Les valeurs inférieures sont meilleures. Le tableau 5.7 montre que le modèle avec intervention fournit des prédictions beaucoup plus proches de la réalité. C'est donc celui qui sera utilisé pour l'estimation de l'impact de la réforme sur les taux d'incidences et leurs projections sur les 10 prochaines années.

	Ajustement		Prédiction	
Modèle ARMA sans intervention	MAE	0,9	AIC	-292,33
Modèle ARMA avec intervention	MAE	0,6	AIC	-437,03
Modèle ARMA sans intervention	RMSE	0,012	BIC	-286,78
Modèle ARMA avec intervention	RMSE	0,009	BIC	-428,58

Table 5.7 Comparaison de modèles

L'hypothèse fondamentale de la modélisation de l'intervention est que la structure du modèle est la même avant et après intervention. Ainsi, pour mesurer l'amplitude de la réforme des retraites sur les taux d'incidence, la modélisation s'est déroulée en quatre étapes :

Étape 1 : Utiliser les données avant l'intervention (janvier 2014 à décembre 2015) pour déterminer un modèle $SARIMA_1$ pour la série

Étape 2 : Utiliser le modèle $SARIMA_1$ pour prédire les valeurs de la période post-intervention (janvier 2016 à décembre 2018).

Étape 3 : Intégrer au modèle $SARIMA_1$ l'intervention en utilisant l'ensemble des données (janvier 2014 à décembre 2018)

Étape 4 : La différence entre les valeurs prédites suite à l'étape 3 et les valeurs prédites à l'étape 2 représente l'effet de l'intervention (ici, la réforme des retraites de 2010)

Sur l'ensemble de la période post-intervention, la réforme aura contribué à la hausse des taux d'incidence à hauteur de 21,62% en moyenne. Le détail par année de survenance est donné dans le tableau 5.8 ci-dessous.

ANNÉE	EFFET RÉFORME
2016	+16,68%
2017	+17,38%
2018	+30,79%

Table 5.8 Effets de la réforme des retraites sur les taux d'incidence des seniors entre 2016 et 2018

À présent la problématique initiale a été résolue. Le modèle est utilisé pour faire des prévisions sur le long terme. Les prévisions indiquent que, d'ici 2030, les taux augmenteront jusqu'à +10 points par rapport à 2018 (cf. figure 5.11). Cette évolution tient compte de la structure et de l'évolution du portefeuille AXA mais aussi du recul de l'âge de départ à la retraite de 2 ans qui concerne aujourd'hui toute la population active.

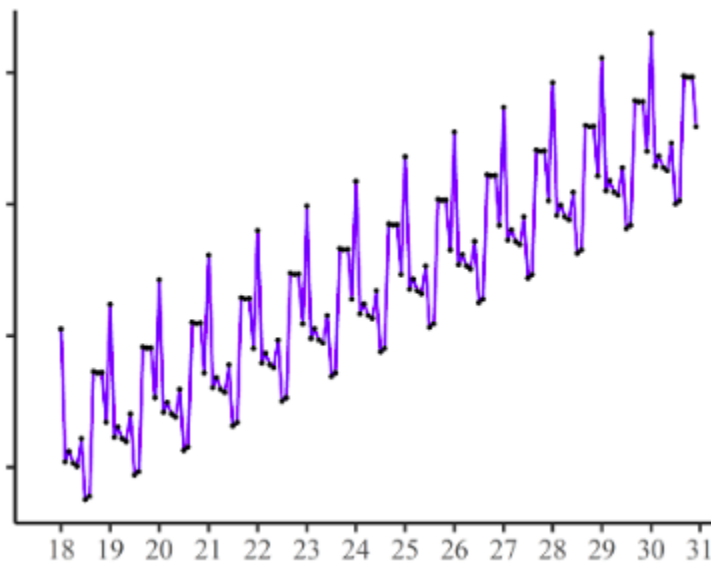


Figure 5.11 Projection des taux d'incidence des seniors de 2019 à 2030

Limites de la modélisation

Les modèles de séries temporelles sont d'autant meilleurs que l'historique de données est important. En particulier, l'analyse de l'intervention n'est possible que lorsqu'on dispose d'au moins deux années d'historique avant et après la perturbation. Dans le cadre de cette étude, la base de données répond au critère minimal requis. En effet, nous avons utilisé un historique de 4 années, ce qui est relativement faible et constitue une limite à notre modélisation.

Lors de l'analyse d'intervention, les effets externes susceptibles d'influer sur la sinistralité (autre que la réforme) doivent être neutralisés afin de renforcer la qualité du modèle d'intervention. Pour cela, la composition du portefeuille a été rajoutée au modèle. L'ajout de ces variables n'a révélé aucune amélioration sur le modèle. Il a donc été conclu qu'une analyse d'intervention, lorsque les paramètres estimés sont significatifs, est pertinente et juste. Cependant, en plus de l'effet de composition du portefeuille, il serait intéressant d'analyser l'évolution de l'état de santé de la population. En effet, la croissance de la sinistralité pourrait être due à une dégradation progressive de l'état de santé des salariés. Malheureusement, les ressources nécessaires à une telle analyse ne sont pas à portée. Leur mise en place requiert plus de temps et s'éloigne du contexte des travaux actuels.

Une précaution particulière doit être portée sur les survenances 2020 et 2021 pour lesquelles les prévisions sont susceptibles d'être drastiquement différentes de ce qui sera observé à cause de la pandémie COVID 19. L'apparition sur le territoire français du coronavirus a conduit le gouvernement à déclencher l'état d'urgence sanitaire et à instaurer, pour les salariés, des arrêts de travail dérogatoires plus favorables que les arrêts de travail classiques. Ainsi, les conditions d'octroi de ces arrêts ont été assouplies et les délais de carence supprimés. La durée d'application de ces assouplissements et l'évolution de la pandémie, auront des répercussions sur la sinistralité du portefeuille.

Conclusion et perspectives

Ce mémoire a été l'occasion pour l'équipe Business Support France d'évaluer l'impact de la réforme des retraites de 2010 sur la sinistralité des garanties incapacité et invalidité. Au terme des travaux réalisés, les points clefs à retenir sont les suivants :

Arrêts impactés par la réforme : En invalidité, 20,7% de sinistres sont dus à la réforme dont 91% concernent le décalage de 2 ans. En incapacité, 1,7% des arrêts maladie sont dus à la réforme dont 90% sont portés par le décalage de 2 ans

Impact PM IP : La provision mathématique en invalidité subit une hausse de 25% suite à l'allongement de la durée d'indemnisation, dont +319,6% en plus pour les seniors. Pour le risque incapacité, la PM pour invalidité en attente est majorée de 1,3%

Réforme 2020 : La projection des impacts de la réforme de 2020 indique un effet composé de la réforme de 2010 et de la nouvelle réforme sur le stock de sinistres encore ouverts. En incapacité, 1,84% de sinistres en plus sont attendus, contre 20,84% en invalidité

Impact sur les taux d'incidence des seniors : Entre 2016 et 2018 la réforme contribue à la hausse des taux d'incidence des seniors à hauteur de 21,6% . Sur le long terme, une évolution de 10 points supplémentaires sont attendus entre 2019 - 2030.

Les impacts du report de l'âge de départ à la retraite sur les garanties incapacité et invalidité sont nombreux. En effet, l'étude des aspects suivants aurait été tout aussi intéressante :

- Impacts de l'augmentation de l'âge sur le bilan prudentiel : l'effet sur les provisions techniques

Solvabilité 2, effet sur le SCR⁴ et les effets sur le ratio de couverture du SCR

- Pour faire face à la hausse des provisions, une nouvelle tarification peut être mise en place. L'impact sur la tarification pourrait révéler des informations importantes bien qu'il soit en partie pris en compte puisque les calculs d'impacts réalisés pourraient être des appuis pour permettre des majorations des cotisations à venir.

4. Solvency Capital Required

Table des figures

2.1	Organisation du système de retraite en France	34
2.2	Évolution de la population de la France de 1990 à 2070 (scénario central de projection). <i>Source : Projections de population à l’horizon 2070, Insee</i>	42
2.3	Population active observée et stimulée par âge en France métropolitaine. <i>Source : Projections de population active 2006-2050, Enquêtes Emploi, Insee.</i>	43
2.4	Taux d’activité des 55-59 ans et des 60-64 ans en France et dans l’Union Européenne à 27 pays depuis 2000. <i>Source : Emploi et chômage des seniors en 2019, DARES</i> .	43
3.1	Evolutions du nombre de sinistres observé et attendu, du nombre de contrats (RPP) et du montant des cotisations par année de survenance	53
3.2	Évolutions de la part des assurés âgés de plus de 60 ans à la survenance du sinistre .	54
3.3	Composition du portefeuille par sexe	55
3.4	Pyramide des âges	56
3.5	Distribution des arrêts selon l’ancienneté en incapacité	59
3.6	Comparaison de la durée moyenne des seniors et de l’ensemble de la population en incapacité	60
3.7	Distribution des arrêts selon l’ancienneté en invalidité	61

3.8	Comparaison de la durée moyenne des seniors et de l'ensemble de la population en invalidité	62
3.9	Indemnité journalière moyenne par classe d'âge et par survenance	63
4.1	Illustration des arrêts dus à la réforme	68
4.2	Illustration des arrêts prolongés par la réforme	69
4.3	PM avant/après réforme par survenance	81
4.4	PM avant/après réforme par tranches d'âge à la survenance	83
4.5	Évolution de l'impact de chaque réforme sur la PM du stock de sinistres en invalidité	86
5.1	Présentation de la série d'incidences	90
5.2	Boxplot des incidences mensuelles	91
5.3	Illustration de l'intervention ponctuelle	108
5.4	Illustration de l'Intervention à effet rémanent	108
5.5	Illustration de l'intervention à effet rémanent et changement de niveau	109
5.6	ACF et PACF du processus résiduel	110
5.7	QQ-plot et ACF des résidus du processus ARMA	111
5.8	Ajustement du modèle SARIMA	112
5.9	QQ-plot et ACF des résidus du modèle ARMA avec intervention	116
5.10	Ajustement du modèle ARMA avec intervention	117
5.11	Projection des taux d'incidence des seniors de 2019 à 2030	119
B.1	Répartition du volume de sinistres (incapacité + invalidité) par tranche d'âge et par survenance	131
B.2	Durée moyenne en invalidité par survenance et par tranche d'âge	131

B.3	Durée moyenne en incapacité par survenance et par tranche d'âge	132
B.4	Durée moyenne en incapacité avant le passage en invalidité par survenance et par tranche d'âge	132
B.5	Durée moyenne en invalidité suite à une incapacité par survenance et par tranche d'âge	133

Liste des tableaux

2.1	Âges de départ à la retraite par génération après la réforme de 2010	37
2.2	Evolution de l'âge légal de départ et de la date de départ avant et après la réforme par génération	38
2.3	Nouvelles mesures : Âges de départ à la retraite par génération suite à la réforme de 2010	39
3.1	Variables de la base de données MRM	48
3.2	Proportion de sinistres encore ouverts à la date du 30 mars 2020 par année de survenance	58
3.3	Évolution des durées médianes et moyennes en fonction des survenances	60
3.4	Durée moyenne par tranche d'âge	60
3.5	Évolution des durées médianes et moyennes en fonction des survenances	61
3.6	Durée moyenne par tranche d'âge	62
3.7	Passage IT - IP par tranche d'âge	63
4.1	Âges de départ à la retraite par génération après la réforme de 2010	66
4.2	Invalidité - Proportions des sinistres additionnels par génération	70
4.3	Incapacité - Proportion de sinistres additionnels par génération	71
4.4	Invalidité - Résultats sur les durées moyennes par génération	72

4.5	Incapacité - Résultats sur les durées moyennes par génération	73
4.6	Invalidité - Résultats sur les prestations moyennes versées par génération	74
4.7	Incapacité - Résultats sur les prestations moyennes versées par génération	74
4.8	Evolution de la PM invalidité par année de survenance	82
4.9	Impact de la PM des 60 ans et plus dans la PM globale par survenance	84
4.10	Proposition de calendrier de mise en place de l'âge pivot	85
4.11	Proportion de sinistres additionnels liés à chaque réforme	87
4.12	Décomposition de la PM des arrêts en cours	87
5.1	Choix de la quantité à modéliser	89
5.2	Critères de détermination des modèles en fonction de l'ACF et du PACF	97
5.3	Significativité des paramètres du processus ARMA	111
5.4	Significativité des paramètres du processus ARMA avec variables explicatives	114
5.5	Significativité des paramètres du modèle ARMA avec intervention	115
5.7	Comparaison de modèles	118
5.8	Effets de la réforme des retraites sur les taux d'incidence des seniors entre 2016 et 2018	119

APPENDICES

Annexe A

Algorithme de Jenks

Principe de la méthode : Proposer un découpage où les individus d'une même classe sont les plus proches possibles (faible variance intra-classes) et bien isoler les différents groupes (forte variance inter-classes). C'est une méthode qui respecte la structure des données et s'adapte à tous types de distribution. Les résultats sont robustes car elle s'appuie sur l'analyse de la variance. Largement utilisées dans les progiciels GIS, ce sont des formes de classification à variance minimale. Les ruptures sont généralement inégales et sont sélectionnées pour séparer les valeurs lorsque de grands changements de valeur se produisent. Elles peuvent être fortement affectées par le nombre de classes sélectionnées et ont tendance à avoir des limites de classe inhabituelles.

Étape 1 : L'utilisateur sélectionne l'attribut, x , à classer et précise le nombre de classes requises, k

Étape 2 : Un ensemble de $k-1$ valeurs aléatoires ou uniformes est généré dans la plage $[\min x, \max x]$. Ces valeurs sont utilisées comme limites de classe initiales

Étape 3 : Les valeurs moyennes pour chaque classe initiale sont calculées et la somme des écarts au carré des membres de la classe par rapport aux valeurs moyennes est calculée. La somme totale des carrés des écarts (STSE) est enregistrée

Étape 4 : Les valeurs individuelles de chaque classe sont ensuite systématiquement attribuées aux classes adjacentes en ajustant les limites des classes pour voir si la STSE peut être réduite. Il s'agit d'un processus itératif qui se termine lorsque l'amélioration de la STSE passe en dessous d'un niveau seuil, c'est-à-dire lorsque la variance à l'intérieur de la classe est aussi

faible que possible et que la variance entre les classes est aussi importante que possible. La véritable optimisation n'est pas assurée. L'ensemble du processus peut être éventuellement répété à partir de l'étape 1 ou 2 et les valeurs de la STSE sont comparées

Annexe B

Statistiques descriptives

B.0.1 Volume

Proportions du nombre d'arrêts par tranche et par survenance					
	14 - 32 ans	33 - 41 ans	42 - 49 ans	50 - 57 ans	+58 ans
2011	5%	16%	26%	43%	11%
2012	10%	19%	25%	36%	10%
2013	18%	22%	24%	28%	9%
2014	17%	22%	24%	27%	10%
2015	17%	22%	24%	28%	10%
2016	18%	22%	23%	27%	10%
2017	19%	22%	23%	26%	10%
2018	20%	22%	23%	25%	10%
2019	22%	23%	21%	24%	10%

Figure B.1 Répartition du volume de sinistres (incapacité + invalidité) par tranche d'âge et par survenance

B.0.2 Durées moyennes

Durée moyenne IP						
	14 - 32 ans	33 - 41 ans	42 - 49 ans	50 - 57 ans	+58 ans	Global
2011	13,8	13,6	9,1	5,4	1,8	7,1
2012	16,3	12,1	9,4	5,4	2,0	7,1
2013	15,4	12,7	9,2	5,5	1,7	6,8
2014	18,5	13,9	9,6	5,6	1,9	7,1
2015	15,2	14,4	10,0	5,7	1,9	7,0
2016	18,7	14,8	9,9	5,6	2,1	7,0
2017	16,9	14,1	10,4	5,6	2,2	6,7
2018	19,2	14,9	10,0	5,5	2,5	6,6
2019	12,1	13,2	10,1	5,4	2,7	6,5

Figure B.2 Durée moyenne en invalidité par survenance et par tranche d'âge

Durée moyenne IT						
	14 - 32 ans	33 - 41 ans	42 - 49 ans	50 - 57 ans	+58 ans	Global
2014	4,7	6,3	7,6	8,5	7,6	7,0
2015	4,6	6,5	7,5	8,5	7,8	7,0
2016	3,7	5,4	6,5	7,5	6,7	6,0
2017	3,6	5,5	6,3	7,2	6,8	5,9
2018	3,7	5,4	6,7	7,4	6,8	5,9
2019	3,8	5,0	6,1	6,7	6,3	5,5

Figure B.3 Durée moyenne en incapacité par survenance et par tranche d'âge

B.0.3 Passage de l'incapacité vers l'invalidité

IT	14 - 32 ans	33 - 41 ans	42 - 49 ans	50 - 57 ans	+58 ans
2011	33,65	33,08	31,87	30,79	30,65
2012	29,90	27,96	28,22	25,87	21,47
2013	26,42	26,61	26,27	22,94	18,99
2014	28,18	26,75	25,95	23,22	19,02
2015	26,68	26,16	25,38	22,66	17,88
2016	26,06	25,13	23,93	22,24	17,70
2017	17,84	18,00	18,95	18,46	15,57
2018	11,74	12,08	12,37	12,68	10,66
2019	10,32	7,75	7,01	6,72	7,03

Figure B.4 Durée moyenne en incapacité avant le passage en invalidité par survenance et par tranche d'âge

IT	14 - 32 ans	33 - 41 ans	42 - 49 ans	50 - 57 ans	+58 ans
2011	33,65	33,08	31,87	30,79	30,65
2012	29,90	27,96	28,22	25,87	21,47
2013	26,42	26,61	26,27	22,94	18,99
2014	28,18	26,75	25,95	23,22	19,02
2015	26,68	26,16	25,38	22,66	17,88
2016	26,06	25,13	23,93	22,24	17,70
2017	17,84	18,00	18,95	18,46	15,57
2018	11,74	12,08	12,37	12,68	10,66
2019	10,32	7,75	7,01	6,72	7,03

Figure B.5 Durée moyenne en invalidité suite à une incapacité par survenance et par tranche d'âge

Annexe C

Analyse des résidus

Dans tous les modèles cités précédemment, une attention particulière a été portée aux résidus. En un point d'observation k , l'écart entre Y observé et Y estimé par le modèle est le résidu au point k : $e_k = Y_k - \hat{Y}_k$. Ces résidus e_k sont considérés comme les erreurs observées des vraies erreurs inconnues ε_k telles que $\varepsilon_k = Y_k - E(Y)$. Les hypothèses faites sur les ε_k pour élaborer les tests statistiques se résument ainsi « les erreurs doivent être indépendantes et identiquement distribuées. Dans les méthodes d'estimation utilisant la méthode des moindres carrés, il faut en plus que les résidus suivent une loi normale. En effet, les résidus contiennent d'une part un aléa d'espérance nulle et de variance σ^2 , et d'autre part une information concernant l'inadéquation du modèle aux données (c'est-à-dire l'écart entre le modèle postulé et le modèle correct inconnu). Ce que l'on veut c'est que l'importance de cette deuxième partie soit moindre que celle due à l'aléa. Si le modèle est approprié aux données, les résidus observés e_k doivent refléter les propriétés des vraies erreurs inconnues ε_k . On doit donc s'assurer de la conformité de ces hypothèses.

C.1 Tests de bruit blanc

C.1.1 Fonction d'autocorrélation empirique (ACF)

La fonction d'autocorrélation d'une suite de variables aléatoires indépendantes et identiquement distribuées de variance finie, vérifie $\rho(h) = 0, \forall h > 0$. L'estimation de l'ACF, $\hat{\rho}$, doit par conséquent être proche de 0.

Soit Y_1, \dots, Y_n v.a. i.i.d. de variance finie. Alors $(\hat{\rho}_Y^{(n)}(h), h > 0)$ est approximativement une

suite de v.a. i.i.d. de loi $\mathcal{N}(0, \frac{1}{n})$, pour n grand.

Par conséquent approximativement 95% des $\hat{\rho}^{(n)}$ devraient se situer entre les bornes $\pm \frac{q(0,975)}{\sqrt{n}} = \pm \frac{1,96}{\sqrt{n}}$ pour satisfaire l'hypothèse d'indépendance.

En pratique, si l'on calcule $\hat{\rho}^{(n)}(h)$, avec $h = 1, \dots, 40$, et que l'on trouve plus de 2 ($= 40 * 0,05$) ou 3 valeurs en dehors de ces bornes (ou qu'une valeur soit vraiment très loin des bornes), alors on rejettera l'hypothèse d'avoir une suite i.i.d.

C.1.2 Test du Porte-manteau

On considère la statistique associée à l'ACF $\hat{\rho}^{(n)}(\cdot)$ définie par :

$$Q_n = n \sum_{h=1}^k \left(\hat{\rho}^{(n)}(h) \right)^2$$

On remarque, en utilisant le modèle (*), que si Y_1, \dots, Y_n est une suite de v.a. i.i.d. de variance finie, alors la loi de Q_n est, pour n grand, approximativement une loi du khi-deux à k degrés de liberté. Une grande valeur de Q_n nous indique donc que les autocorrélations des données sont trop grandes pour que les données soient celles d'une suite i.i.d. d'où le test au niveau de région critique (i.e. rejet de l'hypothèse i.i.d.) $Q_n > \chi_k^2(1 - \alpha)$; avec $\chi_k^2(1 - \alpha) =$ quantile d'ordre $(1 - \alpha)$ du Khi-deux d'ordre k .

C.2 Tests de normalité des résidus

Nous utilisons la technique graphique du Q-Q plot et le test de Shapiro-Wilk. Le seuil limite retenu est 5%

Q-Q plot

Le Q-Q plot, quantile-quantile plot est une technique graphique qui permet de comparer les distributions de deux ensembles de données. Dans notre cas, nous voulons tester l'ajustement de nos résidus à une loi normale $N(0, 1)$. Soit $x_{(i)}$ tel que $x_{(1)} < x_{(2)} < \dots < x_{(n)}$. On associe à chaque

valeur $x_{(i)}$, le $i/(n + 1)$ -quantile d'une loi normale centrée réduite noté $x_{(i)}^*$. Si les données sont compatibles avec une loi normale, les points $(x_{(i)}, x_{(i)}^*)$ forment une droite dite droite de Henry ; ils sont alignés sur la diagonale principale.

Test de Shapiro-Wilk :

Il est basé sur la statistique W . Il teste si la réalisation x_1, x_2, \dots, x_n de la suite de variables aléatoires X_1, \dots, X_n est normalement distribuée. Il est particulièrement puissant pour les petits échantillons de taille ≤ 50 . La statistique W est la suivante :

$$W = \frac{\left[\sum_{i=1}^{\lfloor n/2 \rfloor} a_i (x_{(n-i+1)} - x_{(i)}) \right]^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

où $x_{(i)}$ est la $i^{\text{ème}}$ observation de la série des observations ordonnées ; $\lfloor n/2 \rfloor$ est la partie entière du rapport $n/2$ les termes a sont des constantes générées à partir de la moyenne et de la matrice de covariance des quantiles d'un échantillon de taille n suivant la loi normale. Ces constantes sont disponibles dans des tables spécifiques.

On rejette l'hypothèse de normalité si $W < W_{\text{critique}}$. Les seuils W_{critique} pour différents risques α et effectifs n sont disponibles dans la table de Shapiro-Wilk.

Bibliographie

- [1] Samaneh AMINIKHANGHAHI et Diane J COOK. “A survey of methods for time series change point detection”. In : *Knowledge and information systems* 51.2 (2017), p. 339–367.
- [2] Franck ARNAUD et al. “Les retraités et les retraites : édition 2019”. In : (2019).
- [3] Ruth BERGEL-HAYAT. “La prise en compte de variables explicatives dans les modèles de séries temporelles : application à la demande de transport et au risque routier”. Thèse de doct. 2008.
- [4] Nathalie BLANPAIN et Guillemette BUISSON. “Projections de population à l’horizon 2070. Deux fois plus de personnes de 75 ans ou plus qu’en 2013”. In : (2016).
- [5] George EP BOX et George C TIAO. “Intervention analysis with applications to economic and environmental problems”. In : *Journal of the American Statistical association* 70.349 (1975), p. 70–79.
- [6] George EP BOX et al. *Time series analysis : forecasting and control*. John Wiley & Sons, 2015.
- [7] Elise COUDIN. “Projections de population active 2006-2050 : des actifs en nombre stable, une population âgée toujours plus nombreuse”. In : (2007).
- [8] Préconisations de Jean-Paul DELEVOYE. “Pour un système universel de retraite”. In : (2019).
- [9] Thomas DERUYON. “En 2018, l’espérance de vie sans incapacité est de 64, 5 ans pour les femmes et de 63, 4 ans pour les hommes”. In : (2019).
- [10] Yves DUBOIS et Malik KOUBI. “La réforme des retraites de 2010 : quel impact sur l’activité des séniors ?” In : *Economie prevision* 2 (2017), p. 61–90.
- [11] Yves DUBOIS et Malik KOUBI. “Report de l’âge de la retraite et taux d’emploi des séniors : le cas de la réforme des retraites de 2010”. In : (2017).
- [12] LH KOOPMANS. “The Spectral Analysis of Time Series Academic Press”. In : *New York, San Francisco, London* (1974).
- [13] Claude MINNI. “Emploi et chômage des seniors en 2018. Hausse du taux d’emploi et du taux de chômage”. In : (2019).

- [14] Stefan STAUBLI et Josef ZWEIMÜLLER. “Does raising the early retirement age increase employment of older workers?” In : *Journal of public economics* 108 (2013), p. 17–32.