



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

Mars 2024

Par : Valentin LECOINTRE

Titre : Les traitements envisageables sur les scénarios bicentennaires de longévité permettant
l'intégration des données COVID-19

Confidentialité : Non

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

**Membre présent du jury de l'Institut
des Actuares :**

.....

Signature :

Entreprise :

AXA France

Signature :

Membres présents du jury de l'EURIA :

.....

Directeur de mémoire en entreprise :

Aurélie HAYNES

Signature :

Invité :

.....

Signature :

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

(après expiration de l'éventuel délai de confidentialité)

Signature du responsable entreprise :

Signature du candidat :

La norme Solvabilité 2 impose aux assureurs de l'Union Européenne de détenir un niveau de fonds propres suffisant afin de couvrir un événement défavorable ayant une probabilité d'occurrence bicentenaire. Sous ce régime, le régulateur définit différents modules de risques, dont le risque de longévité, lié à la diminution anormale de la mortalité des assurés. Concernant ce risque en France, on observe depuis de nombreuses années une constante amélioration de l'espérance de vie. Cependant, 2019 fut l'apparition de la pandémie de COVID-19 entraînant la plus importante baisse d'espérance de vie des 50 dernières années.

Or, afin d'évaluer les scénarios bicentenaires de longévité, les assureurs utilisent des modèles de durée ayant pour vocation d'estimer les tendances de mortalité futures en se basant sur les données passées. Ainsi, l'intégration des observations liées à la COVID-19 dans ces modèles a pour effet de déformer les scénarios.

Ce mémoire a pour but d'identifier les impacts de l'intégration de la pandémie dans l'historique de calibrage sur les scénarios bicentenaires de longévité et de proposer des solutions permettant une modélisation adéquate tenant compte de ces observations.

Après avoir présenté notre approche de modélisation, on expose dans un premier temps les impacts de l'intégration de la COVID-19 sur cette dernière. On observe notamment une translation et un rehaussement de la pente des améliorations de mortalité projetées, suggérant une moindre tendance aux améliorations dans le temps. On note également une hausse de la volatilité, indicateur de diffusion des scénarios stochastiques, liée à l'éloignement des nouvelles observations en comparaison aux tendances passées.

Pour faire face à ces constats, on propose divers traitements afin d'absorber ces impacts, en mettant tout d'abord en avant la robustesse du modèle alternatif de Börger-Fleischer-Kuksin proposant un paramètre de tendance spécifique aux âges élevés et permettant une meilleure absorption de la dérive à ces âges. Par la suite, on étudie les opportunités de retraitements des données de calibrage et notamment la suppression des décès liés à la pandémie. Enfin, on expose les ajustements envisageables sur les paramètres de pondération afin d'atténuer l'impact de la COVID-19 sur les scénarios bicentenaires de longévité.

Mots clefs: Longévité, Mortalité, COVID-19, Pandémie, Solvabilité 2, Capital de Solvabilité Requis, Modèle Interne

The Solvency II standard requires European Union insurers to hold a sufficient level of own funds to cover an adverse event with a bicentennial occurrence probability. Under this regime, the regulator defines various risk modules, including longevity risk, associated with the unusual decrease in insured mortality rates. Regarding this risk in France, there has been a consistent improvement in life expectancy for many years. However, 2019 saw the emergence of the COVID-19 pandemic, leading to the most significant drop in life expectancy in the last fifty years.

To assess bicentennial longevity scenarios, insurers use duration models intended to estimate future mortality trends based on past data. Integrating COVID-19 observations into these models has the effect of distorting the scenarios.

This thesis aims to identify the impacts of integrating the pandemic into the calibration history on bicentennial longevity scenarios and to propose solutions for adequate modeling that takes these observations into account.

After presenting our modeling approach, we first discuss the impacts of integrating COVID-19 on it. We observe, in particular, a shift and an increase in the slope of projected mortality improvements, suggesting a lesser tendency for improvements over time. We also note an increase in volatility, an indicator of the spread of stochastic scenarios, linked to the divergence of new observations from past trends.

To address these findings, we propose various treatments to absorb these impacts, first highlighting the robustness of the alternative Börger-Fleischer-Kuksin model, which proposes a specific trend parameter for older ages, allowing better absorption of the drift at these ages. Subsequently, we examine opportunities for recalibrating data, particularly the removal of pandemic-related deaths. Finally, we discuss possible adjustments to weighting parameters to mitigate the impact of COVID-19 on bicentennial longevity scenarios.

Keywords: Longevity, Mortality, COVID-19, Pandemic, Solvency 2, Solvency Capital Requirement, Internal Model

Contexte

La réforme Solvabilité 2 est mise en place pour les assureurs de l'Union Européenne afin d'assurer la stabilité du système financier en établissant un cadre prudentiel strict. Cette réforme introduit notamment la notion de Capital de Solvabilité Requis (SCR), correspondant au montant minimum de fonds propres économiques dont doit disposer la compagnie afin de couvrir un événement défavorable de probabilité d'occurrence bicentenaire. Dans ce cadre, le régulateur définit différents facteurs de risques parmi lesquels on retrouve le risque de longévité, se référant au risque de sous-estimation des probabilités de survie des assurés.

Depuis plusieurs décennies, la France bénéficie d'une augmentation régulière de l'espérance de vie attribuable aux avancées médicales ainsi qu'aux évolutions sociales et comportementales. Toutefois, en 2020, la France a connu la baisse d'espérance de vie à la naissance la plus conséquente des soixante-dix dernières années, un phénomène principalement expliqué par la pandémie mondiale de COVID-19.

Afin de définir localement les scénarios bicentennaires de longévité sous Solvabilité 2, l'usage est d'utiliser des modèles de durée, ces modèles permettent d'estimer les tendances de mortalité futures à l'aide des observations passées. Par conséquent, l'inclusion des données relatives à la pandémie dans l'historique de calibrage a pour effet de modifier sensiblement les trajectoires de ces scénarios.

L'objectif de ce mémoire est d'évaluer les conséquences de l'intégration des données relatives à la COVID-19 sur les scénarios bicentennaires de longévité et de suggérer des approches permettant une modélisation adéquate en tenant compte de ces informations.

Méthodologie

Afin d'évaluer la sensibilité de plusieurs modèles à l'intégration des données de COVID-19 et de mesurer différemment les effets des approches proposées, trois modèles de durée sont étudiés dans ce mémoire.

Le modèle de Lee-Carter (1992) est retenu en tant que référence de la littérature sur les modèles de durée, donnant une expression du taux instantané de mortalité de la forme suivante :

$$\ln \mu_{x,t} = \alpha_x + \beta_x \times \kappa_t + \epsilon_{x,t}$$

Ensuite, le modèle de Cairns-Blake-Dowd (2006), également bien établi dans la littérature, se distinguant notamment par l'inclusion d'un second paramètre temporel, offre une expression du quotient de mortalité de la forme suivante :

$$\text{logit}q_{x,t} = \kappa_t^{(1)} + \kappa_t^{(2)} \times (x - \bar{x}) + \epsilon_{x,t}$$

Enfin, le modèle de Börger-Fleischer-Kuksin alternatif, nettement moins répandu, a l'avantage d'intégrer un troisième paramètre temporel spécifique aux âges élevés, le rendant particulièrement adapté dans le cadre de cette étude. Le modèle donne une expression du quotient de mortalité de la forme suivante :

$$\text{logit}q_{x,t} = \alpha_x + \kappa_t^{(1)} + \kappa_t^{(2)} \times (x - \bar{x}) + \kappa_t^{(3)} \times (x - x_{old})^+ + \epsilon_{x,t}$$

Pour mesurer l'évolution des paramètres temporels sur l'horizon de projection, des méthodes distinctes seront utilisées entre le paramètre de tendance générale, soit κ pour Lee-Carter et $\kappa^{(1)}$ pour Cairns-Blake-Dowd et Börger-Fleischer-Kuksin alternatif, et les autres paramètres temporels, soit $\kappa^{(2)}$ pour Cairns-Blake-Dowd et $\kappa^{(2)}$, $\kappa^{(3)}$ Börger-Fleischer-Kuksin alternatif.

La méthode de projection de la tendance générale consiste en une régression linéaire pondérée avec intégration d'un aléa, établi par erreur quadratique moyenne pondérée, pour les scénarios stochastiques :

$$\kappa_{t_{proj}}^{(i)} = \kappa_{t_M}^{(i)} + (t_{proj} - t_M) \times \text{Slo}_h^{(i)}(t_M) + \text{Int}_h^{(i)}(t_M) + \epsilon_t \times \sigma_{Global,h}$$

Quant aux autres paramètres temporels, l'approche consiste en une projection de l'écart moyen historique :

$$\kappa_{t_{proj}}^{(i)} = \kappa_{t_M}^{(i)} + (t_{proj} - t_M) \times \frac{\sum_{k=t_m}^{t_M-1} \kappa_{k+1}^{(i)} - \kappa_k^{(i)}}{(t_M - 1) - t_m + 1}$$

Pour améliorer la qualité des ajustements aux âges avancés, où la quantité de données est souvent limitée, une fermeture de tables est nécessaire. Dans le cadre de cette étude, la méthode de Denuit-Goderniaux (2005) est retenue, consistant à ajuster un modèle log-quadratique en imposant des conditions de concavité et d'âge limite :

$$\ln(q_x) = a + bx + cx^2 + \epsilon_x$$

Grâce à ces méthodes, on dispose de quotients de mortalité bicentennaires projetés, néanmoins, pour un assureur exerçant dans divers domaines, plusieurs tables de mortalité *Best Estimate* peuvent être nécessaires. Afin de définir un choc adapté à chacune des tables, on décide d'utiliser les *Reduction Rates*, correspondant aux taux d'améliorations cumulées :

$$RR_{x,t}^{Bicentenaire} = \frac{q_{x,t}^{Bicentenaire}}{q_{x,t_M}}$$

Impact de la COVID-19 sur les scénarios bicentennaires de longévité

L'intégration des données de 2020, année singulière en raison de la pandémie, à l'historique de calibrage affecte le paramètre de tendance générale. Cette modification se manifeste à travers trois effets distincts : La translation, la pentification et la diffusion.

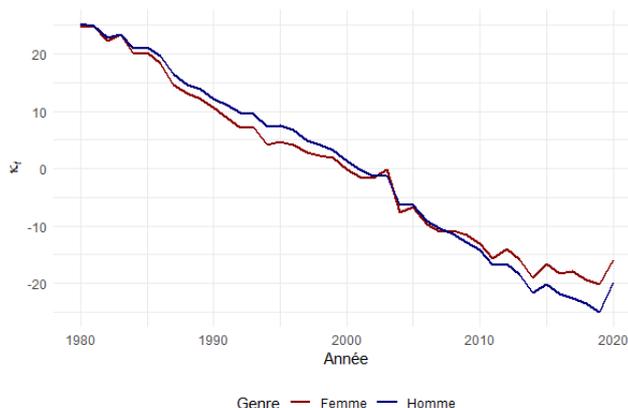


FIGURE 1: Paramètre κ_t par genre pour le modèle de Lee-Carter

Comme l'illustre le graphique précédent, la nouvelle observation, marquée par la pandémie, ajoutée à la série des paramètres de tendance générale montre un écart notable par rapport aux tendances historiques. La projection de ce paramètre démarrante à partir de cette nouvelle donnée, cela entraîne une translation de la tendance générale projetée.

Ensuite, la projection du paramètre étant établie par régression linéaire pondérée, l'intégration de 2020 ajuste la pentification du paramètre projeté à la hausse, suggérant une tendance aux améliorations dans le futur amoindrie.

Enfin, l'impact de diffusion est notable sur le paramètre de volatilité des scénarios stochastiques. La volatilité étant mesurée par erreur quadratique moyenne pondérée, les données liées à la COVID-19 contribuent sensiblement à l'augmentation de cette volatilité. Cela est dû à l'importante erreur d'estimation provoquée par l'écart de cette observation par rapport aux tendances passées, et au fait qu'une pondération conséquente est accordée à cette erreur.

L'application de ces scénarios bicentennaires au portefeuille de retraite d'AXA France a pour effet de réduire le besoin en capital pour le risque de longévité :

	LC		CBD		BFK alternatif	
	2019	2020	2019	2020	2019	2020
SCR Longévité	177 061	167 858	148 504	144 615	163 955	153 445
Évolution 20/19	-5.2%		-2.6%		-6.4%	

TABLE 1: Impact de l'intégration de la COVID-19 sur le SCR de Longévité par modèle

Ce constat peu intuitif s'explique par des effets de translation et pentification supérieurs à la hausse de volatilité :

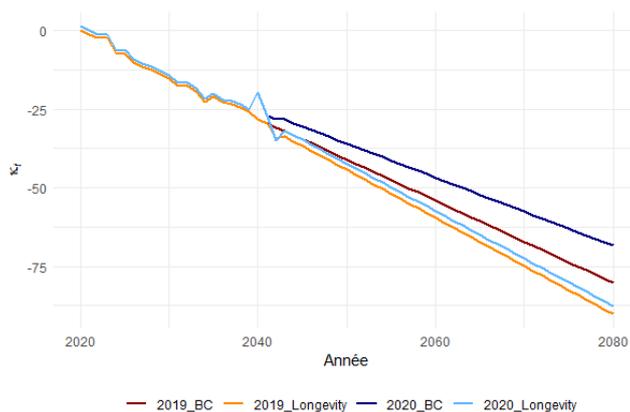


FIGURE 2: Illustration de l'effet compensatoire pour le modèle de Lee-Carter

La hausse de volatilité se manifeste par un écart plus important entre le scénario central (ou *Base Case*) et le scénario bicentenaire. La translation et la pentification sont quant à elles observables par une tendance du scénario central sensiblement plus élevée lorsque les données relatives à la pandémie sont intégrées dans l'historique de calibrage.

Toutefois, l'effet d'une pandémie ne devrait pas être intégré dans des scénarios de tendance de mortalité car la norme Solvabilité 2 prévoit un module dédié à ce risque, et inclure l'impact de la pandémie pourrait conduire à un double comptage. En outre, adopter des scénarios moins conservateurs, surtout dans le cadre de la norme, peut être contestable compte tenu de la particularité de l'année intégrée. L'étude propose alors diverses approches permettant d'élaborer des scénarios robustes qui intègrent les données liées à la COVID-19.

Traitements envisageables permettant l'intégration des données COVID-19

Tout d'abord, l'attention s'est portée sur la sélection du modèle le plus adapté pour élaborer les scénarios bicentennaires compte tenu de l'impact de la pandémie sur les données de calibrage. Le modèle alternatif de Börger-Fleischer-Kuksin, mis en avant dans cette étude, se distingue comme étant particulièrement approprié à cet égard. Les paramètres proposés par ce dernier minimisent l'impact de la pandémie sur la tendance générale, entraînant ainsi une augmentation de la volatilité de 7% et ainsi un taux d'absorption de la pandémie de 99%. Cela contraste des modèles de Lee-Carter et Cairns-Blake-Dowd pour lesquels l'augmentation de volatilité est respectivement de 54% et 23% et le taux d'absorption de 80% et 93%.

La réduction de prudence des scénarios bicentennaires de longévité suite à l'ajout d'une année marquée par la COVID-19 est peu recevable. La première approche proposée vise à ajuster théoriquement le nombre de décès de 2020 en le supposant égal à celui estimé par nos modèles pour cette même année. Cependant, cette méthode génère des résultats incohérents pour certains genres et modèles et a donc été écartée pour ces derniers. Ces incohérences proviennent majoritairement des lacunes dans la méthode de fermetures de tables, en effet, une sur-estimation de la mortalité au point de jonction a été détectée lors de la sélection des modèles. Cet écart influence l'estimation des décès dus à la COVID-19, entraînant un nombre accru de décès au voisinage de la jonction :

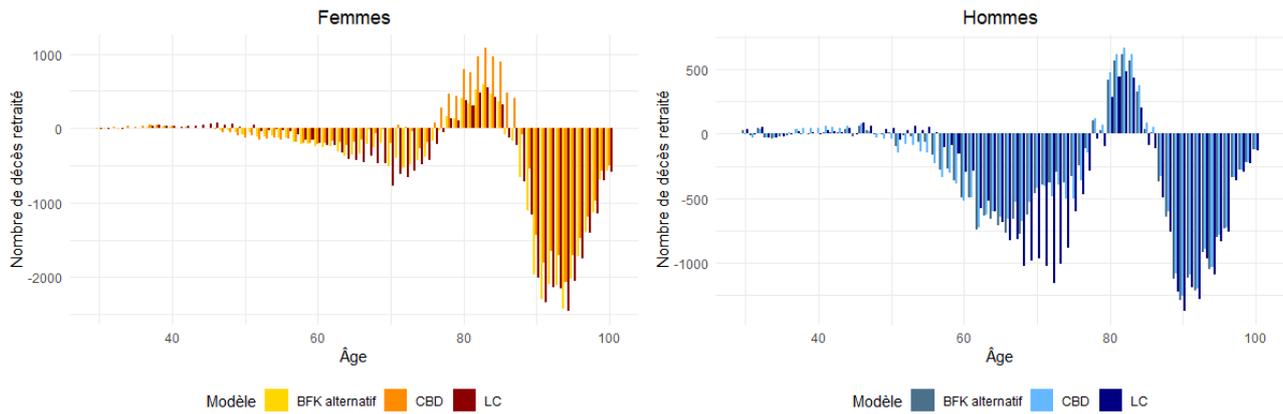


FIGURE 3: Nombre de décès retraités en 2020 par modèle

Les retraitements sélectionnés améliorent légèrement l'adéquation des scénarios bicentennaires, en rapprochant davantage le capital de solvabilité requis pour le risque de longévité à celui d'avant pandémie :

	LC		CBD		BFK alternatif	
	2019	2020	2019	2020	2019	2020
SCR Longévité	177 061	171 497	148 504	146 310	163 955	156 439
Évolution 20/19	-3.1%		-1.5%		-4.6%	

TABLE 2: Impact du retraitement du nombre de décès sur le SCR de Longévité par modèle

Néanmoins, les incertitudes soulevées par les constats précédents nécessiterait l'ajout d'un surplus de volatilité. Étant donné la complexité engendrée par l'inclusion de cette volatilité additionnelle, il serait préférable d'exclure complètement les années impactées par la pandémie. Cette approche a pour avantage d'être à la fois simple et permet d'éviter l'ajout du complément de volatilité destiné à compenser le manque de fiabilité des hypothèses de retraitements.

La dernière approche suggérée, qui ne requiert pas d'ajout de volatilité et n'introduit donc pas de complexité opérationnelle supplémentaire, consiste à affiner la modélisation en proposant des paramètres de pondération ajustés dans le cadre de la pandémie. Une pondération déterministe réduite est ainsi choisie lorsqu'une donnée atypique, qui pourrait fausser les tendances futures, est intégrée dans l'historique de calibrage. En ce qui concerne les pondérations utilisées pour la tendance post-événement bicentenaire et pour le calcul de la volatilité, les niveaux actuels sont considérés comme appropriés et devraient alors être conservés. L'ajustement du paramètre de pondération déterministe vise à retrouver des améliorations de mortalité semblables à celles observées avant la pandémie. Cet ajustement permet notamment d'obtenir des niveaux de capital de solvabilité requis nettement plus cohérents et spécifiquement pour le modèle de Börger-Fleischer-Kuksin alternatif :

	LC		CBD		BFK alternatif	
	2019	2020	2019	2020	2019	2020
SCR Longévité	177 061	173 199	148 504	147 227	163 955	167 831
Évolution 20/19	-2.2%		-0.9%		+2.4%	

TABLE 3: Impact de l'ajustement de la pondération sur le SCR de Longévité par modèle

Conclusion

Ce mémoire présente diverses méthodes permettant une modélisation adéquate dans un contexte où les données de calibrage sont perturbées par la pandémie de COVID-19.

Le modèle de Börger-Fleischer-Kuksin alternatif permet une absorption quasi-totale de la hausse de volatilité liée à la pandémie et est donc à l'origine, tout comme les autres modèles étudiés, d'une réduction de prudence des scénarios bicentennaires de longévité.

Face à cela, une méthode d'évaluation des décès liés à la COVID-19 est établie, en supposant le nombre de décès en 2020 égal à celui prédit par nos modèles. Cette méthode démontre certaines limites puisqu'elle réplique les potentielles faiblesses de modélisation, une exclusion totale des années affectées par la pandémie est alors suggérée. Cela permet notamment d'éviter d'introduire une volatilité additionnelle destinée à pallier les incertitudes des estimations.

La dernière approche, ne nécessitant pas d'hypothèses supplémentaires, consiste en une revue des paramètres de pondération utilisés dans la modélisation. Cette revue mène à l'adoption d'une pondération déterministe de remplacement et conduit à des résultats plus en ligne avec ceux pré-pandémie, en particulier pour le modèle mis en avant dans ce mémoire.

L'étude se concentrant sur un cadre théorique spécifique, il en résulte que plusieurs éléments n'ont pas été abordés, ouvrant ainsi des perspectives pour des études ultérieures. Ces perspectives incluent l'élargissement de l'analyse à d'autres modèles ou méthodes de fermeture de tables, la revue de la méthode de définition des chocs bicentennaires, l'intégration des autres années impactées par la pandémie ou encore la paramétrisation de la volatilité additionnelle.

Context

The Solvency 2 reform is implemented for insurers in the European Union to ensure the financial system's stability by establishing a strict prudential framework. This reform notably introduces the concept of Solvency Capital Requirement (SCR), which corresponds to the minimum amount of economic own funds a company must have to cover an unfavorable event with a bicentennial occurrence probability. In this context, the regulator defines different risk factors, including longevity risk, referring to the risk of underestimating the survival probabilities of the insured.

For several decades, France has benefited from a regular increase in life expectancy due to medical advances and social and behavioral evolutions. However, in 2020, France experienced the most significant decrease in life expectancy at birth of the last seventy years, a phenomenon mainly explained by the global COVID-19 pandemic.

To define local bicentennial longevity scenarios under Solvency 2, the practice is to use duration models. These models help estimate future mortality trends using past observations. Consequently, including pandemic-related data in the calibration history significantly alters the trajectories of these scenarios.

The aim of this thesis is to assess the consequences of integrating COVID-19 related data on bicentennial longevity scenarios and to suggest approaches that allow for adequate modeling by taking this information into account.

Methodology

To assess the sensitivity of various models to the integration of COVID-19 data and to measure the effects of the proposed approaches differently, three duration models are studied in this thesis.

The Lee-Carter model (1992) is selected as a reference in the literature on duration models, providing an expression of the instantaneous mortality rate of the following form :

$$\ln \mu_{x,t} = \alpha_x + \beta_x \times \kappa_t + \epsilon_{x,t}$$

Then, the Cairns-Blake-Dowd model (2006), also well-established in the literature, distinguishes itself particularly by including a second time parameter, providing an expression for the mortality quotient of the following form :

$$\text{logit}q_{x,t} = \kappa_t^{(1)} + \kappa_t^{(2)} \times (x - \bar{x}) + \epsilon_{x,t}$$

Finally, the alternative Börger-Fleischer-Kuksin model, significantly less common, has the advantage of incorporating a third time parameter specific to older ages, making it particularly suitable in the context of this study. The model provides an expression for the mortality quotient of the following form :

$$\text{logit}q_{x,t} = \alpha_x + \kappa_t^{(1)} + \kappa_t^{(2)} \times (x - \bar{x}) + \kappa_t^{(3)} \times (x - x_{old})^+ + \epsilon_{x,t}$$

To measure the evolution of the temporal parameters over the projection horizon, distinct methods will be used between the general trend parameter, namely κ for Lee-Carter and $\kappa^{(1)}$ for Cairns-Blake-Dowd and the alternative Börger-Fleischer-Kuksin, and the other temporal parameters, namely $\kappa^{(2)}$ for Cairns-Blake-Dowd and $\kappa^{(2)}$, $\kappa^{(3)}$ for Börger-Fleischer-Kuksin alternative.

The method for projecting the general trend consists of a weighted linear regression with the integration of a random component, established by weighted mean square error, for the stochastic scenarios :

$$\kappa_{t_{proj}}^{(i)} = \kappa_{t_M}^{(i)} + (t_{proj} - t_M) \times \text{Slo}_h^{(i)}(t_M) + \text{Int}_h^{(i)}(t_M) + \epsilon_t \times \sigma_{Global,h}$$

As for the other temporal parameters, the approach consists of projecting the historical average deviation :

$$\kappa_{t_{proj}}^{(i)} = \kappa_{t_M}^{(i)} + (t_{proj} - t_M) \times \frac{\sum_{k=t_m}^{t_M-1} \kappa_{k+1}^{(i)} - \kappa_k^{(i)}}{(t_M - 1) - t_m + 1}$$

To improve the quality of adjustments at advanced ages, where data is often limited, a table closure is necessary. In this study, the Denuit-Goderniaux (2005) method is adopted, which involves fitting a log-quadratic model by imposing concavity conditions and a limit age :

$$\ln(q_x) = a + bx + cx^2 + \epsilon_x$$

Thanks to these methods, we have projected bicentennial mortality quotients, however, for an insurer operating in various fields, several Best Estimate mortality tables may be necessary. In order to define a unique shock applicable to each table, it was decided to use the Reduction Rates, corresponding to cumulative improvement rates :

$$RR_{x,t}^{Bicentennial} = \frac{q_{x,t}^{Bicentennial}}{q_{x,t_M}}$$

Impact of COVID-19 on bicentennial longevity scenarios

The integration of data from 2020, a unique year due to the pandemic, into the calibration history affects the general trend parameter. This modification manifests through three distinct effects : Translation, steepening, and diffusion.

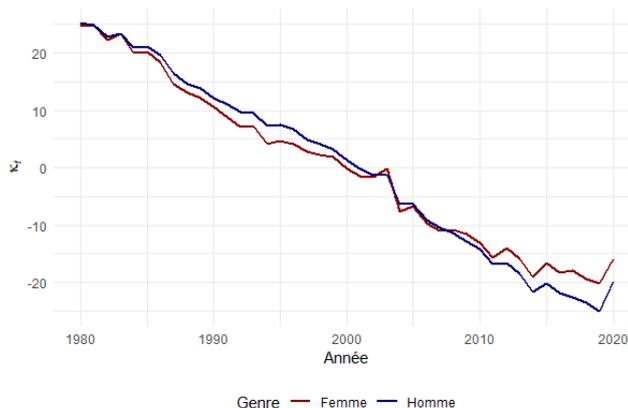


FIGURE 4: Parameter κ_t by gender for the Lee-Carter model

As illustrated by the previous graph, the new observation, marked by the pandemic, added to the series of general trend parameters shows a notable deviation from historical trends. The projection of this parameter starting from this new data leads to a translation of the projected general trend.

Then, since the projection of the parameter is established by weighted linear regression, the integration of 2020 adjusts the steepening of the projected parameter upwards, suggesting a trend of diminished improvements in the future.

Finally, the diffusion impact is significant on the volatility parameter of the stochastic scenarios. Since volatility is measured by weighted mean square error, the COVID-19 related data substantially contribute to increasing this volatility. This is due to the significant estimation error caused by the deviation of this observation from past trends and the fact that considerable weighting is given to this error.

The application of these bicentennial scenarios to the AXA France pension portfolio has the effect of reducing the solvency capital requirement for longevity risk :

	LC		CBD		BFK alternative	
	2019	2020	2019	2020	2019	2020
Longevity SCR	177 061	167 858	148 504	144 615	163 955	153 445
Evolution 20/19	-5.2%		-2.6%		-6.4%	

TABLE 4: Impact of COVID-19 integration on Longevity SCR by model

This counterintuitive observation is explained by translation and steepening effects that outweigh the increase in volatility :

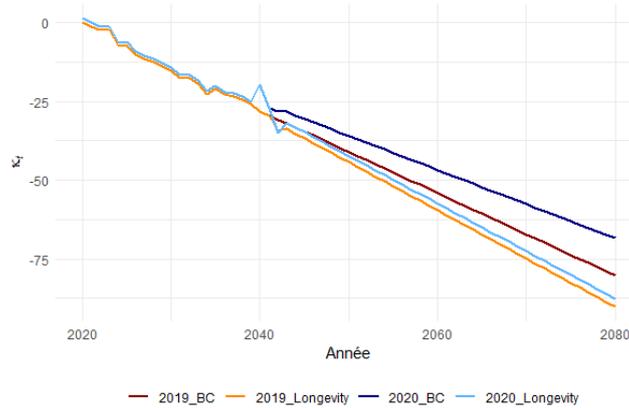


FIGURE 5: Illustration of the compensatory effect for the Lee-Carter model

The increase in volatility is manifested by a larger gap between the central scenario (or Base Case) and the bicentennial scenario. The translation and steepening, meanwhile, are observable by a substantially higher trend of the central scenario when pandemic-related data are integrated into the calibration history.

However, the effect of a pandemic should not be integrated into mortality trend scenarios because the Solvency 2 standard includes a dedicated module for this risk, and incorporating the pandemic's impact could lead to double counting. Moreover, adopting less conservative scenarios, especially within the framework of the standard, can be questionable given the uniqueness of the integrated year. The study then proposes various approaches to develop robust scenarios that incorporate COVID-19 related data.

Potential treatments for COVID-19 data integration

Firstly, attention was focused on selecting the most suitable model to develop bicentennial scenarios considering the impact of the pandemic on calibration data. The alternative model of Börger-Fleischer-Kuksin, highlighted in this study, stands out as particularly appropriate in this regard. The parameters proposed by this model minimize the pandemic's impact on the general trend, thereby resulting in a 7% increase in volatility and hence a 99% pandemic absorption rate. This contrasts with the Lee-Carter and Cairns-Blake-Dowd models, for which the volatility increase is 54% and 23% respectively, and the absorption rate is 80% and 93%.

The reduction in prudence of bicentennial longevity scenarios following the addition of a year marked by COVID-19 is not very acceptable. The first proposed approach aims to theoretically adjust the number of deaths in 2020, assuming it to be equal to that estimated by our models for the same year. However, this method generates inconsistent results for certain genders and models and has therefore been discarded for them. These inconsistencies mostly arise from flaws in the table closure method, indeed, an overestimation of mortality at the junction point was detected during model selection. This gap influences the estimation of COVID-19 related deaths, leading to an increased number of deaths in the vicinity of the junction :

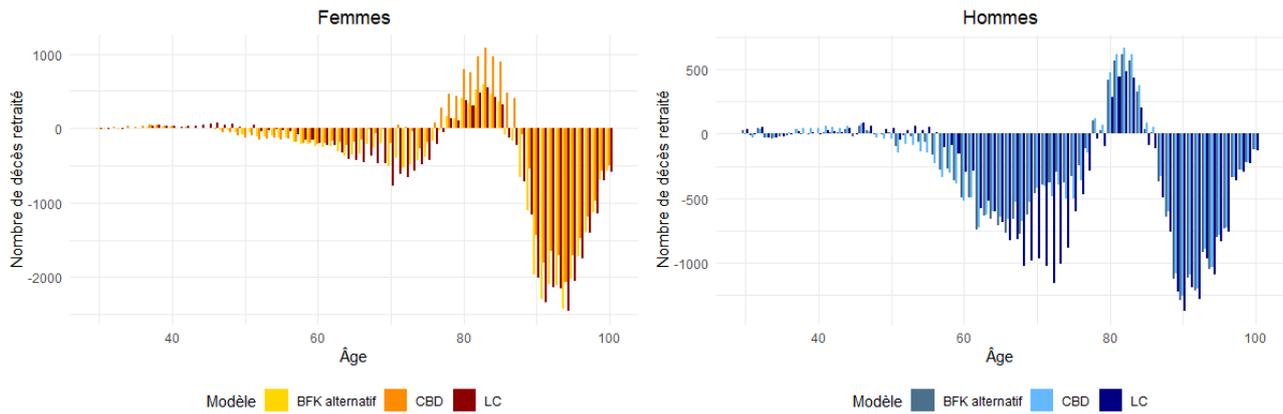


FIGURE 6: Number of deaths suppressed in 2020 by model

The selected adjustments slightly improve the adequacy of the bicentennial scenarios, bringing the solvency capital requirement for longevity risk closer to that before the pandemic :

	LC		CBD		BFK alternative	
	2019	2020	2019	2020	2019	2020
Longevity SCR	177 061	171 497	148 504	146 310	163 955	156 439
Evolution 20/19	-3.1%		-1.5%		-4.6%	

TABLE 5: Impact of adjusting for the number of deaths on Longevity SCR by model

Nevertheless, the uncertainties raised by the previous findings would require the addition of extra volatility. Given the complexity caused by including this additional volatility, it would be preferable to completely exclude the years affected by the pandemic. This approach has the advantage of being simple and avoids adding extra volatility intended to compensate for the unreliability of the treatment assumptions.

The last suggested approach, which does not require the addition of volatility and thus does not introduce additional operational complexity, consists of refining the modeling by proposing adjusted weighting parameters in the context of the pandemic. A reduced deterministic weighting is thus chosen when an atypical data point, which could skew future trends, is integrated into the calibration history. Regarding the weightings used for the post-bicentennial event trend and for calculating volatility, the current levels are considered appropriate and should then be maintained. The adjustment of the deterministic weighting parameter aims to recover mortality improvements similar to those observed before the pandemic. This adjustment notably allows obtaining solvency capital requirement levels that are significantly more consistent and specifically for the alternative Börger-Fleischer-Kuksin model :

	LC		CBD		BFK alternative	
	2019	2020	2019	2020	2019	2020
Longevity SCR	177 061	173 199	148 504	147 227	163 955	167 831
Evolution 20/19	-2.2%		-0.9%		+2.4%	

TABLE 6: Impact of weighting adjustment on Longevity SCR by model

Conclusion

This thesis presents various methods for adequate modeling in a context where calibration data are disrupted by the COVID-19 pandemic.

The alternative Börger-Fleischer-Kuksin model allows for an almost total absorption of the increased volatility related to the pandemic and is therefore, like the other models studied, at the origin of a reduction in the prudence of bicentennial longevity scenarios.

In response, a method for assessing COVID-19 related deaths is established, assuming the number of deaths in 2020 to be equal to that predicted by our models. This method demonstrates certain limits since it replicates potential modeling weaknesses, a total exclusion of the years affected by the pandemic is then suggested. This notably helps avoid introducing additional volatility intended to compensate for the uncertainties of the estimates.

The last approach, which does not require additional assumptions, involves a review of the weighting parameters used in the modeling. This review leads to the adoption of an alternative deterministic weighting and results in outcomes more in line with pre-pandemic results, particularly for the model highlighted in this thesis.

The study, focusing on a delimited theoretical context, results in several elements not being addressed, thus opening prospects for future studies. These prospects include expanding the analysis to other models or table closure methods, revising the methodology for defining bicentennial shocks, integrating other years impacted by the pandemic, and parameterizing additional volatility.

Je tiens avant tout à remercier la direction du Risk Management (RMF) d'AXA France, et tout particulièrement Aurélie HAYNES, ma tutrice en entreprise, pour son accompagnement précieux et ses conseils avisés. Sa confiance en mes capacités et son expertise ont grandement contribué à mon développement professionnel et ont significativement influencé l'orientation de mes recherches.

Je souhaite également remercier l'EURO Institut d'Actuariat (EURIA), et spécifiquement à Marine HABART, ma tutrice académique, pour sa disponibilité et son engagement. Ses critiques constructives et son aspiration à l'excellence m'ont permis de développer ma capacité d'analyse et d'affiner mon étude.

Un merci particulier à ma compagne, pour son encouragement constant et son soutien indéfectible, fondamental pour l'accomplissement de ce mémoire.

Enfin, je souhaite adresser mes remerciements les plus sincères à toutes les personnes qui, de près ou de loin, ont contribué à la réalisation de ce mémoire. J'exprime ainsi ma gratitude envers ma famille, mes amis, mes collègues et l'ensemble des professeurs pour leur soutien. Ce travail, couronnant des années d'études et de rigueur, est le fruit d'un effort collectif et je suis profondément reconnaissant envers chaque contribution qui a permis sa réalisation.

Résumé	1
Abstract	2
Note de synthèse	3
Executive Summary	9
Remerciements	15
Introduction	18
I Contexte de l'étude	19
1 Solvabilité 2	21
1.1 Cadre général	21
1.2 Pilier 1 : Les exigences quantitatives	23
1.3 Pilier 2 : Les exigences qualitatives	24
1.4 Pilier 3 : Les informations à destination du public et du superviseur	25
1.5 Capital de Solvabilité Requis	25
2 Risque de mortalité	29
2.1 Définition	29
2.2 Évolution de la mortalité en France	30
2.3 COVID-19	31
II Méthodes d'estimation de la mortalité prospective	34
3 Notations, définitions et hypothèses de travail	36
3.1 Notations et définitions	36
3.2 Hypothèses de travail	39
4 Modèles de durée	40
4.1 Lee-Carter	40
4.2 Cairns-Blake-Dowd	43
4.3 Börger-Fleischer-Kuksin alternatif	45
4.4 Projection de la tendance	47
4.4.1 Projection déterministe	47

4.4.2	Projection stochastique	48
4.5	Fermeture de tables de mortalité	49
4.6	Chocs bicentennaires	50
5	Critères de sélection de modèle	51
5.1	Critères statistiques	51
5.1.1	BIC : Critère d'information bayésien	51
5.1.2	SMR : Ratio de mortalité standardisé	52
5.2	Critères de prudence	52
5.2.1	Amélioration d'espérance de vie	52
5.2.2	<i>Backtesting</i>	53
III	Étude de l'impact COVID-19 sur l'estimation de la mortalité prospective	54
6	Analyse de l'impact de la COVID-19 sur l'estimation de la mortalité prospective	56
6.1	Données de l'étude	56
6.2	Paramètres de modélisation	58
6.3	Sélection de modèle	59
6.4	Impact de la COVID-19 sur l'estimation de la mortalité prospective	65
7	Traitements envisageables pour la prise en compte de la COVID-19 pour les scénarios bicentennaires	79
7.1	Définition du portefeuille de mesure d'impact	79
7.2	Robustesse du modèle de Börger-Fleischer-Kuksin alternatif	81
7.3	Retraitement des données de mortalité	83
7.4	Ajustement des paramètres et de la modélisation	90
	Conclusion	96

Afin de faire face aux particularités du secteur de l'assurance et notamment son cycle de production inversé, la réforme Solvabilité 2 encadre les entreprises d'assurance de l'Union Européenne, visant à garantir leur solvabilité dans un cadre prudentiel structuré en trois piliers : Les exigences quantitatives, qualitatives et de communication d'informations. Le pilier quantitatif introduit la notion de Capital de Solvabilité Requis (SCR), correspondant au montant de fonds propres économiques dont doit disposer la compagnie afin d'avoir une probabilité de ruine inférieure à 0.5% sur 1 an. Parmi les facteurs de risques identifiés par le superviseur, on retrouve le risque de longévité, correspondant au risque de sous-estimation des probabilités de survie des assurés.

Depuis plusieurs décennies, l'espérance de vie en France augmente régulièrement, soutenue non seulement par les avancées médicales dans le domaine de la prévention, la détection et le traitement des maladies, mais aussi par les évolutions sociales et comportementales incluant une alimentation et des conditions de travail améliorées. Toutefois, la population Française a connu en 2020 la plus importante baisse d'espérance de vie à la naissance des soixante-dix dernières années, principalement due à la pandémie de COVID-19. Ce virus, entraînant des complications respiratoires, s'est avéré particulièrement fatal pour les individus âgés et/ou vulnérables.

Or, dans l'évaluation locale des scénarios bicentennaires de longévité sous Solvabilité 2, les assureurs utilisent des modèles de durée se basant sur les données passées, avec pour vocation d'estimer les tendances de mortalité futures. Ainsi, l'intégration des observations liées à la pandémie a pour effet de déformer les trajectoires des scénarios. Ce mémoire s'articulera autour de la problématique suivante :

Quels sont les traitements envisageables sur les scénarios bicentennaires de longévité permettant l'intégration des données COVID-19 ?

Pour répondre à cette problématique, le mémoire se divise en trois parties, la première explorera la norme Solvabilité 2 en mettant l'accent sur la notion de SCR. Elle examinera également le risque de longévité et sa déclinaison en France, avec une attention particulière sur la pandémie de COVID-19.

La seconde partie approfondira le cadre théorique, en détaillant les modèles de Lee-Carter, Cairns-Blake-Dowd et une variante de Börger-Fleischer-Kuksin, tout en expliquant les méthodes de projection des paramètres temporels, de fermeture de tables et de définition des chocs bicentennaires. Cette partie traitera également des critères de sélection de modèle et de validation des scénarios bicentennaires.

Enfin, dans la troisième partie, on examinera l'impact de la mortalité COVID-19 dans chaque étape de l'élaboration des scénarios bicentennaires de longévité. En réponse à ces observations, diverses alternatives seront proposées. Ces alternatives incluent le choix du modèle le plus approprié parmi ceux étudiés, les méthodes de retraitement des décès dus à la COVID-19, et une section consacrée aux ajustements des paramètres et aux axes d'amélioration de la modélisation.

NB : Pour des raisons de confidentialité, toutes les valeurs numériques d'AXA France présentées dans ce mémoire ne correspondent pas à des valeurs réelles mais à des valeurs modifiées par une méthode non précisée. Pour autant, ces valeurs ajustées fournissent une indication fiable des tendances observées et les résultats restent comparativement pertinents entre eux.

Première partie
Contexte de l'étude

Dans cette première partie, on présentera la norme Solvabilité 2 qui est le régime réglementaire de l'Union Européenne s'appliquant aux compagnies d'assurance et de réassurance en mettant spécifiquement l'accent sur les exigences en matière de capital de solvabilité requis. Par la suite, notre attention se portera sur le risque de mortalité tel que définit sous cette norme, et plus généralement, la traduction de ce risque en France avec une attention particulière à la pandémie de COVID-19. Enfin, les interactions entre ces deux chapitres permettront d'établir la problématique de ce mémoire.

1.1 Cadre général

Toute entreprise d'assurance se doit d'être solvable auprès de ses assurés afin de garantir la pérennité du système financier. De part les spécificités de l'assurance, et notamment son fonctionnement en cycle inversé, les assureurs sont plus que quiconque dans la nécessité d'être en mesure d'évaluer leurs risques.

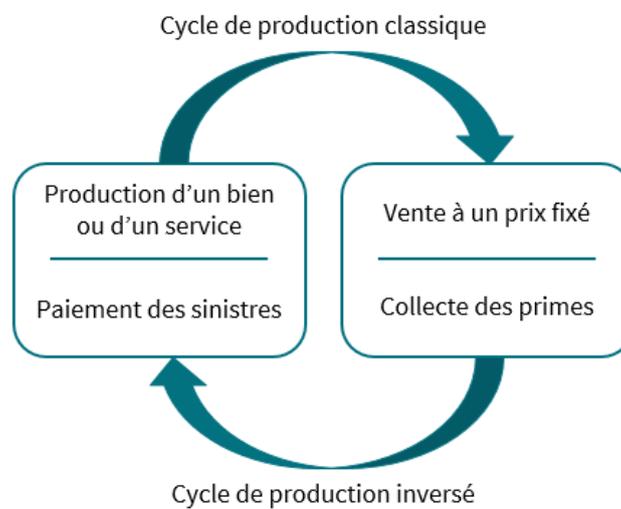


FIGURE 1.1: Comparaison entre un cycle de production classique et inversé

Afin d'encadrer les assureurs dans leur gestion des risques, le superviseur a mis en place une norme prudentielle s'appliquant aux entreprises d'assurances de l'Union Européenne, Solvabilité 2 [1], entrée en application le 1^{er} janvier 2016. Cette norme avait pour objectif de remplacer Solvabilité 1 qui admettait certaines limites :

- Un régime prudentielle trop simpliste, sans prise en compte de la diversité des risques portés par les entreprises d'assurances ;
- Les aspects groupe peu ou mal appréhendés, le régime s'appuyant principalement sur le contrôle "solo" ;
- Une communication financière et harmonisation européenne insuffisante, génératrice de limites au développement d'un marché européen unique de l'assurance.

Les quatre principaux objectifs de Solvabilité 2 sont les suivants :

- Renforcer la résistance et la solvabilité des assureurs, la sécurité des souscripteurs et donc la stabilité du système financier européen ;
- Améliorer la prise en compte de la pluralité des risques auxquels sont confrontés les organismes d'assurance et adapter les exigences réglementaires aux risques encourus ;
- Harmoniser la réglementation prudentielle du secteur de l'assurance de l'espace économique européen ainsi que les principes et pratiques de contrôle ;
- Harmoniser les *reportings* en intégrant dans la conception du dispositif les exigences réglementaires en matière de qualité des données.

Pour respecter ces objectifs, la norme se décline en 3 piliers :

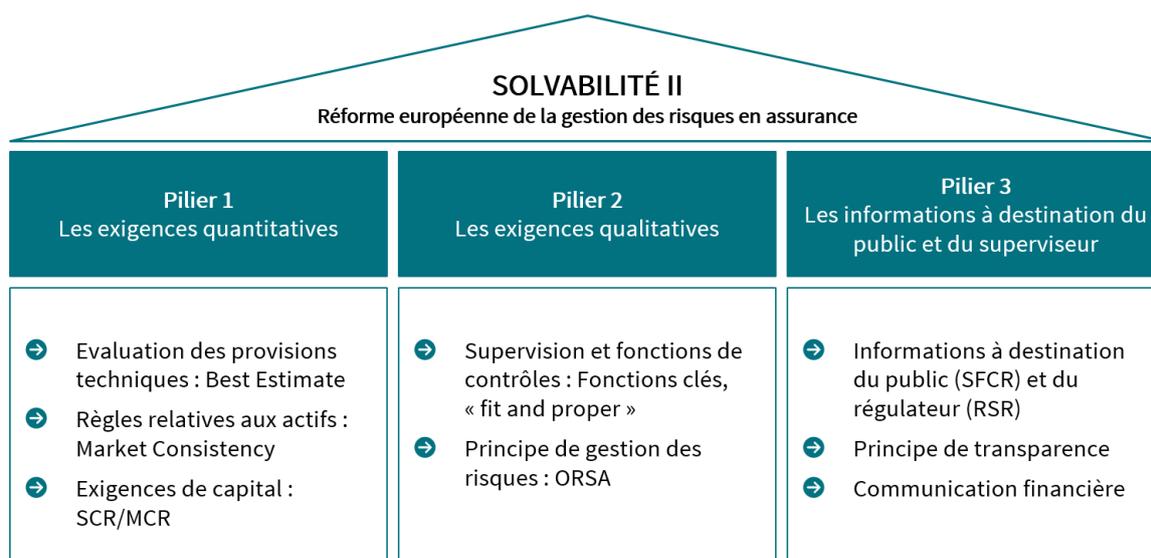


FIGURE 1.2: Piliers de Solvabilité 2

1.2 Pilier 1 : Les exigences quantitatives

Le pilier 1 regroupe les exigences quantitatives, soit les règles de valorisation des actifs et des passifs, ainsi que les exigences de capital et leur mode de calcul.

Le bilan sous Solvabilité 2 se présente de la forme suivante :

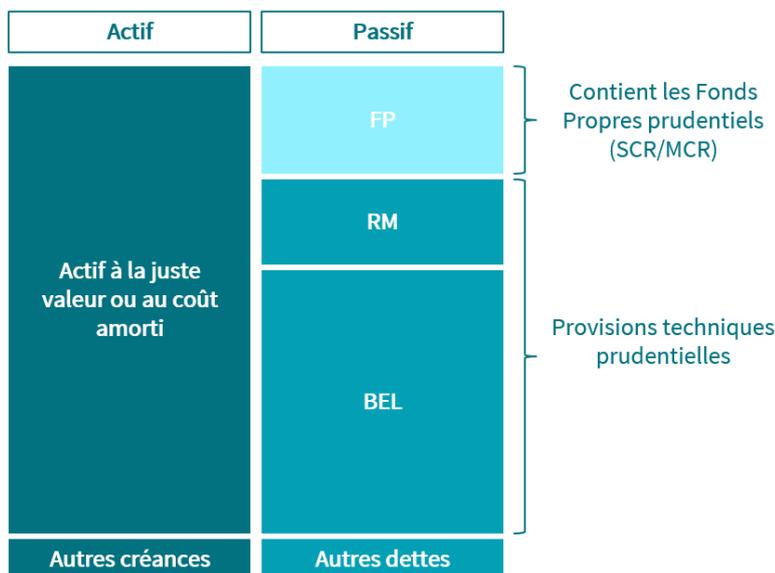


FIGURE 1.3: Bilan sous Solvabilité 2

Le bilan représente la situation patrimoniale d'une compagnie à un instant T , ce dernier montre ce que la compagnie possède, à l'actif, et ce qu'elle doit, au passif. Les fonds propres (FP), inscrits au passif, définissent les ressources stables de l'entreprise.

Sous Solvabilité 2, les fonds propres correspondent aux ressources de l'assureur immédiatement disponibles afin d'honorer ses engagements auprès des assurés dans un contexte défavorable. Ils doivent être en mesure de couvrir les exigences en capital de la norme, soit :

- Le SCR¹ : Niveau de capital cible nécessaire pour que la probabilité de faillite de l'organisme d'assurance soit inférieure à 0,5 % à l'horizon d'un an. Il est calculé à partir d'une formule standard ou d'un modèle interne, pouvant être partiel ;
- Le MCR² : Exigence minimale de fonds propres en dessous de laquelle l'intervention de l'autorité de contrôle est automatique.

L'actif représente les investissements effectués par l'entreprise, majoritairement obligataires en assurance vie, valorisé à la juste valeur ou au coût amorti et tient donc compte des évolutions de marché :

- Juste valeur : Montant auquel l'actif pourrait être échangé entre des parties bien informées et consentantes dans le cadre d'une transaction effectuée dans des conditions de concurrence normales ;
- Coût amorti : Montant auquel est évalué l'actif lors de sa comptabilisation initiale, diminué des remboursements en principal, majoré ou diminué de l'amortissement cumulé, calculé par la méthode du taux d'intérêt effectif, de toute différence entre ce montant initial et le montant à l'échéance et ajusté au titre de la correction de valeur pour pertes, le cas échéant.

En pratique, les actifs voués à être revendus avant leur échéance sont valorisés à la juste valeur, le reste au coût amorti.

1. *Solvency Capital Requirement* ou Capital de Solvabilité Requis

2. *Minimum Capital Requirement* ou Minimum de Capital Requis

Le BEL³ correspond à la meilleure vision économique des engagements de l'assureur. Ce dernier est égal à la moyenne pondérée par leur probabilité des flux de trésorerie futurs, compte tenu de la valeur temporelle de l'argent, déterminée à partir de la courbe des taux sans risque pertinente. Il est fondé sur des informations actuelles crédibles et des hypothèses réalistes et fait appel à des méthodes actuarielles et des techniques statistiques adéquates.

La RM⁴ correspond au montant réclamé par une entité tierce pour supporter l'incertitude intrinsèque aux flux de trésorerie d'origine non-financière.

Elle est calculée via une approche coût du capital, c'est-à-dire par la valeur actuelle de l'immobilisation du capital de solvabilité requis sur chaque période au coût du capital.

1.3 Pilier 2 : Les exigences qualitatives

Le pilier 2 regroupe l'ensemble des règles de gouvernance et de gestion des risques, ce pilier exige une organisation claire et bien définie à travers la mise en place de processus efficaces pour identifier, évaluer et gérer les risques existants et potentiels.

L'un des points principaux de ce pilier est la mise en place de l'ORSA⁵. Il s'agit d'un processus interne d'évaluation des risques et de la solvabilité par l'organisme. Celui-ci doit illustrer la capacité de l'organisme ou du groupe à identifier, mesurer et gérer les éléments de nature à modifier sa solvabilité ou sa situation financière. L'ORSA comporte obligatoirement trois évaluations :

- L'évaluation du besoin global de solvabilité ;
- L'évaluation du respect permanent des obligations réglementaires concernant la couverture du SCR, du MCR et des exigences concernant le calcul des provisions techniques ;
- L'évaluation de l'écart entre le profil de risque de l'entreprise et les hypothèses qui sous-tendent le capital de solvabilité requis.

La norme prévoit également de mettre en place des fonctions clés, celles-ci ont pour but de garantir une gestion saine et prudente de l'activité d'assurance. Ces fonctions clés sont au nombre de quatre et doivent respecter des exigences de compétences et d'honorabilité (*fit and proper*) :

- Gestion des risques : Porte la responsabilité de l'efficacité du système de gestion des risques, en fixant l'appétence au risque et les limites de tolérance générale au risque de l'entreprise ;
- Audit interne : Établit le plan d'audit, évalue l'organisation et l'efficacité du contrôle interne et peut être amenée à alerter le régulateur en cas de besoin ;
- Contrôle interne : Établit un plan et une politique de contrôle interne avec pour but d'identifier et d'évaluer le risque de non conformité ;
- Actuariat : Coordonne le calcul des provisions techniques, soutient la gestion des risques et donne son avis sur la politique de souscription et de réassurance.

3. *Best Estimate Liabilities*

4. *Risk Margin* ou Marge pour Risque

5. *Own Risk and Solvency Assessment*

1.4 Pilier 3 : Les informations à destination du public et du superviseur

Le dernier pilier de la norme concerne les règles relatives à la communication d'informations au public et au superviseur et vise à harmoniser au niveau européen les informations publiées par les organismes d'assurance ainsi que celles remises aux superviseurs. On peut distinguer deux types de *reportings* :

- Les *reportings* qualitatifs, tels que le SFCR⁶ destiné au public ou le RSR⁷ destiné au superviseur, livrés annuellement, dans lesquels on retrouve par exemple l'activité de la compagnie, ses résultats, sa gouvernance ;
- Les *reportings* quantitatifs, tels que les QRT⁸, regroupant les informations relatives aux fonds propres, SCR ou provisions techniques. Ce rapport est à destination du superviseur et est livré trimestriellement.

1.5 Capital de Solvabilité Requis

Le SCR est le montant de fonds propres économiques, soit la différence entre la valeur de marché des actifs et de la valeur économique des passifs, dont doit disposer la compagnie afin d'avoir une probabilité de ruine inférieure à 0.5% sur 1 an. C'est également la VaR⁹ au quantile 99.5% à horizon 1 an ou le montant de pertes économiques que la compagnie est susceptible d'encourir sur une année basé sur des risques pouvant se produire au moins une fois tous les 200 ans.

En formule standard, le SCR est calculé via une approche modulaire, c'est-à-dire en déterminant tout d'abord le montant de fonds propres économique à détenir pour chaque sous-module. Ensuite, pour tenir compte de la faible probabilité de réalisation simultanée de tous ces événements, la formule standard introduit des corrélations entre ces facteurs de risque et permet donc de réduire leurs impacts. Cette corrélation s'applique à deux niveaux, premièrement entre les sous-modules, pour obtenir le SCR du module et, deuxièmement, entre les modules pour obtenir le BSCR¹⁰. Enfin, intègre le risque opérationnel ne bénéficiant pas de diversification, puis, afin de refléter au mieux la réalité économique, la perte brute observée par l'organisme, soit le BSCR, peut être atténuée par deux effets avant d'obtenir le SCR final :

- La prise en compte de la capacité de transfert d'une partie de sa perte aux assurés via une moindre participation aux bénéfices que celle qu'il avait escomptée avant le choc ;
- L'imputation de la perte au résultat fiscal qui conduira dans la plupart des cas à payer moins d'impôts dans le futur que ceux qui avaient été comptabilisés au bilan initial.

6. *Solvency and Financial Conditions Reports*

7. *Regular Supervisory Report*

8. *Quantitative Reporting Template*

9. *Value at Risk*

10. *Basic Solvency Capital Requirement* ou SCR de base

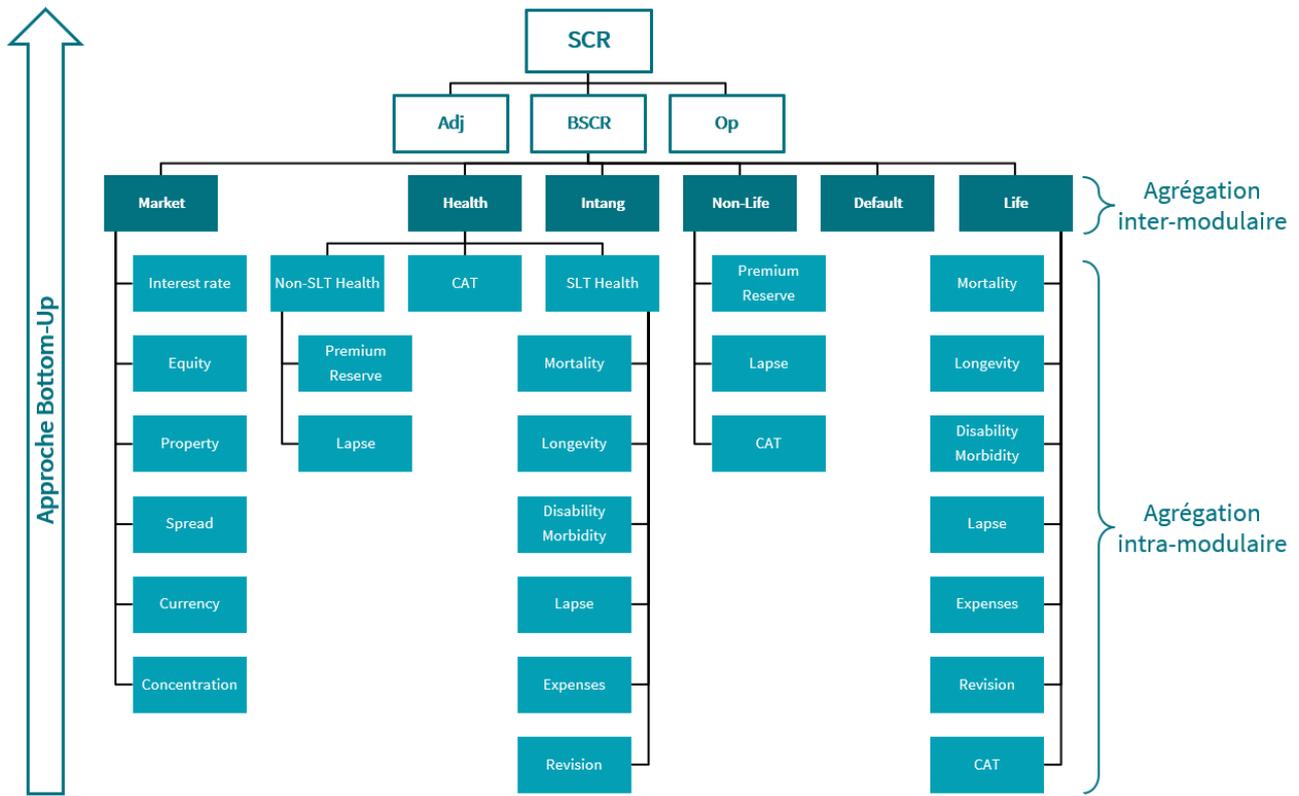


FIGURE 1.4: Approche *Bottom-Up* pour le calcul du SCR

Le SCR est donc défini ainsi :

$$SCR = BSCR + Adj + SCR_{Op}$$

Avec :

$$BSCR = \sqrt{SCR_x^{module} \times \sum_{x \times y} Corr SCR_{x,y}^{module} \times SCR_y^{module}}$$

$$\text{et } Adj = Adj_{FDB} + Adj_{DT}$$

Où :

- SCR_x^{module} = Capital de solvabilité requis pour le module x (par exemple *Market*, *Health* ou *Life*)
- $Corr SCR_{x,y}$ = Corrélation entre les modules x et y
- Adj_{FDB} = Ajustement lié à l'absorption par la participation aux bénéfices
- Adj_{DT} = Ajustement lié à l'absorption dû aux impôts différés
- SCR_{Op} = Capital de solvabilité requis pour le risque opérationnel

Pour chaque module, on a la relation suivante :

$$SCR^{module} = \sqrt{SCR_i^{sous-module} \times \sum_{i \times j} Corr SCR_{i,j}^{sous-module} \times SCR_j^{sous-module}}$$

Avec :

- $SCR_i^{sous-module}$ = Capital de solvabilité requis pour le sous-module i (par exemple *Longevity*, *Equity* ou *Expenses*)
- $CorrSCR_{i,j}^{sous-module}$ = Corrélation entre les sous-modules i et j

Le $SCR_i^{sous-module}$ est quant à lui obtenu par différence entre les fonds propres d'un scénario central et les fonds propres d'un scénario x choqué¹¹, soit :

$$SCR_x^{sous-module} = FP_0^{central} - FP_0^{choc_x}$$

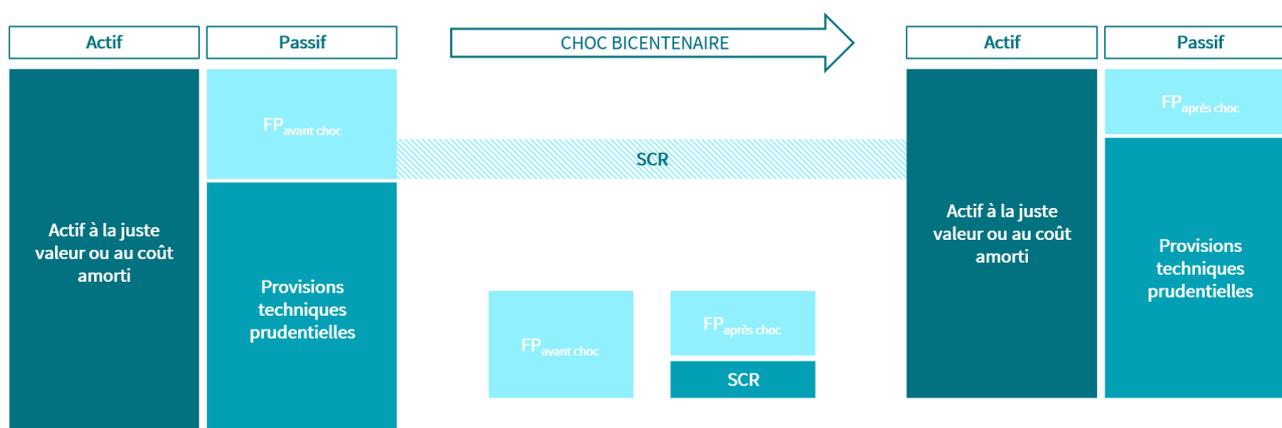


FIGURE 1.5: Calcul du SCR d'un sous-module

Un modèle interne, qu'il soit complet ou partiel, peut également être utilisé par un organisme, en remplacement de la formule standard, pour calculer son capital de solvabilité requis. Cette utilisation est soumise à l'approbation de l'ACPR¹² selon plusieurs critères :

- Critères techniques : La pertinence du périmètre et de la structure du modèle interne, l'adéquation de l'évaluation du bilan à 1 an pour le calcul du SCR ou la justesse des facteurs de risques utilisés ;
- Critères qualitatifs : La qualité de la gouvernance autour du modèle interne, une politique de changement de modèle rigoureuse ou l'intégration du dispositif modèle interne dans le contrôle interne.

Pour les assureurs, les intérêts de disposer d'un modèle interne sont multiples, cela constitue une aide au pilotage de la société, leur permet d'étudier et de maîtriser leurs risques et est un outil essentiel de communication à destination de la communauté financière et des agences de notation.

Un indicateur clé de la norme est le ratio de solvabilité :

$$\text{Ratio de solvabilité} = \frac{\text{Fonds propres économiques}}{\text{Capital de solvabilité requis}}$$

Cet indicateur est une mesure de la robustesse de la compagnie, en cas de ratio inférieur à 100%, la compagnie doit prouver au superviseur qu'elle a mis en place des plans d'action pour revenir à un niveau satisfaisant. En cas de dégradation du ratio jusqu'à atteindre le niveau du MCR, soit $\frac{FP}{MCR} \leq 100\%$, l'organisme doit produire un plan de financement dans un délai d'un mois qui sera soumis à l'approbation l'ACPR avec 3 mois pour rétablir la conformité du MCR. En cas de non retour à un niveau de couverture suffisant dans les délais impartis, l'organisme peut se voir retirer son agrément d'assurance.

11. L'aspect pratique du fonctionnement d'un modèle de projection et donc du calcul du $SCR_i^{sous-module}$ sera précisé en troisième partie

12. Autorité de Contrôle Prudentiel et de Résolution

L'approche en formule standard étant commune à tous les assureurs, elle n'est donc, par définition, ajustée à aucun assureur. L'utilisation d'un modèle interne, et notamment dans le cadre du calcul du capital de solvabilité requis, résulte généralement en une réduction du coût de l'événement bicentenaire en raison d'une modélisation et d'un calibrage adapté aux risques de l'assureur.

Dans ce mémoire, notre attention sera portée sur l'optimisation des chocs de mortalité en modèle interne, et plus spécifiquement sur le risque de longévité. Par la suite et à des fins de simplification, le risque de mortalité fera référence à la fois à la surmortalité et à la sous-mortalité.

2.1 Définition

En assurance, et plus précisément sous Solvabilité 2, le risque de mortalité correspond au risque de surestimation des probabilités de survie des assurés exposés à la surmortalité. Un assuré est exposé à la surmortalité lorsque le raccourcissement de sa durée de vie impacte défavorablement l'assureur. Par exemple, pour des produits de prévoyance décès, si les assurés décèdent plus que prévu initialement, l'assureur se devra d'effectuer les prestations prévues au contrat, sans pour autant avoir collecté les primes nécessaires.

Quant au risque de longévité, il correspond au risque de sous-estimation des probabilités de survie des assurés exposés à la longévité. Un assuré est exposé à la longévité lorsque le rallongement de sa durée de vie impacte défavorablement l'assureur. Par exemple, pour des produits de retraite, si les assurés décèdent moins que prévu initialement, l'assureur se devra de verser davantage d'arrérages qu'il ne l'avait estimé.

Du point de vu de la modélisation, le risque de mortalité est décomposable en deux sous-risques :

- Le risque de niveau ou l'erreur d'estimation : Cela correspond au risque de mauvaise estimation des probabilités de décès initiales. Cette mauvaise estimation peut provenir de l'hétérogénéité entre la population servant de base pour estimer les probabilités de décès et la population assurée ou bien de la volatilité induite par l'échantillon utilisé potentiellement non représentatif de la population entière ;
- Le risque de tendance ou l'erreur de processus : Cela correspond au risque de mauvaise estimation des futures améliorations de mortalité. Cette mauvaise estimation peut provenir d'une évolution des facteurs régissant les tendances de mortalité, comme l'évolution des soins de santé, les améliorations médicales ou le changement de style de vie.

Dans le cadre de ce mémoire, on se focalisera sur ce dernier point, soit les estimations de tendances futures pour les scénarios de mortalité de la population française.

2.2 Évolution de la mortalité en France

Depuis des dizaines d'années voire des siècles, la mortalité en France a considérablement diminué. Au début du XX^e siècle, l'espérance de vie française à la naissance était d'environ 45 ans, tandis qu'aujourd'hui, elle s'élève à plus de 80 ans.

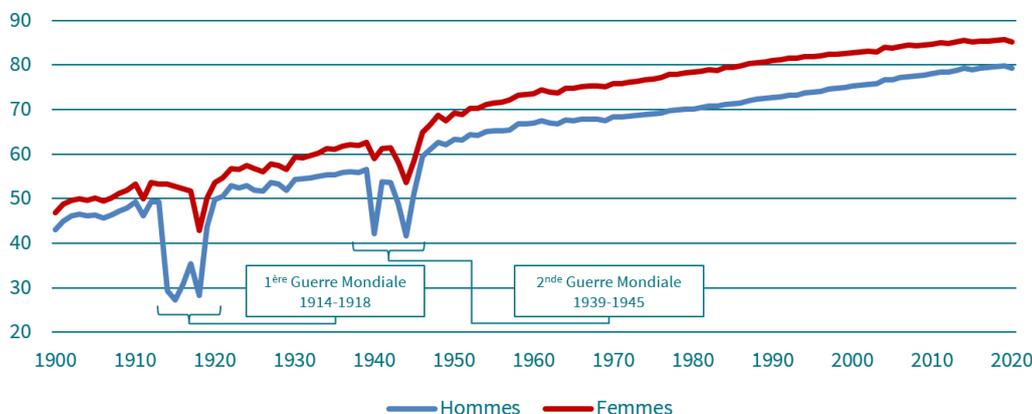


FIGURE 2.1: Évolution de l'espérance de vie à la naissance en France

Le graphique ci-dessus, établi sur les données de la Human Mortality Database [2], montre que l'espérance de vie croît presque linéairement depuis plus de 100 ans. Deux événements remettent en cause cette linéarité, les guerres mondiales de 1914-18 et 1939-45. En effet, un événement de cette ampleur a un impact fort sur la démographie car les jeunes soldats, souvent en âge de procréer, décèdent au combat ce qui engendre à la fois une baisse de l'espérance de vie et de la natalité. Par ailleurs, la dégradation des conditions de vie, avec des rationnements, des famines et de mauvaises conditions sanitaires contribuent à la diminution de l'espérance de vie. Ces événements étant atypiques, ils seront exclus de ce mémoire afin d'éviter d'introduire un biais dans nos résultats et dans notre analyse.

En dehors de ces deux périodes, l'amélioration de l'espérance de vie française est liée à divers facteurs :

- Les avancées médicales en matière de prévention, détection et traitement des maladies ;
- L'amélioration des conditions sanitaires, réduisant notamment la transmission de maladies infectieuses ;
- L'augmentation du niveau de vie, avec une meilleure alimentation, des conditions de travail plus sûres et des conditions de vie plus confortables ;
- Les évolutions sociales et comportementales en matière de santé avec par exemple une meilleure prise de conscience des risques du tabagisme pour la santé ou l'importance d'effectuer une activité physique régulière.

Entre 1950 et 2000, les gains d'espérance de vie sont surtout dus à la baisse de la mortalité infantile, cela s'explique par une amélioration des conditions sanitaires lors de la grossesse et de l'accouchement mais également par les soins aux nouveau-nés tels que la vaccination, l'administration d'antibiotiques en cas d'infections, etc... Néanmoins, la lutte pour la baisse de la mortalité infantile ne se fait pas uniquement sur le plan médical mais également sur plan sociétal et économique au travers d'une nutrition adaptée ou même grâce à la réduction de la fécondité.

Au cours des vingt dernières années, les gains d'espérance de vie sont désormais portées par la baisse de mortalité aux âges élevés. Les progrès de la prévention des maladies liées au vieillissement contribuent à cette amélioration avec notamment un accès à des programmes de dépistage, de vaccination ou de prévention aux maladies cardiovasculaires ou neurodégénératives. Les progrès de la médecine, avec une meilleure prise en charge des maladies chroniques et donc une meilleure qualité de vie ou la prise de conscience du vieillissement avec une population âgée encouragée à rester active physiquement et mentalement, sont également des facteurs contribuant à la baisse de mortalité.

Malgré la prise de conscience, le dépistage et le suivi des maladies chroniques, le vieillissement de l'organisme et l'affaiblissement du système immunitaire font qu'elles restent les premières causes de décès en France :

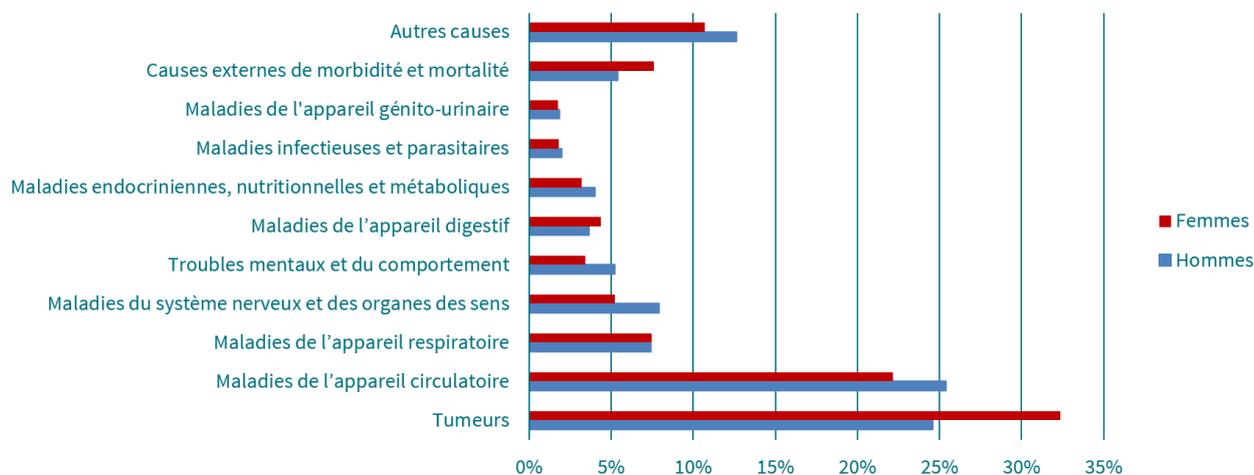


FIGURE 2.2: Causes de décès en 2017 en France

Comme constaté, depuis 1950, l'espérance de vie ne cesse d'augmenter ou *a minima* de rester stable, mais en 2020, la population Française a connu la plus forte baisse d'espérance de vie à la naissance des soixante-dix dernières années, cette baisse s'explique par un événement : la pandémie mondiale liée à la COVID-19.

2.3 COVID-19

Les coronavirus sont une famille de virus susceptibles d'être à l'origine d'un large éventail de maladies. Chez l'Homme, ils provoquent des infections respiratoires, allant d'un rhume banal à une infection pulmonaire sévère, responsable de détresse respiratoire.

La COVID-19 [3] fait partie de la famille des coronavirus et est due au virus SARS-CoV-2. La plupart des personnes infectées par ce virus présentent une maladie respiratoire d'intensité légère à modérée et se rétablissent sans avoir besoin d'un traitement particulier. Néanmoins, certaines tombent gravement malades et ont besoin de soins médicaux. Les personnes âgées et celles qui ont un problème médical sous-jacent, tel qu'une maladie cardiovasculaire, un diabète, une maladie respiratoire chronique ou un cancer, ont plus de chance de présenter une forme grave. Peu importe l'âge, tout individu peut être infecté par la COVID-19 et risquer de développer une forme grave de la maladie ou d'en décéder.

La particularité du virus est sa vitesse de propagation, en effet, il se transmet d'une personne infectée à d'autres par des gouttelettes respiratoires et des aérosols lorsqu'une personne infectée respire, tousse, éternue ou parle. Le virus peut également se propager au contact d'une personne, surface ou objet contaminé.

Le virus est apparu pour la première fois en décembre 2019 à Wuhan, puis, au cours des premiers mois de 2020, il s'est propagé rapidement en Chine puis dans d'autres pays pour finalement être déclaré comme étant une pandémie par l'OMS¹ en mars 2020.

En France, le premier cas de COVID-19 a été signalé le 24 janvier 2020. Le virus s'étant propagé dans tout le pays à grande vitesse, le gouvernement a mis en place un certain nombre de mesures afin de tenter de freiner sa diffusion, comme la mise en place de gestes barrière, la vaccination ou les confinements (liste des restrictions non exhaustive) :

1. Organisation Mondiale de la Santé

- Premier confinement :
 - 17 mars 2020 : Début du confinement, avec des restrictions de déplacement, la fermeture des écoles et des entreprises non-essentiels ;
 - 11 mai 2020 : Phase 1 du déconfinement, avec la réouverture des écoles et des entreprises non-essentiels sous certaines conditions ;
 - 2 juin 2020 : Phase 2 du déconfinement, avec la fin de la restriction des déplacements à plus de 100km de son domicile ;
 - 22 juin 2020 : Phase 3 du déconfinement, avec la réouverture des cinémas et la reprise des activités de sports collectifs.

- Deuxième confinement :
 - 29 octobre 2020 : Remise en place d'un confinement, avec le maintien des écoles ouvertes ;
 - 28 novembre 2020 : Phase 1 du déconfinement, avec la mise en place d'un couvre-feu de 21 à 6h ;
 - 15 décembre 2020 : Phase 2 du déconfinement, avec la fin des attestations confinement mais avec un couvre-feu maintenu de 20h à 6h ;
 - 16 janvier 2021 : Avancement du couvre-feu à 18h.

- Troisième confinement :
 - 3 avril 2021 : Remise en place d'un confinement, avec un couvre-feu de 19h à 6h ;
 - 3 mai 2021 : Phase 1 du déconfinement, avec la fin des attestations et des restrictions de déplacement en journée ;
 - 19 mai 2021 : Phase 2 du déconfinement, avec le couvre-feu repoussé à 21h et la réouverture des commerces, musées et cinémas avec des jauges limitées ;
 - 9 juin 2021 : Phase 3 du déconfinement, avec la réouverture des restaurants et cafés en intérieur et des salles de sport.

La COVID-19 a eu des conséquences importantes sur la santé publique, l'économie et la vie quotidienne des citoyens. Malgré les mesures établies par le gouvernement, sa progression n'a pu qu'être atténuée temporairement, comme le montre l'indicateur d'évolution du nombre de cas COVID-19 en France :



FIGURE 2.3: Nombre de cas COVID-19 en France

Le nombre de cas de COVID-19 ne doit pas être analysé isolément car la détection dépend de la quantité de tests effectués. En cas de grand nombre de tests réalisés, on aura davantage de chance de détecter des cas, ne signifiant pas nécessairement une situation plus grave. Inversement, si les tests

sont limités, le nombre de cas sera plus faible mais l'épidémie pas pour autant moins présente. Par ailleurs, le nombre de cas ne tient pas compte de la gravité de la maladie, on peut avoir des contaminés asymptomatiques comme hospitalisés, par conséquent, une hausse de cas n'implique pas toujours une hausse de la mortalité.

Cette étude visant à mesurer l'impact de la COVID-19 sur la mortalité prospective en France, il convient plutôt d'analyser le nombre de décès :

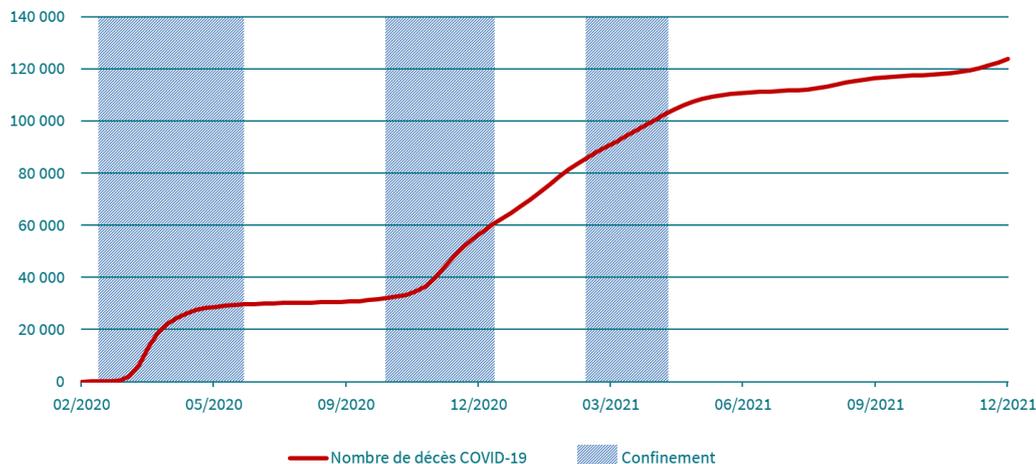


FIGURE 2.4: Nombre de décès COVID-19 en France

Le nombre de décès de COVID-19 doit également être analysé avec précaution, en effet, les décès sont souvent basés sur les certificats décès qui peuvent ne pas refléter la cause exacte du décès. Tout d'abord, il est probable que certaines personnes soient décédées de la COVID-19 sans que cela ait clairement été identifié ou signalé dans le certificat de décès. D'autre part, on peut également avoir des cas où des décès sont signalés comme étant dus à la COVID-19 alors que la cause réelle est une autre maladie.

Enfin, le nombre de décès ne tient pas compte des conséquences indirectes de la pandémie, telles que les retards de diagnostic ou de traitement d'autres maladies, les effets psychologiques sur la population ou les conséquences économiques et sociales de la crise.

La pandémie de COVID-19 aura permis de se questionner sur l'avenir en ce qui concerne les évolutions de mortalité. On peut alors se demander si les années COVID-19 reflètent l'avenir en termes de mortalité et plus spécifiquement, dans le cadre de la norme Solvabilité 2, on peut formuler la problématique suivante :

Quels sont les traitements envisageables sur les scénarios bicentennaires de longévité permettant l'intégration des données COVID-19 ?

Deuxième partie

Méthodes d'estimation de la mortalité
prospective

Dans cette seconde partie, on présentera l'exhaustivité du cadre théorique de ce mémoire. Pour cela, on définira au préalable les principales notions et hypothèses de travail essentielles à cette étude. Par la suite on introduira les modèles de durée employés, en précisant leur calibrage, en passant par les méthodes d'évaluation prospective utilisées, jusqu'à la définition des chocs bicentennaires. Finalement, on établira les critères de sélection pertinents permettant de juger de la qualité des estimations et de la robustesse des scénarios bicentennaires.

3.1 Notations et définitions

Dans le cadre de la théorie de survie [4] [5] [6] [7], on s'intéresse à la variable aléatoire T représentant la durée de vie d'un individu. Sachant que les individus étudiés ont généralement atteint un certain âge, on considérera plutôt la variable T_x comme étant la durée de vie résiduelle d'un individu conditionnellement au fait qu'il soit vivant à l'âge x , soit $T_x = [T - x | T > x]$.

On peut donc définir la probabilité qu'un individu d'âge x survive y années supplémentaires, avec $y > 0$:

$${}_y p_x = \mathbb{P}(T_x > y) = \mathbb{P}(T > x + y | T > x)$$

Symétriquement, on peut définir le quotient de mortalité, soit la probabilité qu'un individu d'âge x ne survive pas y années supplémentaires, avec $y > 0$:

$${}_y q_x = 1 - {}_y p_x = \mathbb{P}(T_x \leq y) = \mathbb{P}(T \leq x + y | T > x)$$

A noter que par simplification, lorsque $y = 1$, on utilisera les notations p_x et q_x .

En utilisant les propriétés des lois de probabilité, on obtient la relation suivante :

$${}_y p_x = \frac{p_{x+y}}{p_x}$$

Considérons désormais à l'intérieur d'un groupe homogène, à un instant pris comme origine, l'ensemble des individus d'âge x en nombre L_x . Supposons qu'ils décèdent, en nombre $D_x = L_x - L_{x+1}$, indépendamment les uns des autres c'est-à-dire que leur état de santé n'est pas affecté par des phénomènes contagieux. Dans ce cas on peut attacher à chaque élément du groupe une variable aléatoire $X_i(y)$ qu'on appellera indicateur de survie et qui prend à l'époque y la valeur 1 si l'individu est vivant, et la valeur 0 s'il est mort.

Les variables $X_i(y)$ sont en nombre égal à L_x et elles sont mutuellement indépendantes, on a :

$$\mathbb{E}(X_i(y)) = {}_y p_x$$

A l'époque y le nombre de survivants du groupe initial composé de L_x individus est :

$$L_{x+y} = X_1(y) + X_2(y) + \dots + X_{L_x}(y)$$

D'où :

$$\mathbb{E}(L_{x+y}) = \sum_{i=1}^{L_x} \mathbb{E}(X_i(y)) = L_x \times {}_y p_x$$

Soit $l_{x+y} = \mathbb{E}(L_{x+y})$ le nombre probable de vivants à l'âge $x + y$, lorsque y tend vers 0 on a :

$$l_x = \mathbb{E}(L_x) = L_x$$

D'où :

$${}_y p_x = \frac{l_{x+y}}{l_x}$$

Si ω est l'âge extrême de la vie humaine, on a :

$$l_{\omega+\epsilon} = l_x \times {}_{\omega-x+\epsilon} p_x = 0$$

L'ensemble des valeurs de l_y pour $x \leq y \leq \omega$ constitue une loi de survie. Elle est ainsi définie à une constante de proportionnalité $L_x = l_x$

Prenons un individu d'âge x et supposé vivant à l'époque y , soit à l'âge $x + y$, la probabilité qu'il décède dans l'intervalle de temps $[y, y + \Delta y]$ est la suivante :

$$\mathbb{P}(y < T_x < y + \Delta y | T_x > y) = \frac{\mathbb{P}(y < T_x < y + \Delta y)}{\mathbb{P}(T_x > y)}$$

Or :

$$\mathbb{P}(y < T_x < y + \Delta y) = {}_y p_x - {}_{y+\Delta y} p_x$$

Lorsque $\Delta y \rightarrow 0$, cette dernière expression tend vers 0, mais on peut l'écrire, pour Δy infiniment petit et si la fonction ${}_y p_x$ est dérivable par rapport à y , de la façon suivante :

$$\mathbb{P}(y < T_x < y + \Delta y) = -{}_y p'_x \times \Delta y$$

Dans ces conditions, on a donc :

$$\mathbb{P}(y < T_x < y + \Delta y | T_x > y) = -\frac{{}_y p'_x}{{}_y p_x} \times \Delta y$$

Ensuite, si on remarque que ${}_y p_x = \frac{l_{x+y}}{l_x}$, alors ${}_y p'_x = \frac{l'_{x+y}}{l_x}$, on a :

$$\lim_{\Delta y \rightarrow 0} \frac{\mathbb{P}(y < T_x < y + \Delta y | T_x > y)}{\Delta y} = \frac{l'_{x+y}}{l_{x+y}}$$

C'est une fonction μ_{x+y} de l'âge atteint $x + y$ que l'on appelle le taux instantané de mortalité ou la force de mortalité à l'âge $x + y$. Pour un individu d'âge z , on a donc :

$$\mu_z = -\frac{l'_z}{l_z} = -\frac{\partial}{\partial z} [\ln(l_z)]$$

Inversement, si on connaît la fonction μ_z , on aura par intégration entre x et $x + y$:

$$\ln(l_{x+y}) - \ln(l_x) = - \int_x^{x+y} \mu_z \times dz$$

Soit :

$$\ln({}_y p_x) = - \int_x^{x+y} \mu_z \times dz$$

D'où :

$${}_y p_x = \exp^{-\int_x^{x+y} \mu_z \times dz}$$

A partir de ces notations, on peut définir la durée de vie résiduelle d'un individu d'âge x , indicateur caractéristique des tables de mortalité. On note :

$$E_x = \int_0^{\omega-x} y \times \mathbb{P}(y < T_x < y + dy)$$

Et comme :

$$\mathbb{P}(y < T_x < y + dy) = -{}_y p'_x \times dy = -\frac{l'_{x+y}}{l_x} \times dy$$

On a alors :

$$E_x = -\frac{1}{l_x} \times \int_0^{\omega-x} y \times l'_{x+y} \times dy$$

On observe que :

$$d(y \times l_{x+y}) = (t \times l'_{x+y} + l_{x+y}) \times dy$$

Sachant que $l_\omega = 0$, on en déduit finalement les relations suivantes :

$$\begin{aligned} E_x &= -\frac{1}{l_x} \times \int_0^{\omega-x} d(y \times l_{x+y}) + \frac{1}{l_x} \times \int_0^{\omega-x} l_{x+y} \times dy \\ &= \frac{1}{l_x} \times \int_x^\omega l_z \times dz \end{aligned}$$

Ce mémoire portant sur l'évaluation de la mortalité prospective et donc son évolution au cours du temps, il est donc nécessaire de compléter les variables définies ci-dessus en intégrant la dimension temporelle, soit l'année calendaire :

Notation	Définition
$L_{x,t}$	Nombre d'individus d'âge x en année calendaire t
$D_{x,t}$	Nombre de décès à l'âge x en année calendaire t
${}_y p_{x,t}$	Pour l'année calendaire t , probabilité qu'un individu d'âge x survive y années supplémentaires
${}_y q_{x,t}$	Pour l'année calendaire t , probabilité qu'un individu d'âge x ne survive pas y années supplémentaires
$\mu_{x,t}$	Taux instantané de mortalité pour un individu d'âge x en année calendaire t
$E_{x,t}$	Espérance de vie résiduelle pour un individu d'âge x en année calendaire t

TABLE 3.1: Extension des notations à la dimension temporelle

3.2 Hypothèses de travail

Le taux de décès observé, soit $m_{x,t} = \frac{D_{x,t}}{L_{x,t}}$, et le quotient de mortalité $q_{x,t}$ sont généralement proches en valeur. Par une hypothèse simple, on peut formaliser cette relation de manière plus précise :

Pour tout x et t et, pour tout $0 \leq u, s < 1$, $\mu_{x+u,t+s} = \mu_{x,t}$, ainsi, on suppose que la force de mortalité reste constante au cours de chaque année d'âge entier et au cours de chaque année calendaire. Cela implique les relations suivantes :

$$\begin{aligned}m_{x,t} &= \mu_{x,t} \\q_{x,t} &= 1 - \exp^{-\mu_{x,t}} = 1 - \exp^{-m_{x,t}}\end{aligned}$$

La première relation est souvent utilisée dans l'analyse des données relatives aux taux de mortalité. La deuxième relation est utile dans l'analyse des modèles paramétriques de mortalité qui sont formulés en termes de $q_{x,t}$. L'hypothèse de constance de mortalité entre deux entiers n'est pas toujours vérifiée, mais la relation entre $m_{x,t}$ et $q_{x,t}$ est généralement considérée comme une approximation précise.

4.1 Lee-Carter

4.1.1 Définition

Le modèle de Lee-Carter (1992) [8] ou LC en abrégé, est un modèle permettant d'évaluer la mortalité prospective en se basant sur les tendances passées. Ce modèle est rapidement devenu un standard dans le domaine et reste une référence aujourd'hui. Le modèle donne une expression du taux instantané de mortalité de la forme suivante :

$$\ln \mu_{x,t} = \alpha_x + \beta_x \times \kappa_t + \epsilon_{x,t} \text{ avec } \epsilon_{x,t} \sim \mathcal{N}(0, \sigma^2) \text{ i.i.d.}$$

Le paramètre α_x s'interprète comme étant la valeur moyenne des $\ln \mu_{x,t}$ au cours du temps. Par ailleurs, on vérifie que :

$$\frac{\partial \ln \mu_{x,t}}{\partial t} = \beta_x \times \frac{\partial \kappa_t}{\partial t}$$

On en déduit que le coefficient β_x traduit la sensibilité de la mortalité instantanée à l'âge x par rapport à l'évolution générale κ_t , notamment car $\frac{\partial \ln \mu_{x,t}}{\partial \kappa_t} = \beta_x$. De par la non dépendance au temps du paramètre β_x , le modèle de LC suppose implicitement la constance du paramètre au cours du temps, ce qui peut apparaître comme étant une hypothèse forte.

Ensuite, le modèle de LC impose deux contraintes supplémentaires afin de le rendre identifiable ; pour toute constante c non nulle, le modèle doit être invariant aux transformations suivantes :

$$\begin{aligned} (\alpha_x, \beta_x, \kappa_t) &\rightarrow (\alpha_x, \frac{\beta_x}{c}, c \times \kappa_t) \\ (\alpha_x, \beta_x, \kappa_t) &\rightarrow (\alpha_x - c \times \beta_x, \beta_x, \kappa_t + c) \end{aligned}$$

Les contraintes retenues sont donc les suivantes :

$$\sum_x \beta_x = 1 \text{ et } \sum_t \kappa_t = 0$$

On doit résoudre le problème des moindres carrés suivant :

$$(\hat{\alpha}_x, \hat{\beta}_x, \hat{\kappa}_t) = \arg \min \sum_{x,t} (\ln \mu_{x,t} - \alpha_x - \beta_x \times \kappa_t)^2$$

Il s'agit donc d'un problème à $2 \times (x_M - x_m + 1) + (t_M - t_m + 1) - 2$ paramètres, avec x_m (resp. x_M) l'âge minimum (resp. maximum) de l'historique utilisé, t_m (resp. t_M) l'année la plus ancienne (resp. récente) de l'historique utilisé et -2 en raison des contraintes d'identifiabilité.

4.1.2 Estimation des paramètres

Tout d'abord, afin d'estimer α_x , si on remarque que :

$$\frac{\partial}{\partial \alpha_x} \sum_{y,t} (\ln \mu_{y,t} - \alpha_y - \beta_y \times \kappa_t)^2 = 2 \times (t_M - t_m + 1) \times \alpha_x - 2 \times \sum_t (\ln \mu_{x,t} - \beta_x \times \kappa_t)$$

Et en tenant compte de la contrainte $\sum_t \kappa_t = 0$, on obtient finalement :

$$\hat{\alpha}_x = \frac{1}{t_M - t_m + 1} \times \sum_t \ln \mu_{x,t}$$

Cela signifie que, pour chaque âge x , $\hat{\alpha}_x$ est la moyenne historique du logarithme des taux instantanés de mortalité.

Maintenant, pour l'estimation des paramètres restants, l'approche la plus couramment utilisée consiste à définir une matrice $Z = (z_{x,t})$ avec :

$$z_{x,t} = \ln \mu_{x,t} - \hat{\alpha}_x$$

On cherche deux vecteur $\hat{\beta}$ et $\hat{\kappa}$ tels que $Z \approx \hat{\beta} \times \hat{\kappa}$, pour cela, on utilise une méthode basée sur la décomposition en éléments propres de Z . On a $Z = \sum_{i \geq 1} \sqrt{\lambda_i} \times \nu_i \times u'_i$, avec :

$$\begin{aligned} \lambda_1 \geq \lambda_2 \geq \dots \geq 0 \text{ les valeurs propres de } Z' \times Z \\ u_i \text{ le vecteur propre normé de } Z' \times Z \text{ associé à } \lambda_i \\ \nu_i \text{ le vecteur propre normé de } Z \times Z' \text{ associé à } \lambda_i \end{aligned}$$

A noter que les valeurs propres sont identiques pour $Z' \times Z$ et $Z \times Z'$. De plus, si u_i est un vecteur propre de $Z' \times Z$ alors $Z \times u_i$ est un vecteur propre de $Z \times Z'$ associé à la même valeur propre. Ensuite, si on pose $\nu_i = \frac{1}{\sqrt{\lambda_i}} \times Z \times u_i$, on a alors $\sqrt{\lambda_i} \times \nu_i \times u'_i = Z \times u_i \times u'_i$, et donc, en sommant et en tenant compte de l'orthogonalité des vecteurs propres, on obtient une décomposition de Z :

$$\begin{aligned} \sum_{i \geq 1} \sqrt{\lambda_i} \times \nu_i \times u'_i &= \sum_{i \geq 1} Z \times u_i \times u'_i \\ &= Z \times \sum_{i \geq 1} u_i \times u'_i \\ &= Z \times \mathbb{1} \end{aligned}$$

Cela permet donc de proposer comme approximation $Z \approx \sqrt{\lambda_1} \times \nu_1 \times u'_1$ avec comme mesure de qualité d'approximation $\frac{\lambda_1}{\sum_{i \geq 1} \lambda_i}$. Après décomposition en éléments propres, les estimateurs retenus sont les suivants :

$$\begin{aligned} \hat{\beta} &= \frac{1}{\sum_{j \geq 1} \nu_{1,j}} \times \nu_1 \\ \hat{\kappa} &= \sqrt{\lambda_1} \times \left(\sum_{j \geq 1} \nu_{1,j} \right) \times u_1 \end{aligned}$$

A ce stade, on dispose d'une première estimation des paramètres du modèle. Néanmoins, pour améliorer la qualité de ses estimations, on va supposer que le nombre de décès $D_{x,t}$ suit une loi de Poisson de moyenne $L_{x,t} \times m_{x,t}$. Ensuite, posons θ comme étant l'ensemble des paramètres, la notation $m_{x,t}$ devient donc $m_{x,t|\theta}$ pour indiquer la dépendance entre ces paramètres. L'expression générale de la log-vraisemblance est alors la suivante (à une constante additive près) :

$$l_{\theta|D,L} = \sum_{x,t} D_{x,t} \times \log[L_{x,t} \times m_{x,t|\theta}] - L_{x,t} \times m_{x,t|\theta} - \log[D_{x,t}!^1]$$

Tout d'abord, à l'aide de l'expression de la log-vraisemblance et on reprenant l'hypothèse $m_{x,t} = \mu_{x,t}$, on va chercher à estimer les paramètres β et κ par maximum de vraisemblance, soit en trouvant la racine de la dérivée première. Cela revient à résoudre, pour tout x pour β et, pour tout t pour κ , les égalités suivantes :

$$\begin{aligned} \frac{\partial l_{\theta|D,L}}{\partial \beta_x} &= \sum_t \kappa_t \times [D_{x,t} - L_{x,t} \exp^{\alpha_x + \beta_x \times \kappa_t}] = 0 \\ \frac{\partial l_{\theta|D,L}}{\partial \kappa_t} &= \sum_x \beta_x \times [D_{x,t} - L_{x,t} \exp^{\alpha_x + \beta_x \times \kappa_t}] = 0 \end{aligned}$$

Posons alors les fonctions $F^{(1)}(b) = \sum_t \kappa_t \times [D_{x,t} - L_{x,t} \exp^{\alpha_x + b \times \kappa_t}]$ et $F^{(2)}(k) = \sum_x \beta_x \times [D_{x,t} - L_{x,t} \exp^{\alpha_x + \beta_x \times k}]$. La forme des fonctions F assure l'unicité de la racine si elle existe. Pour déterminer cette racine, on utilise un algorithme de type Newton-Raphson, en posant, pour $i \geq 1$, avec $b_0 = \hat{\beta}$ et $k_0 = \hat{\kappa}$ et en définissant un critère d'arrêt ϵ tel que l'évolution entre deux itérations successives soit inférieur à ϵ :

$$\begin{aligned} b_{i+1} &= b_i - \frac{F^{(1)}(b_i)}{F^{(1)'}(b_i)} \\ k_{i+1} &= k_i - \frac{F^{(2)}(k_i)}{F^{(2)'}(k_i)} \end{aligned}$$

On obtient ainsi deux nouveaux estimateurs $\hat{\beta}$ et $\hat{\kappa}$ qui doivent être ajustés pour vérifier les contraintes $\sum_x \beta_x = 1$ et $\sum_t \kappa_t = 0$. Les paramètres β_x^* et κ_t^* retenus pour le modèle sont donc les suivants :

$$\begin{aligned} \beta_x^* &= \frac{\hat{\beta}_x}{\sum_j \hat{\beta}_j} \\ \kappa_t^* &= \sum_j \hat{\beta}_j \times \left[\hat{\kappa}_t - \frac{\sum_i \hat{\kappa}_i}{t_M - t_m + 1} \right] \end{aligned}$$

Finalement, à l'aide des paramètres β_x^* et κ_t^* , on ajuste, par maximum de vraisemblance, l'estimateur $\hat{\alpha}_x$. On pose donc, pour tout x , α_x^* , tel que :

$$\alpha_x^* = \log \sum_t D_{x,t} - \log \left[\sum_t E_{x,t} \times \exp^{\beta_x^* \times \kappa_t^*} \right]$$

Le modèle de LC est désormais calibré en tenant compte des contraintes qui lui sont associées, cependant, il n'est pas encore utilisable en l'état. Des étapes supplémentaires sont nécessaires afin de traiter correctement la problématique de ce mémoire, celles-ci seront détaillées par la suite :

- Fermeture de tables : L'estimation des paramètres du modèle de LC nécessite que l'on dispose de suffisamment d'observations pour limiter l'instabilité des estimations. Pour les âges élevés, les observations sont souvent insuffisantes pour se baser sur les prédictions du modèle. De nombreuses méthodes permettent d'évaluer la mortalité aux âges élevés ;

- Projection de κ : L'intérêt du modèle de LC, et particulièrement dans le cadre de ce mémoire, est d'évaluer la mortalité prospective, pour cela, il est nécessaire de projeter le paramètre temporel κ . Tout comme pour la fermeture de tables, il existe de nombreuses méthodes de projection de κ .

4.2 Cairns-Blake-Dowd

4.2.1 Définition

Le modèle de Cairns-Blake-Dowd (2006) [9] ou CBD en abrégé, une des variantes principales du modèle de LC, permet, tout comme ce dernier, d'évaluer la mortalité prospective en se basant sur les tendances passées. Il existe de nombreuses versions du modèle de CBD, on s'intéresse ici à l'originelle, bien moins sophistiquée que les versions récentes mais suffisamment robuste pour produire des résultats pertinents. Le modèle donne une expression du quotient de mortalité de la forme suivante :

$$\text{logit}q_{x,t} = \ln \frac{q_{x,t}}{1 - q_{x,t}} = \kappa_t^{(1)} + \kappa_t^{(2)} \times (x - \bar{x}) + \epsilon_{x,t}$$

Avec :

$$\begin{aligned} \epsilon_{x,t} &\sim \mathcal{N}(0, \sigma^2) \text{ i.i.d.} \\ \bar{x} &= \frac{1}{x_M - x_m + 1} \times \sum_{i=x_m}^{x_M} x_i \end{aligned}$$

Le paramètre $\kappa_t^{(1)}$ indépendant des âges, traduit l'évolution temporelle moyenne du logit des quotients de mortalité et reflète ainsi le niveau moyen de la mortalité en fonction du temps. $\kappa_t^{(2)}$, le second paramètre temporel du modèle, correspond à la sensibilité de la mortalité en fonction de l'âge x . Historiquement, on constate que la mortalité baisse plus rapidement aux plus jeunes âges comparé aux âges élevés, bien que la tendance s'inverse depuis quelques années.

Contrairement au modèle de LC, le modèle de CBD ne nécessite pas d'imposer des contraintes d'identifiabilité. On doit donc résoudre le problème des moindres carrés :

$$(\widehat{\kappa_t^{(1)}}_t, \widehat{\kappa_t^{(2)}}_t) = \arg \min_{x,t} \sum (\text{logit}q_{x,t} - \kappa_t^{(1)} - \kappa_t^{(2)} \times (x - \bar{x}))^2$$

Il s'agit donc d'un système à $2 \times (t_M - t_m + 1)$ paramètres, avec t_m (resp. t_M) l'année la plus ancienne (resp. récente) de l'historique utilisé.

4.2.2 Estimation des paramètres

Afin d'initialiser les paramètres par moindres carrés, on cherche à trouver $\kappa_t^{(1)}$ et $\kappa_t^{(2)}$ tels que :

$$\begin{aligned} \frac{\partial}{\partial \kappa_t^{(1)}} \sum_{x,t} [\text{logit}q_{x,t} - \kappa_t^{(1)} - \kappa_t^{(2)} \times (x - \bar{x})]^2 &= 0 \\ \frac{\partial}{\partial \kappa_t^{(2)}} \sum_{x,t} [\text{logit}q_{x,t} - \kappa_t^{(1)} - \kappa_t^{(2)} \times (x - \bar{x})]^2 &= 0 \end{aligned}$$

Après quelques étapes, on obtient finalement pour tout t les égalités suivantes :

$$\widehat{\kappa_t^{(1)}} = \frac{\sum_x \text{logit} q_{x,t}}{x_M - x_m + 1}$$

$$\widehat{\kappa_t^{(2)}} = \frac{\sum_x (x - \bar{x}) \times [\text{logit} q_{x,t} - \kappa_t^{(1)}]}{\sum_x (x - \bar{x})^2}$$

On dispose ainsi d'une première estimation des paramètres, cependant, tout comme pour le modèle de LC, on souhaite améliorer la qualité des estimations en supposant que le nombre de décès $D_{x,t}$ suit une loi de Poisson de moyenne $L_{x,t} \times m_{x,t}$. On rappelle alors l'expression de la log-vraisemblance :

$$l_{\theta|D,L} = \sum_{x,t} D_{x,t} \times \log[L_{x,t} \times m_{x,t|\theta}] - L_{x,t} \times m_{x,t|\theta} - \log[D_{x,t}!]$$

On cherche alors à estimer les paramètres $\kappa_t^{(1)}$ et $\kappa_t^{(2)}$ par maximum de vraisemblance, pour cela, on cherche pour tout t , la racine des dérivées premières de la log-vraisemblance. Afin de simplifier ces dérivées, sous l'hypothèse de constance du taux instantané de mortalité entre deux années calendaires et deux âges entiers, on rappelle les égalités suivantes :

$$m_{x,t} = \mu_{x,t} = -\log(1 - q_{x,t}) = \log(1 + \exp^{\kappa_t^{(1)} + \kappa_t^{(2)} \times (x - \bar{x})})$$

On cherche alors :

$$\frac{\partial l_{\theta|D,L}}{\partial \kappa_t^{(1)}} = \sum_x \frac{D_{x,t}}{\log(1 + \exp^{\kappa_t^{(1)} + \kappa_t^{(2)} \times (x - \bar{x})}) \times (1 + \exp^{-\kappa_t^{(1)} - \kappa_t^{(2)} \times (x - \bar{x})})} - \frac{L_{x,t}}{(1 + \exp^{-\kappa_t^{(1)} - \kappa_t^{(2)} \times (x - \bar{x})})} = 0$$

$$\frac{\partial l_{\theta|D,L}}{\partial \kappa_t^{(2)}} = \sum_x \frac{D_{x,t} \times (x - \bar{x})}{\log(1 + \exp^{\kappa_t^{(1)} + \kappa_t^{(2)} \times (x - \bar{x})}) \times (1 + \exp^{-\kappa_t^{(1)} - \kappa_t^{(2)} \times (x - \bar{x})})} - \frac{L_{x,t} \times (x - \bar{x})}{(1 + \exp^{-\kappa_t^{(1)} - \kappa_t^{(2)} \times (x - \bar{x})})} = 0$$

Tout comme pour le modèle de LC, afin d'utiliser un algorithme de type Newton-Raphson, on pose au préalable les expressions suivantes :

$$F^{(1)}(a) = \sum_x \frac{D_{x,t}}{\log(1 + \exp^{a + \kappa_t^{(2)} \times (x - \bar{x})}) \times (1 + \exp^{-a - \kappa_t^{(2)} \times (x - \bar{x})})} - \frac{L_{x,t}}{(1 + \exp^{-a - \kappa_t^{(2)} \times (x - \bar{x})})}$$

$$F^{(2)}(b) = \sum_x \frac{D_{x,t} \times (x - \bar{x})}{\log(1 + \exp^{\kappa_t^{(1)} + b \times (x - \bar{x})}) \times (1 + \exp^{-\kappa_t^{(1)} - b \times (x - \bar{x})})} - \frac{L_{x,t} \times (x - \bar{x})}{(1 + \exp^{-\kappa_t^{(1)} - b \times (x - \bar{x})})}$$

Afin de déterminer la dérivée première de ces deux expressions, rappelons l'expression générale de la fonction f' , dérivée d'une fonction f :

$$f'(x) = \lim_{\Delta t \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}$$

On a donc, pour h suffisamment petit (par exemple $h = 10^{-6}$) :

$$F^{(1)'}(a) = \frac{F^{(1)}(a + h) - F^{(1)}(a)}{h}$$

$$F^{(2)'}(b) = \frac{F^{(2)}(b + h) - F^{(2)}(b)}{h}$$

Ainsi, pour $i \geq 1$, avec comme initialisation $a_0 = \widehat{\kappa}^{(1)}$ et $b_0 = \widehat{\kappa}^{(2)}$ et en définissant un critère d'arrêt ϵ tel que l'évolution entre deux itérations successives soit inférieur à ϵ :

$$a_{i+1} = a_i - \frac{F^{(1)}(a_i)}{F^{(1)'}(a_i)}$$

$$b_{i+1} = b_i - \frac{F^{(2)}(b_i)}{F^{(2)'}(b_i)}$$

On obtient alors deux nouveaux estimateurs $\widehat{\kappa}^{(1)}$ et $\widehat{\kappa}^{(2)}$, le modèle de CBD est donc calibré. Comme pour tout modèle, une fermeture de tables est nécessaire afin d'ajuster les estimations volatiles aux âges élevés. La projection des deux paramètres temporels sera également nécessaire afin d'évaluer la mortalité prospective et ainsi traiter la problématique de ce mémoire. Les méthodes utilisées, communes pour chacun des modèles étudiés, seront développées par la suite.

4.3 Börger-Fleischer-Kuksin alternatif

4.3.1 Définition

Le modèle de Börger-Fleischer-Kuksin alternatif ou BFK alternatif en abrégé, inspiré du modèle de BFK (2014) [10], est une variante du modèle de CBD, qui permet, tout comme ce dernier, d'évaluer la mortalité prospective en se basant sur les tendances passées. Ce modèle a pour avantage de tenir compte des évolutions de la mortalité pour différentes tranches d'âges. Le modèle donne une expression du quotient de mortalité de la forme suivante :

$$\text{logit}q_{x,t} = \ln \frac{q_{x,t}}{1 - q_{x,t}} = \alpha_x + \kappa_t^{(1)} + \kappa_t^{(2)} \times (x - \bar{x}) + \kappa_t^{(3)} \times (x - x_{old})^+ + \epsilon_{x,t}$$

Avec :

$$\epsilon_{x,t} \sim \mathcal{N}(0, \sigma^2) \text{ i.i.d.}$$

$$\bar{x} = \frac{1}{x_M - x_m + 1} \times \sum_{i=x_m}^{x_M} x_i$$

$$(x - x_{old})^+ = \max(x - x_{old}, 0)$$

Le paramètre α_x s'interprète comme étant la valeur moyenne des $\text{logit}q_{x,t}$ au cours du temps. Le paramètre $\kappa_t^{(1)}$ indépendant des âges, traduit l'évolution temporelle moyenne du logit des quotients de mortalité. Le second paramètre temporel du modèle, soit $\kappa_t^{(2)}$, correspond à la sensibilité de la mortalité en fonction de l'âge x . Enfin, le dernier paramètre temporel $\kappa_t^{(3)}$ tient compte des effets supplémentaires aux âges avancés, avec x_{old} l'âge à partir duquel $\kappa_t^{(3)}$ est défini.

Le modèle originel proposé par BFK intègre deux paramètres supplémentaires à l'expression de du quotient de mortalité, soit :

$$\text{logit}q_{x,t} = \ln \frac{q_{x,t}}{1 - q_{x,t}} = \alpha_x + \kappa_t^{(1)} + \kappa_t^{(2)} \times (x - \bar{x}) + \kappa_t^{(3)} \times (x - x_{old})^+ + \kappa_t^{(4)} \times (x_{young} - x)^+ + \gamma_{t-x} + \epsilon_{x,t}$$

Le paramètre $\kappa_t^{(4)}$, dans le même esprit que $\kappa_t^{(3)}$, a pour vocation de tenir compte des effets supplémentaires mais pour les âges inférieurs à x_{young} . On ne retient pas ce paramètre en raison de son non intérêt dans le cadre de cette étude, en effet, l'impact de la COVID-19 sur les personnes peu âgées est marginal comparativement à celui sur les plus âgées.

Le second paramètre γ_{t-x} , de cohorte, permet de modéliser les influences spécifiques à une cohorte en cherchant à capturer les caractéristiques communes des individus nés à la même période. De même, on retire ce paramètre étant donné que les modèles de LC et CBD n'intègrent pas ce type de paramètre, cela permet de ne pas créer de biais de modélisation et ainsi facilite l'inter-comparabilité des modèles étudiés.

Tout comme pour le modèle de CBD, le modèle de BFK alternatif ne nécessite pas d'imposer des contraintes d'identifiabilité. On doit donc résoudre le problème des moindres carrés :

$$(\widehat{\alpha}_x, \widehat{\kappa}_t^{(1)}, \widehat{\kappa}_t^{(2)}, \widehat{\kappa}_t^{(3)}) = \arg \min_{x,t} \sum (\text{logit}q_{x,t} - \alpha_x - \kappa_t^{(1)} - \kappa_t^{(2)} \times (x - \bar{x}) - \kappa_t^{(3)} \times (x - x_{old})^+)^2$$

Il s'agit donc d'un système à $(x_M - x_m + 1) + 3 \times (t_M - t_m + 1)$ paramètres, avec x_m (resp. x_M) l'âge minimum (resp. maximum) de l'historique utilisé et t_m (resp. t_M) l'année la plus ancienne (resp. récente) de l'historique utilisé. Étant donné qu'on souhaite fixer subjectivement le paramètre α_x et que $\kappa_t^{(3)}$ n'est pas défini pour $x \leq x_{old}$, on peut réduire le système à $2 \times (t_M - t_m + 1) + (t_M - x_{old})$ paramètres.

4.3.2 Estimation des paramètres

La première étape de l'estimation des paramètres consiste à définir α_x comme étant la moyenne historique des $\text{logit}q_{x,t}$, soit :

$$\widehat{\alpha}_x = \frac{1}{t_M - t_m + 1} \times \sum_t \text{logit}q_{x,t}$$

Ceci est principalement fait pour des raisons de temps de calcul et d'unicité des paramètres. Un affinement des paramètres sera effectué en fin de calibrage.

Pour estimer les paramètres $\kappa_t^{(i)}$ pour tout $t = t_m, \dots, t_M$, on effectuera une régression linéaire multiple (ou GLM²) telle que, pour tout $i = x_m, \dots, x_M$, $Y_i = a_0 + a_1 \times X_{1,i} + a_2 \times X_{2,i}$, avec :

$$\begin{aligned} Y_i &= \text{logit}q_{i,t} - \alpha_x \\ X_{1,i} &= i - \bar{x} \\ X_{2,i} &= (i - x_{old})^+ \\ (a_0, a_1, a_2) &= (\widehat{\kappa}_t^{(1)}, \widehat{\kappa}_t^{(2)}, \widehat{\kappa}_t^{(3)}) \end{aligned}$$

Finalement, pour tout $t = t_m, \dots, t_M$, le système à résoudre prend la forme suivante :

$$\begin{pmatrix} \text{logit}q_{x_m,t} - \alpha_{x_m} \\ \dots \\ \text{logit}q_{x_i,t} - \alpha_{x_i} \\ \dots \\ \text{logit}q_{x_M,t} - \alpha_{x_M} \end{pmatrix} = M \times \begin{pmatrix} \kappa_t^{(1)} \\ \kappa_t^{(2)} \\ \kappa_t^{(3)} \end{pmatrix}$$

Avec :

$$M = \begin{pmatrix} 1 & x_m - \bar{x} & (x_m - x_{old})^+ \\ \dots & \dots & \dots \\ 1 & x_i - \bar{x} & (x_i - x_{old})^+ \\ \dots & \dots & \dots \\ 1 & x_M - \bar{x} & (x_M - x_{old})^+ \end{pmatrix}$$

Pour résoudre ce système, on utilisera la méthode d'itération pondérée des moindres carrés (IWLS) ou algorithme de Fisher Scoring. L'algorithme ne sera pas détaillé dans ce mémoire et on renvoie vers Dobson (1990) [11] pour plus de détails.

Le modèle est désormais calibré, néanmoins, afin de faciliter l'interprétation et la comparabilité entre les différents modèles étudiés, on doit procéder à des ajustements sur les paramètres obtenus. Tout d'abord, définissons ψ_1 comme étant la pente de la régression linéaire de α_x , les ajustements sont alors les suivants :

$$\begin{aligned}\kappa_t^{(2)} &\leftarrow \kappa_t^{(2)} + \psi_1 \\ \alpha_x &\leftarrow \alpha_x + \psi_1 \times (x - \bar{x}) \\ \psi_2 &= \bar{\alpha} \\ \alpha_x &\leftarrow \alpha_x - \psi_2 \\ \kappa_t^{(1)} &\leftarrow \kappa_t^{(1)} + \psi_2\end{aligned}$$

A noter que ces ajustements n'impactent pas les $\text{logit}q_{x,t}$ calibrés initialement, ils permettent d'obtenir une moyenne nulle pour le paramètre α_x et que, ainsi, les évolutions de taux de mortalité soient expliquées par les paramètres $\kappa_t^{(i)}$. Le paramètre $\kappa_t^{(1)}$ décrit donc le niveau de mortalité général, $\kappa_t^{(2)}$ la pente du logit des quotients de mortalité et α_x la structure de mortalité par âge.

Enfin, comme pour tout modèle, une fermeture de tables est nécessaire afin d'ajuster les estimations volatiles aux âges élevés. La projection des trois paramètres temporels sera également nécessaire afin d'évaluer la mortalité prospective et ainsi traité la problématique de ce mémoire. Les méthodes utilisées, communes pour chacun des modèles étudiés, seront développées par la suite.

4.4 Projection de la tendance

L'évaluation de la mortalité prospective nécessite de projeter les paramètres temporels, soit κ_t pour le modèle de LC, $\kappa_t^{(1)}$ et $\kappa_t^{(2)}$ pour celui de CBD et $\kappa_t^{(1)}$, $\kappa_t^{(2)}$ et $\kappa_t^{(3)}$ pour le modèle de BFK alternatif. La projection des paramètres temporels se fait à la fois de manière déterministe et stochastique, la différence entre les deux réside dans la prise en compte d'un aléa pour la stochastique, calibré sur les résidus historiques. On va donc détailler, pour chaque paramètre temporel, la méthode de projection déterministe et stochastique retenue.

4.4.1 Projection déterministe

Projection de κ_t (LC) et $\kappa_t^{(1)}$ (CBD et BFK alternatif) : Pour chacun des modèles étudiés, ce paramètre représente le niveau de mortalité général dans le temps, soit la composante d'intérêt pour les scénarios de tendance. Ainsi, une méthode de projection spécifique est utilisée, l'approche retenue, inspirée de celle proposée par BFK (2014), consiste à projeter le paramètre en effectuant une régression linéaire pondérée soit pour tout $t_{proj} \geq t_M$:

$$\kappa_{t_{proj}}^{(i)} = \kappa_{t_M}^{(i)} + (t_{proj} - t_M) \times \text{Slo}_h^{(i)}(t_M) + \text{Int}_h^{(i)}(t_M)$$

Avec :

$\text{Slo}_h^{(i)}(t_M)$: La pente (ou *slope*) de la régression linéaire pondérée de $\kappa^{(i)}$, de paramètre de pondération h et avec t_M la date d'observation la plus récente utilisée pour la régression ;

$\text{Int}_h^{(i)}(t_M)$: L'intercept de la régression linéaire pondérée de $\kappa^{(i)}$, de paramètre de pondération h et avec t_M la date d'observation la plus récente utilisée pour la régression ;

h : Le paramètre de pondération utilisé dans la fonction de pondération $\omega_t = \frac{1}{(1+1/h)^t}$.

En d'autres termes, le paramètre $\kappa^{(i)}$ est projeté linéairement compte tenu de l'historique de t_m à t_M en sur-pondérant les observations récentes. La pondération se fait au travers du paramètre h : au plus ce dernier est grand, au plus la régression tiendra compte des observations historiques et inversement, s'il est faible, alors les observations récentes auront un fort impact sur la régression. Le paramétrage de h sera détaillée dans la troisième partie de ce mémoire.

Projection de $\kappa_t^{(2)}$ (CBD et BFK alternatif) et $\kappa_t^{(3)}$ (BFK alternatif) : Le reste des paramètres temporels ne tiennent pas compte de la tendance générale et sont ainsi supposés insensibles aux scénarios de tendance. On décide donc de retenir une approche de projection simplifiée consistant à projeter l'écart moyen historique, soit pour tout $t_{proj} \geq t_M$:

$$\kappa_{t_{proj}}^{(i)} = \kappa_{t_M}^{(i)} + (t_{proj} - t_M) \times \frac{\sum_{k=t_m}^{t_M-1} \kappa_{k+1}^{(i)} - \kappa_k^{(i)}}{(t_M - 1) - t_m + 1}$$

4.4.2 Projection stochastique

Dans cette partie, on ne s'intéressera qu'à la projection de κ_t (LC) et $\kappa_t^{(1)}$ (CBD et BFK alternatif). En effet, les autres paramètres temporels étant supposés insensibles aux scénarios de tendance, l'aléa sera considéré nul et les projections déterministes et stochastiques seront alors identiques. Pour la projection stochastique, on réutilisera l'approche déterministe en intégrant un aléa tel que pour $t_{proj} > t_M$:

$$\kappa_{t_{proj}}^{(i)} = \kappa_{t_M}^{(i)} + (t_{proj} - t_M) \times \text{Slo}_h^{(i)}(t_M) + \text{Int}_h^{(i)}(t_M) + \epsilon_t \times \sigma_{Global,h}$$

Où :

$$\epsilon_t \sim \mathcal{N}(0, 1) \text{ i.i.d.}$$

$$\sigma_{Global,h} = \sigma_{Additionnel} + \sigma_{Historique,h}$$

Avec :

$\sigma_{Additionnel}$: Un supplément de volatilité facultatif destiné à garantir une volatilité raisonnablement prudente, notamment dans le cadre de la norme Solvabilité 2

$\sigma_{Historique,h} = \sqrt{\sum_{i=t_{m+1}}^{t_M-1} W_{i,h} \times (e_{i,h} - \mu_h)^2}$: La volatilité historique pondérée de paramètre de pondération h

Où :

- $\mu_h = \sum_{i=t_{m+1}}^{t_M-1} W_{i,h} \times e_{i,h}$: La moyenne pondérée des erreurs de régression, de paramètre de pondération h
- $e_{i,h}$: L'erreur d'estimation de la régression linéaire pondérée, soit l'écart entre κ_{i+1} et son estimation par régression linéaire pondérée, de paramètre de pondération h , basée sur les observations de t_m à i
- $W_{i,h} = \frac{\omega_{i,h}}{\sum_{i=t_{m+1}}^{t_M-1} \omega_{i,h}}$: Le vecteur de pondération de paramètre de pondération h
- $\omega_{i,h} = (1 + \frac{1}{h})^i$

Pour notre projection stochastique et dans le cadre de la norme Solvabilité 2, l'intégration de l'aléa ne sera fait qu'en première année de projection. En effet, la norme prévoit d'utiliser des chocs bicentennaires à horizon 1 an dans le cadre du calcul du capital de solvabilité requis. Ainsi, à compter de la seconde année, on repasse à une projection déterministe qui tiendra compte de la nouvelle observation stochastique telle que, pour tout $t_{proj} > t_{M+1}$:

$$\kappa_{t_{proj}}^{(i)} = \kappa_{t_{M+1}}^{(i)} + (t_{proj} - t_{M+1}) \times \text{Slo}_h^{(i)}(t_{M+1}) + \text{Int}_h^{(i)}(t_{M+1})$$

En pratique, le paramètre h évoluera en fonction de l'étape de projection : un paramètre spécifique, que l'on spécifiera et justifiera dans la troisième partie, sera utilisé pour la projection déterministe, stochastique et également pour le calcul de $\sigma_{Historique, h}$.

4.5 Fermeture de tables de mortalité

Tel que mentionné précédemment, à partir d'un certain âge, l'estimation des taux de mortalité se complexifie dans la mesure où les effectifs disponibles sont de plus en plus réduits. Cela conduit à des taux de mortalité estimés très volatils et peu représentatifs.

Pour remédier à ces problèmes, l'usage est d'étudier séparément la mortalité aux grands âges, pour cela, de nombreuses méthodes de fermeture de tables sont disponibles. La méthode retenue dans ce mémoire est celle de Denuit-Goderniaux (2005) [12]. Cette dernière consiste à ajuster un modèle log-quadratique de la forme suivante :

$$\ln(q_x) = a + bx + cx^2 + \epsilon_x \text{ avec } \epsilon_x \sim \mathcal{N}(0, \sigma^2) \text{ i.i.d.}$$

Avec les contraintes suivantes :

$$\begin{aligned} q_{\text{Age Limite}} &= 1 \\ q'_{\text{Age Limite}} &= 0 \end{aligned}$$

La première contrainte impose la fermeture de tables car on suppose que, pour tout t , il existe un âge limite à partir duquel tous les individus décèdent. La seconde contrainte impose une concavité, soit des taux de mortalité croissants avec un ralentissement aux âges élevés. Ces contraintes conduisent à l'égalité suivante :

$$a + bx + cx^2 = c \times (\text{Age Limite}^2 - 2x \times \text{Age Limite} + x^2)$$

Afin de déterminer l'âge de raccord à partir duquel il faut remplacer les quotients bruts par les valeurs ajustées, on effectue plusieurs régression en faisant varier l'âge de raccord x_0 dans un intervalle d'âges $[x_{inf}, x_{supp}]$ puis en sélectionnant l'âge raccord x_0 qui maximise le coefficient de détermination R^2 . Un lissage autour de l'âge raccord x_0 est également nécessaire afin d'éviter une éventuelle cassure au voisinage de ce dernier, pour cela, on procède à une simple moyenne géométrique pour les âges $x = x_0 - n, \dots, x_0 + n$.

Il existe de nombreuses autres méthodes de fermeture de tables, celles-ci ne seront pas détaillées dans ce mémoire, néanmoins, pour plus de détails sur ces différentes méthodes et leurs comparaisons, on renvoie vers Quashie-Denuit (2005) [13].

4.6 Chocs bicentennaires

Dès lors que les tables de mortalité sont projetées pour les quantiles souhaités et fermées à l'aide de la méthode retenue, il reste à définir le niveau de choc bicentenaire à appliquer dans le cadre des modules Mortalité/Longévité du calcul du SCR en modèle interne. L'utilisation immédiate des tables résultantes peut générer un risque de base dans la mesure où ces tables se basent la population des données de calibrage tandis que la population à choquer peut être différente. Par exemple, la sensibilité au risque de tendance ne sera pas systématiquement la même, qu'on soit sur un portefeuille de retraités ou d'emprunteurs.

Pour palier cette problématique, on introduit les *Reduction Rates* (RR) [15], tels que, pour tout $t = t_M, \dots, t_{proj}$ et tout $x = x_m, \dots, x_M$:

$$RR_{x,t}^{Longevity} = \min\left(\frac{q_{x,t}^{Longevity}}{q_{x,t_M}}, \frac{q_{x,t-1}^{Longevity}}{q_{x,t_M}}\right) \text{ dans le cas de scénarios de longévité}$$

$$RR_{x,t}^{Mortality} = \frac{q_{x,t}^{Mortality}}{q_{x,t_M}} \text{ dans le cas de scénarios de mortalité}$$

$$RR_{x,t}^{Central} = \frac{q_{x,t}^{Central}}{q_{x,t_M}} \text{ dans le cas du scénarios central}$$

Cette approche a pour avantage de pouvoir générer des tables bidimensionnelles d'améliorations bicentennaires à appliquer sur l'hypothèse de mortalité *Best Estimate* (BE). Elle est notamment pertinente pour les compagnies exerçant dans plusieurs secteurs d'activité dans un même pays, la compagnie peut alors utiliser les données de mortalité nationales pour le calibrage des modèles puis éventuellement ajuster les niveaux de chocs en ajoutant plus ou moins de prudence aux scénarios obtenus au travers du paramètre $\sigma_{Additionnel}$ présenté auparavant. Dans ce cas, au plus la tendance d'amélioration de la population nationale diffère de la population assurée, au plus la volatilité additionnelle doit être importante.

Finalement, quelle que soit la table de mortalité BE utilisée, la table choquée sera, pour tout $t = t_M, \dots, t_{proj}$ et tout $x = x_m, \dots, x_M$ de la forme suivante :

$$q_{x,t}^{Longevity} = q_{x,t_M}^{Central} \times \min(RR_{x,t}^{Longevity}, RR_{x,t}^{BE}) \text{ dans le cas de scénarios de longévité}$$

$$q_{x,t}^{Mortality} = q_{x,t_M}^{Central} \times \max(RR_{x,t}^{Mortality}, RR_{x,t}^{BE}) \text{ dans le cas de scénarios de mortalité}$$

Avec :

$$RR_{x,t}^{BE} = \frac{q_{x,t}^{BE}}{q_{x,t_M}^{BE}} \text{ les améliorations de mortalité de la table BE}$$

On introduit ici un *floor* (resp. *cap*) sur les scénarios de longévité (resp. mortalité) afin que les améliorations de mortalité soient supérieures (resp. inférieures) ou égales à celles de la table BE. Aussi, on supposera que les améliorations bicentennaires pour les âges non compris dans l'intervalle $[x_m, x_M]$ seront égales à celles de la borne la plus proche de l'intervalle, soit :

$$RR_{x,t}^{Bicentenaire} = RR_{x_m,t}^{Bicentenaire} \text{ pour tout } x < x_m \text{ et tout } t = t_M, \dots, t_{proj}$$

$$RR_{x,t}^{Bicentenaire} = RR_{x_M,t}^{Bicentenaire} \text{ pour tout } x > x_M \text{ et tout } t = t_M, \dots, t_{proj}$$

On a donc vu les modèles utilisés, leur calibrage ainsi que la façon de les projeter dans le temps sous différents scénarios. On a également introduit le concept de fermeture de tables de mortalité et finalement la traduction de ces tables en chocs à appliquer aux tables BE. Néanmoins, avant d'explorer les résultats obtenus, il est essentiel de comprendre les critères de validation de ces modèles afin d'identifier lesquels permettent la robustesse et la fiabilité des projections. Ces critères feront l'objet du cinquième chapitre et seront étudiés sous deux axes : Les critères statistiques et de prudence.

La sélection de modèle est une étape importante dans l'évaluation de la mortalité prospective. En effet, il existe de nombreux autres modèles que ceux présentés dans ce mémoire, et, en fonction de la population étudiée, ces derniers ne donneront pas les mêmes trajectoires de mortalité. Il est donc nécessaire d'introduire des critères de sélection, ces derniers peuvent être qualitatifs ou quantitatifs et permettent de faciliter la comparaison des modèles étudiés afin de définir le plus approprié à la population sous risque.

5.1 Critères statistiques

5.1.1 BIC : Critère d'information bayésien

Le critère d'information bayésien ou *Bayes Information Criterion* (BIC) est une mesure permettant de comparer différents modèles statistiques. Le BIC prend en compte la qualité du modèle ajusté ainsi que la complexité du modèle. Il est couramment utilisé dans le contexte de la sélection de modèle, où plusieurs modèles sont comparés pour déterminer celui qui offre le meilleur compromis entre l'ajustement aux données et la simplicité. La formule du BIC est la suivante :

$$BIC = -2 \times \ln(L) + k \times \ln(n)$$

Avec :

$\ln(L)$: La log-vraisemblance du modèle

k : Le nombre de paramètres du modèle

n : La taille de l'échantillon

Le BIC pénalise les modèles les plus complexes en ajoutant un terme proportionnel au nombre de paramètres. Ainsi, le BIC favorise les modèles qui offrent un bon ajustement tout en étant relativement simples, évitant ainsi le sur-ajustement. En général, lors de la comparaison de modèles, on cherche à minimiser la valeur absolue du BIC car cela indique un meilleur compromis entre l'ajustement aux données et la complexité du modèle.

5.1.2 SMR : Ratio de mortalité standardisé

Le ratio de mortalité standardisé ou *Standardized Mortality Ratio* (SMR) est un indicateur s'exprimant en pourcentage permettant de mesurer l'écart entre la mortalité observée et la mortalité théorique, soit :

$$SMR_{x,t} = \frac{q_{x,t}}{\widehat{q}_{x,t}}$$

Un SMR de 100% indique que le nombre observé de décès est égal au nombre attendu. Un SMR inférieur à 100% suggère une mortalité inférieure à celle attendue, tandis qu'un SMR supérieur à 100% indique une mortalité plus élevée que prévu. Cet indicateur est essentiel pour la gestion des risques assurantiels, notamment en provisionnement pour un portefeuille de rentiers, il est préférable d'utiliser une table de mortalité pour laquelle le SMR est supérieur à 100% afin de s'assurer que les provisions établies permettent de couvrir les engagements en cas de dérive de mortalité défavorable.

L'analyse des SMR fournit des informations importantes sur la qualité du modèle, un modèle bien ajusté devrait produire des SMR qui ne montrent pas de schémas systématiques, ce qui signifie que les erreurs de prédiction sont bien aléatoires. Si des schémas sont détectés dans les SMR, cela peut indiquer des aspects du phénomène de mortalité qui ne sont pas pris en compte par le modèle, suggérant ainsi la nécessité d'améliorer le modèle ou d'inclure d'autres variables explicatives.

5.2 Critères de prudence

5.2.1 Amélioration d'espérance de vie

Ce premier critère de prudence vise à s'assurer que les améliorations d'espérance de vie projetées dans nos scénarios bicentennaires sont pertinentes au regard des améliorations historiques. Pour cela, on définit le taux d'amélioration d'espérance de vie tel que, pour tout $x = x_m, \dots, x_M$ et $t = t_m + 1, \dots, t_M, \dots, t_{proj}$:

$$\text{Imp}_{x,t} = 1 - \frac{q_{x,t}}{q_{x,t-1}}$$

On définit ensuite l'amélioration moyenne, sur une profondeur de 10 ans (hypothèse de duration moyenne des passifs étudiés), soit pour tout $x = x_m, \dots, x_M$ et tout $t = t_m + 10, \dots, t_M, \dots, t_{proj}$:

$$\text{Avg Imp}_{x,t} = \frac{1}{t - (t - 9) + 1} \times \sum_{i=t-9}^t \text{Imp}_{x,i}$$

Finalement, on définit Imp Ratio, correspondant au ratio entre la moyenne des Avg Imp $_{x,t_M+10}$ du scénario bicentenaire pondérés par ω (avec par exemple ω comme étant à la distribution d'âge du portefeuille étudié) et la moyenne des Avg Imp $_{x,t_M}$ historiques pondérés par le même ω . Ce ratio devra par la suite être supérieur au seuil $\epsilon^{(1)}$, soit :

$$\text{Imp Ratio} = \frac{\sum_{i=x_m}^{x_M} \omega_i \times |\text{Avg Imp}_{i,t_M+10}^{\text{Bicentenaire}}|}{\sum_{i=x_m}^{x_M} \omega_i \times |\text{Avg Imp}_{i,t_M}^{\text{Historique}}|} > \epsilon^{(1)}$$

Un second critère sur les améliorations d'espérance de vie est instauré, proche du critère précédent, celui-ci a pour vocation de comparer les améliorations historiques maximales aux améliorations projetées dans nos scénarios bicentennaires, sur une profondeur de 10 ans. Ainsi, dans le même esprit que pour le premier critère, on introduit Imp Ratio max comme étant le ratio entre la moyenne des

Avg Imp $_{x,t_M+10}$ du scénario bicentenaire pondérés par ω et la moyenne des améliorations maximales historiques pondérées par le même ω . Ce ratio devra par la suite être compris entre $\epsilon_{inf}^{(2)}$ et $\epsilon_{sup}^{(2)}$, soit :

$$\epsilon_{inf}^{(2)} < \text{Imp Ratio}^{max} = \frac{\sum_{i=x_m}^{x_M} \omega_i \times |\text{Avg Imp}_{i,t_M+10}^{Bicentenaire}|}{\sum_{i=x_m}^{x_M} \omega_i \times \left| \max_{t_M-9 \leq i \leq t_M} \text{Imp}_{x,i}^{Historique} \right|} < \epsilon_{sup}^{(2)}$$

Avec les seuils déterminés de la manière suivante :

$\epsilon^{(1)}$: Le seuil, basé sur le quantile 90% de la distribution empirique des Imp Ratio historiques, est égal à 170% pour les hommes et femmes (non différenciation par genre en raison de la proximité des seuils obtenus) ;

$\epsilon_{inf}^{(2)}$: Le seuil, basé sur le quantile 90% de la distribution empirique des Imp Ratio max historiques, est égal à 20% pour les hommes et femmes (non différenciation par genre en raison de la proximité des seuils obtenus) ;

$\epsilon_{sup}^{(2)}$: Le seuil est fixé arbitrairement à 100%, en effet, on mesure le risque de tendance et ne souhaitons donc pas que nos améliorations moyennes soient supérieures à celles maximales observées car ces dernières sont ponctuelles et ne représentent ainsi pas une tendance.

5.2.2 Backtesting

Le *backtesting* a pour but de vérifier la qualité de prédiction du modèle quant à la population étudiée, il consiste à comparer les taux de mortalité projetés aux taux de mortalité historiques. Pour cela, on segmente notre intervalle temporel en 2, le premier intervalle $[t_m, \dots, \lfloor t_m + (t_M - t_m) \times \frac{2}{3} \rfloor + 1]$ permet le calibrage du modèle à *backtester*. Le modèle calibré est projeté sur le second intervalle $[\lfloor t_m + (t_M - t_m) \times \frac{2}{3} \rfloor + 2, \dots, t_M]$, les taux de mortalités ainsi projetés sont comparés aux taux de mortalité historiques.

On cherche à s'assurer que le SMR de la mortalité BE projetée à 10 ans (hypothèse de duration moyenne des passifs étudiés) est supérieur (resp. inférieur) au ratio entre la mortalité projetée pour un scénario bicentenaire de longévité (resp. mortalité) et la mortalité BE projetée à 10 ans, avec une éventuelle pondération ω par âge, soit :

$$\frac{\sum_{i=x_m}^{x_M} \frac{q_{i,N}^{Longevity}}{q_{i,N}^{Central}} \times \omega_i}{\sum_{i=x_m}^{x_M} \omega_i} < \frac{\sum_{i=x_m}^{x_M} \frac{q_{i,N}^{Historique}}{q_{i,N}^{Central}} \times \omega_i}{\sum_{i=x_m}^{x_M} \omega_i} < \frac{\sum_{i=x_m}^{x_M} \frac{q_{i,N}^{Mortality}}{q_{i,N}^{Central}} \times \omega_i}{\sum_{i=x_m}^{x_M} \omega_i}$$

Avec :

$$N = \lfloor t_m + (t_M - t_m) \times \frac{2}{3} \rfloor + 1 + 10$$

Par ailleurs, on pourra comparer qualitativement la vision BE du modèle calibré aux observations historiques, par exemple, on aura tendance à favoriser un modèle sous-estimant les probabilités de décès pour un portefeuille de rentiers et inversement pour un portefeuille d'emprunteurs. Il convient de souligner que les critères de prudence servent davantage à écarter les modèles inappropriés qu'à valider les modèles appropriés. Autrement dit, un modèle qui valide les critères de prudence n'est pas nécessairement considéré comme adapté, mais on aura tendance à rejeter ceux qui ne répondent pas à ces critères.

1. Notons $\lfloor x \rfloor$ comme étant la partie entière de x

Troisième partie

Étude de l'impact COVID-19 sur l'estimation de la mortalité prospective

Dans cette troisième et dernière partie, on exploitera le cadre théorique établi auparavant, en conjonction avec les données et hypothèses de modélisation qui seront précisées, pour constater et analyser les répercussions de la COVID-19 sur l'estimation de la mortalité prospective. Face à cette problématique émergente, on proposera des solutions alternatives visant à atténuer voire neutraliser les effets de la pandémie sur l'évaluation de la mortalité future.

Analyse de l'impact de la COVID-19 sur l'estimation de la mortalité prospective

6.1 Données de l'étude

Dans le cadre de ce mémoire, on utilisera les données disponibles sur la *Human Mortality Database* (HMD), cette base de données résulte d'une collaboration internationale entre des experts en démographie, en statistiques et en santé publique. Son objectif principal est de centraliser et d'harmoniser les données sur la mortalité issues de diverses sources et périodes, facilitant ainsi l'accès à des informations fiables et comparables.

Les données disponibles dans la HMD englobent une variété de mesures, notamment des taux de mortalité par tranche d'âges, l'espérance de vie et d'autres indicateurs démographiques. La base couvre un large éventail de pays, bien que la disponibilité des données puisse varier en fonction des régions et des périodes, son accessibilité en ligne offre une opportunités aux chercheurs qui souhaitent explorer et intégrer ces données dans leurs études sur les modèles de mortalité.

Ces données fournissent une base solide et requise pour ajuster les paramètres des modèles étudiés, au travers des données d'expositions et de décès, et permettant ainsi d'évaluer de manière approfondie l'impact de la pandémie sur l'évaluation de la mortalité prospective.

Au cours de cette section, on présentera certaines statistiques descriptives des données utilisées, portant sur la période 1979-2020 et pour la tranche d'âges 30-100 ans. Ces statistiques fourniront un aperçu approfondi des dynamiques de mortalité au fil du temps et permettront une compréhension plus fine de l'impact de la COVID-19 sur nos modèles :

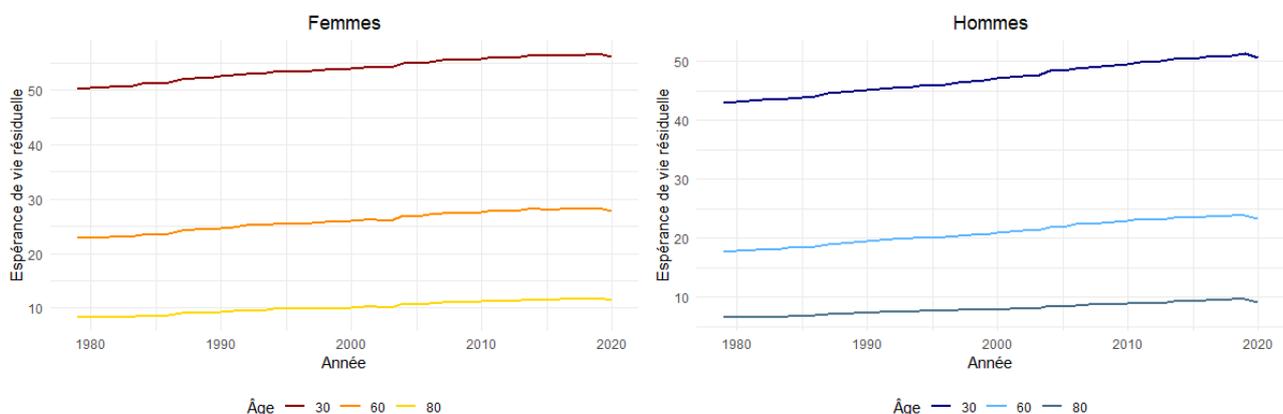


FIGURE 6.1: Évolution historique de l'espérance de vie à 30, 60 et 80 ans

Comme on l'évoquait en première partie, l'espérance de vie française croît presque linéairement depuis de nombreuses années, cependant cette croissance n'est pas identique entre les femmes et les hommes. En effet, on observe un rattrapage pour les hommes moins âgés se traduisant par une croissance plus accentuée pour 30 et 60 ans avec respectivement 1,7 et 0,5 ans de gain d'espérance de vie résiduelle supplémentaire pour les hommes comparé aux femmes sur la période 1979-2020. En revanche, le phénomène inverse est notable pour l'espérance de vie résiduelle à 80 ans, avec gain d'espérance de vie supplémentaire pour les femmes de 0,6 sur la même période.

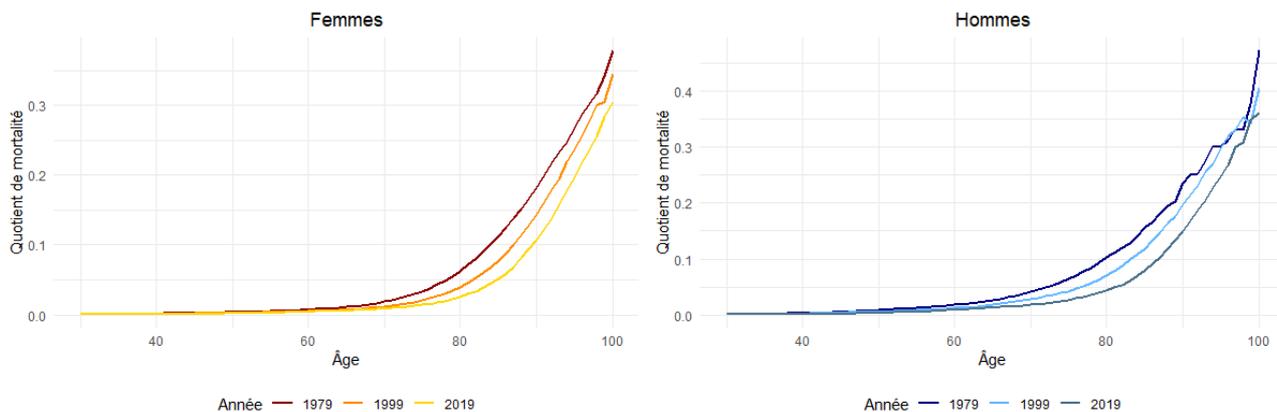


FIGURE 6.2: Distribution des taux de mortalité par âge en 1979, 1999 et 2019

Ces graphiques permettent d'illustrer les améliorations d'espérances de vie, on observe notamment une translation de la distribution des taux de mortalité par âge. Cette dernière est observable dans une moindre mesure pour les âges peu élevés en raison du faible niveau des taux de mortalité. Toutefois, pour les âges plus élevés, le niveau de translation peut être contesté en raison du nombre limité d'observations, une statistique moyenne sur une profondeur d'historique plus grande, supérieure à 3 points, serait alors plus adaptée.

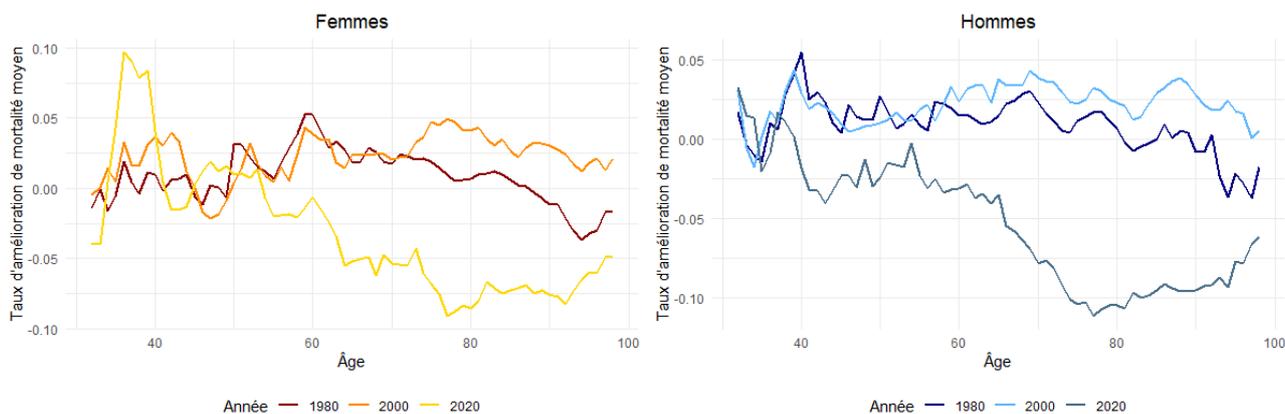


FIGURE 6.3: Moyenne sur 5 ans glissants des taux d'amélioration de mortalité par âge en 1980, 2000 et 2020

La moyenne des améliorations de mortalité, glissante sur 5 ans afin de modérer les variations ponctuelles, permet de mesurer et confirmer l'impact de la COVID-19, principalement pour les âges plus élevés. On constate que ces améliorations par âge, soit le rapport entre le quotient de mortalité de l'année rapporté à celui de l'année précédente, est globalement supérieur à 0% pour les années 1980 et 2000. Cependant, on remarque pour l'année 2020 une détérioration chez les femmes et les hommes à partir de 60 ans, plus prononcée pour ces derniers. Cette détérioration est en grande partie attribuable à la COVID-19 et ce mémoire aura pour vocation de proposer des solutions permettant l'atténuation voire le retraitement de ces effets.

6.2 Paramètres de modélisation

Dans cette section, on détaillera l'exhaustivité des paramètres utilisés dans notre modélisation pour la calibration des modèles, la projection de tendance déterministe et stochastique, la fermeture de tables et la production des chocs bicentennaires :

- **Profondeur d'historique de calibrage : 40 ans.** La profondeur d'historique est pertinente comparativement à ce que l'on souhaite capter, on veille à ce que l'historique soit suffisamment profond afin de ne pas percevoir des phénomènes ponctuels, tout en évitant d'aller trop loin pour ne pas inclure des événements exceptionnels tels que la Seconde Guerre mondiale. Il convient de noter que les variations modérées dans les profondeurs d'historiques ont un impact relativement faible ;
- **Tranche d'âges de calibrage : 30-100 ans.** La sélection de la tranche d'âges est alignée avec le contexte de l'étude. Dans le cas d'une population assurée, il est judicieux de fixer une borne inférieure relativement élevée étant donné que les personnes plus jeunes sont généralement moins assurées. Quant à la borne supérieure, celle-ci sera complétée par fermeture de tables, il est donc préférable de ne pas retenir un âge trop élevé car cela pourrait dégrader la qualité des ajustements en raison de la volatilité des observations à ces âges. Il convient de noter que les variations modérées dans les tranches d'âges ont un impact relativement faible ;
- **Seuil x_{old} du modèle de BFK alternatif : 60 ans.** Dans le cadre de cette étude, le paramètre temporel utilisant x_{old} a pour vocation de capter l'effet COVID-19 sur les personnes les plus âgées. L'approche pour définir son niveau est de projeter de façon déterministe les modèles n'utilisant pas ce paramètre (soit LC et CBD) sur les années impactées par la COVID-19 en les calibrant sur les données non impactées. Ainsi, on est en mesure d'identifier l'âge à partir duquel les SMR dérivent significativement ;
- **Paramètre de pondération, h de $\sigma_{Historique,h}$: 10 ou 30.** La paramètre $\sigma_{Historique,h}$ détermine le niveau de diffusion des scénarios stochastiques. Par mesure de prudence, on opte pour la volatilité maximale entre une volatilité moyen-terme, associée à une pondération à $h = 10$, et une volatilité long-terme, associée à une pondération à $h = 30$. La volatilité court-terme, avec par exemple $h = 3$, n'est pas retenue afin d'éviter la sur-pondération d'événements ponctuels ne reflétant pas une nouvelle tendance et n'étant donc pas destinés à être intégrés ;
- **Paramètre de pondération, h de Slo_h et Int_h pour la projection déterministe : 10.** Pour la projection de la tendance déterministe ou avant événement bicentenaire, on décide de retenir un paramètre de pondération moyen-terme, cela a pour effet de tenir compte des observations récentes sans pour autant sur-pondérer les éventuelles nouvelles tendances ponctuelles, mais également de tenir compte observations plus anciennes sans pour autant biaiser les paramètres de régression par des tendances passées résiduelles ;
- **Paramètre de pondération, h de Slo_h et Int_h après événement bicentenaire : 3.** A compter de l'événement bicentenaire, on décide d'adopter un paramètre de pondération court-terme afin d'assurer une influence significative des tendances futures par cet événement. Les scénarios résultants seront ainsi suffisamment prudents pour répondre aux exigences de Solvabilité 2 et pour valider nos divers critères ;
- **Paramètre $\sigma_{Additionnel}$: 0%.** Dans le cadre de cette étude, on décide de ne pas utiliser ce paramètre à des fins de simplifications et d'interprétabilité des résultats. En pratique, ce paramètre a diverses utilités et doit être évalué en cohérence avec les paramètres de pondération h . Il permet notamment de refléter l'écart d'évolution entre la population assurée, celle que l'on cherche à choquer, et la population nationale, celle qui sert de base de calibrage, afin que les scénarios obtenus soient suffisamment prudents, notamment dans le cadre de Solvabilité 2 ou de sorte à ce que les critères de prudence soient respectés ;
- **Nombre de scénarios stochastiques : 5 000.** Le nombre de scénarios stochastiques est évalué de sorte à offrir le meilleur compromis entre stabilité des résultats et réduction de temps de simulation. On observe que la volatilité des résultats devient relativement stable à compter de 5 000 scénarios, on décide alors de retenir cette valeur ;

- **Durée de projection de la mortalité prospective : 60 ans.** Les scénarios de mortalité prospective sont évalués sur 60 ans afin de s'aligner avec l'horizon de projection de la majorité des modèles ALM¹ ;
- **Paramètres de fermeture de tables, Age Limite : 130 ans.** L'âge limite que l'on retient est identique à celui proposé par Denuit-Goderniaux (2005), en effet, bien qu'à ce stade rien ne montre que la durée de vie humaine admette un maximum, il semble raisonnable de retenir un âge limite à 130 ans ;
- **Paramètres de fermeture de tables, $[x_{inf}, x_{sup}]$: 60-85 ans.** L'âge raccord à partir duquel on applique la fermeture de table est identifié sur un intervalle d'âges suffisamment large afin d'optimiser l'âge raccord retenu ;
- **Paramètres de fermeture de tables, n : 5 ans.** Le lissage est effectué sur 5 ans autour de l'âge raccord afin de lisser l'éventuelle cassure au voisinage de ce dernier.

Dans le cadre de cette étude, on sera amenés à ajuster certains de ces paramètres, ces ajustements seront explicités, justifiés et quantifiés dans la suite de ce mémoire. Ensuite, concernant le développement des approches présentées en deuxième partie, celles-ci seront implémentées à l'aide du logiciel *RStudio*.

6.3 Sélection de modèle

Une fois que tous les paramètres de la modélisation sont définis, on est prêts à appliquer les critères de sélection et de validation des modèles sur les résultats obtenus. A ce stade, on supposera que nos données ne sont pas influencées par la COVID-19, on limitera donc notre analyse à la période de 1979-2019. Certains critères de sélection seront réexaminés lorsque cela sera pertinent.

Il est important de noter qu'on observe historiquement une corrélation positive entre les améliorations de l'espérance de vie des femmes et celles des hommes, en particulier pour les tranches d'âges élevées. Cela pousse, dans la mesure du possible, à privilégier l'utilisation d'un modèle unique pour les deux genres :

Tranche d'âges	Corrélation des améliorations
30-50 ans	1%
51-65 ans	16%
66-80 ans	59%
81-100 ans	84%

TABLE 6.1: Corrélation historique des améliorations d'espérance de vie entre femmes et hommes par tranche d'âges

6.3.1 BIC : Critère d'information bayésien

Pour rappel, le critère d'information bayésien, que l'on cherche à minimiser, permet de comparer les modèles entre eux et favorise les modèles offrant le meilleur compromis entre ajustement aux données et simplicité. Les résultats obtenus sur les modèles étudiés sont les suivants :

	Femmes			Hommes		
	LC	CBD	BFK alternatif	LC	CBD	BFK alternatif
BIC	36 577	84 629	76 098	45 447	49 706	65 096
k	181	82	122	181	82	122
n	2 991					

TABLE 6.2: BIC pour les femmes par genre

1. *Asset Liability Management*

Les résultats montrent que le modèle de LC offre le meilleur compromis entre qualité des ajustements et simplicité pour les deux genres. En revanche, il est plus complexe d'établir un classement entre le modèle de CBD et BFK alternatif. Pour les hommes, le modèle de CBD offre un meilleur BIC que le BFK alternatif, néanmoins, l'inverse est notable, d'un moindre écart, pour les femmes. En termes de BIC, on peut donc considérer le modèle de LC comme étant le meilleur et ceux de CBD et BFK alternatif, en fonction de la pondération entre les femmes et les hommes, comme étant les moins performants.

6.3.2 SMR : Ratio de mortalité standardisé

Pour rappel, le ratio de mortalité standardisé permet de mesurer la qualité des ajustements du modèle au travers du niveau du SMR, en effet, un modèle parfaitement ajusté donnerait un SMR égal à 100% pour tous les âges et années. En pratique, cette égalité est difficilement atteignable, on cherchera alors à s'approcher au plus de 100% et sera privilégié un modèle donnant des SMR supérieurs (resp. inférieurs) à 100% pour un périmètre exposé à la longévité (resp. mortalité) afin que les estimations soient au minimum prudentes.

De plus, l'étude graphique des SMR centrés permet de confirmer ou rejeter l'existence de schémas systématiques. Ces schémas peuvent mettre en lumière des aspects de la mortalité non pris en compte dans la modélisation actuelle et suggèrent ainsi l'amélioration du modèle ou l'inclusion d'autres variables explicatives.

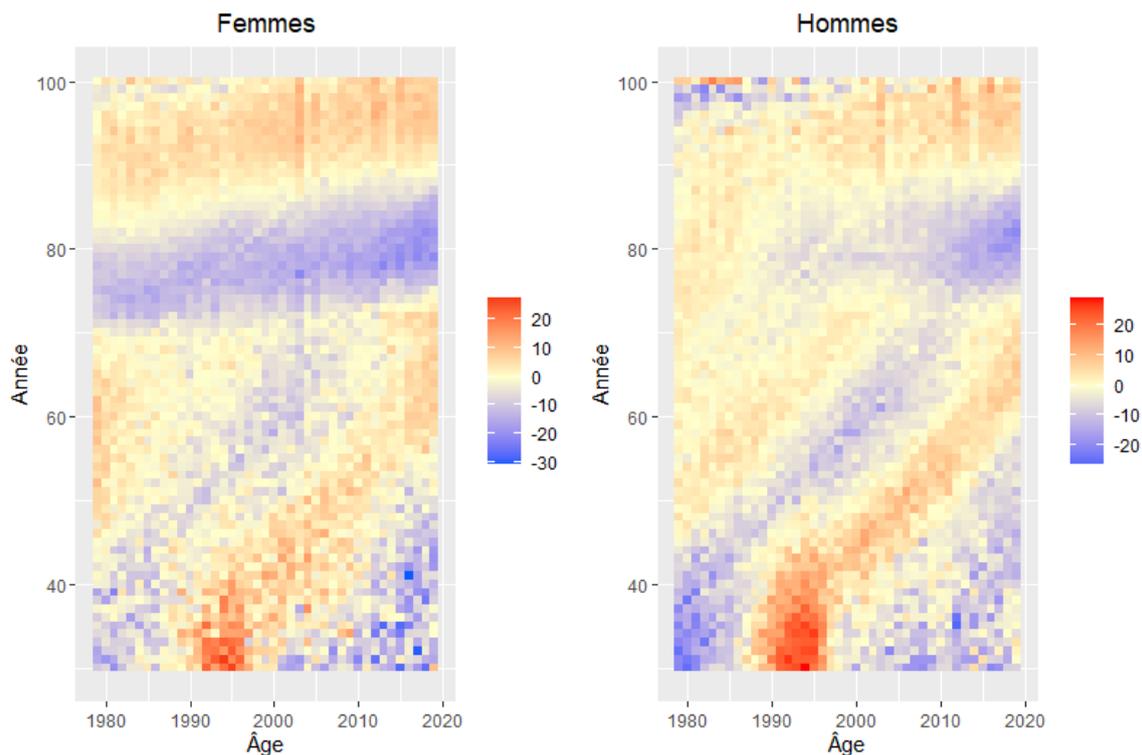


FIGURE 6.4: *Heatmap* des $SMR_{x,t}$ centrés pour le modèle de LC

Pour le modèle de LC, les SMR centrés donnent des écarts modérés avec des oscillations entre 20% et -20% pour les femmes et les hommes avec un écart moyen à respectivement -2% et -1%. En ce qui concerne les schémas systématiques, on constate deux effets :

- Les cohortes : Plus accentué pour les hommes, on observe deux diagonales pour les générations avoisinant 1960 et 1980, cela s'explique par une légère évolution dans la tendance de mortalité dans cet intervalle qui n'est pas capté par le modèle ;
- La fermeture de tables : Plus accentué pour les femmes, on remarque que la mortalité estimée par fermeture de tables aux âges proches de la jointure est sur-estimée. La méthode de fermeture

de tables utilisée offre une croissance de la mortalité modérée pour les âges moins élevés tandis qu'elle est plus importante en réalité, expliquant l'effet la sur-estimation à ces âges.

En dehors de ces deux effets, on note que le niveau global des SMR centrés est relativement bon, avec une légère tendance à la sous-estimation de la mortalité et donc préférable pour les périmètres exposés à la longévité. Le modèle reste néanmoins applicable aux périmètres exposés à la mortalité en raison de la faible matérialité de cette tendance.

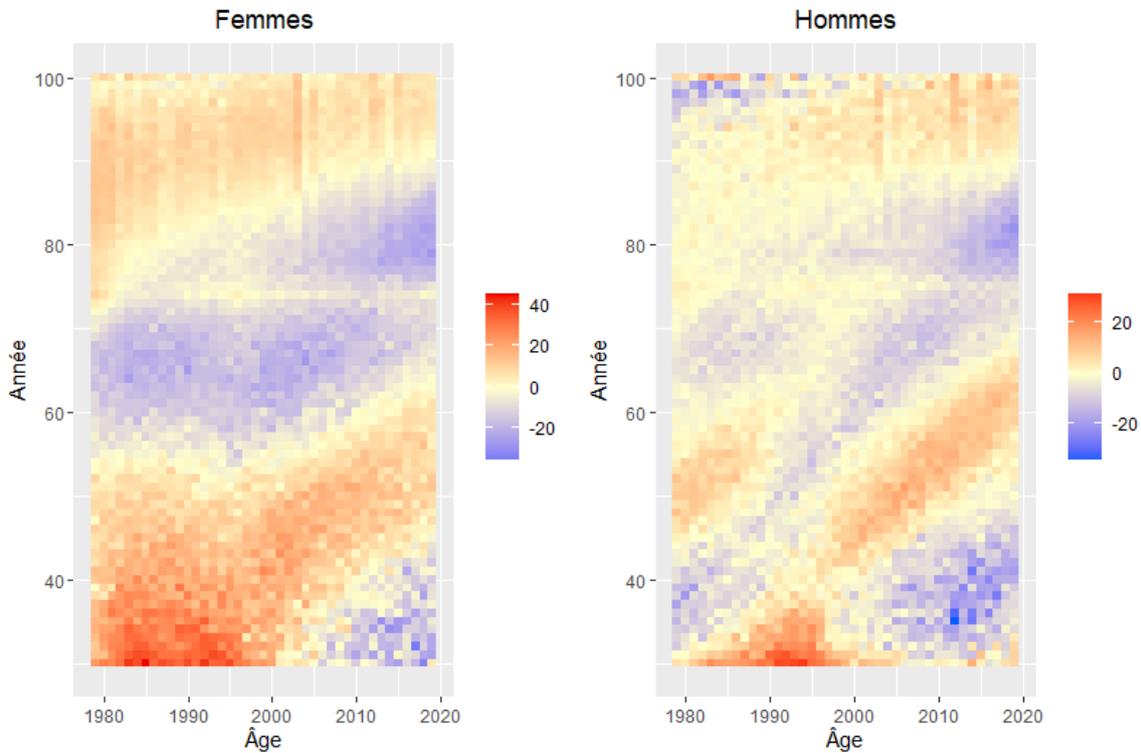


FIGURE 6.5: *Heatmap* des $SMR_{x,t}$ centrés pour le modèle de CBD

Pour le modèle de CBD, les SMR centrés donnent des écarts non négligeables avec des oscillations entre 40% et -20% pour les femmes et 20% et -20% pour les hommes avec un écart moyen à respectivement 1% et -1%. En ce qui concerne les schémas systématiques, pour les hommes, on ne constate que l'effet sur les cohortes avec un effet négligeable sur la fermeture de tables. En revanche, pour les femmes, on constate les deux effets, cependant, ils peuvent être considérés comme étant de second ordre au vu de la mauvaise qualité des estimations.

On constate que le modèle donne des estimations globalement peu fiables pour les femmes où l'on observe une sous-estimation de la mortalité pour les tranches d'âges 30-50 et 90-100 et l'inverse pour la tranche d'âges 50-90. En revanche, pour les hommes ce dernier reste cohérent avec des effets de cohortes et fermeture de tables relativement peu significatifs.

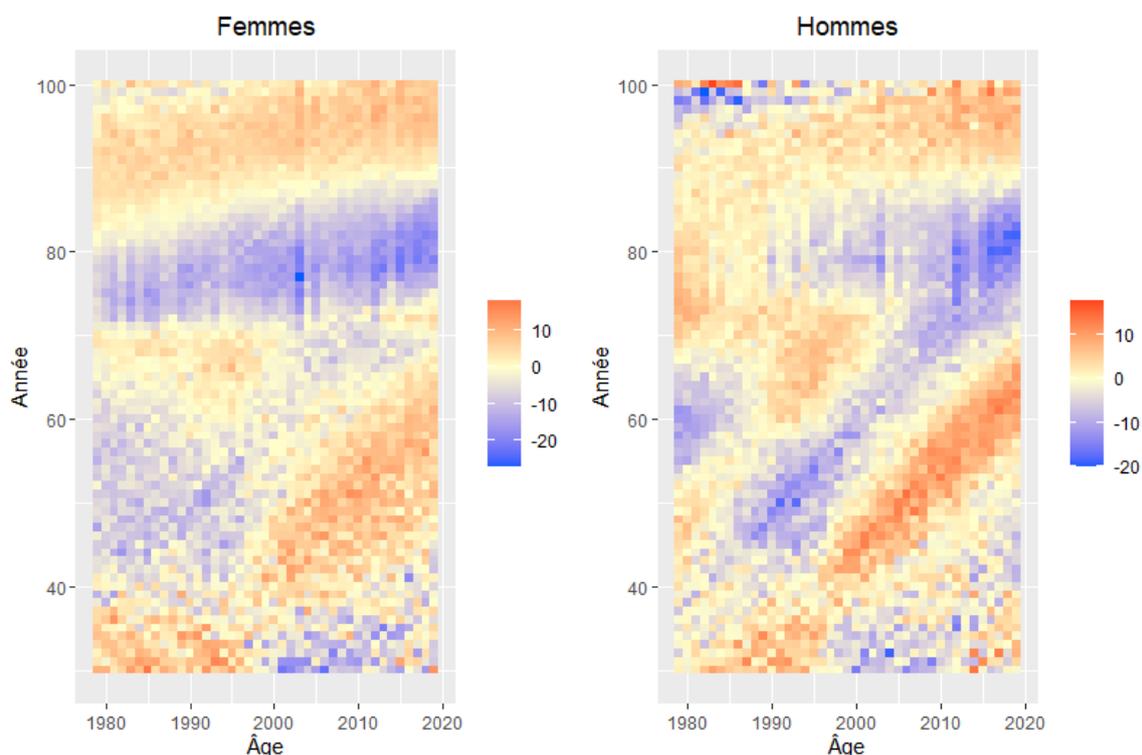


FIGURE 6.6: *Heatmap* des $SMR_{x,t}$ centrés pour le modèle de BFK alternatif

Pour le modèle de BFK alternatif, les SMR centrés donnent des écarts similaires à ceux du modèle de LC avec des oscillations entre 10% et -20% pour les femmes et les hommes avec un écart moyen à respectivement -1% et 0%. En ce qui concerne les schémas systématiques, les effets de cohorte et fermeture de tables sont observés, avec l'effet cohorte plus accentué pour les hommes et celui de fermeture de tables pour les femmes.

Au-delà des effets mentionnés ci-dessus, le niveau global des SMR centrés est relativement bon et ainsi applicable à des périmètres exposés à la fois à la longévité et mortalité.

En synthèse, on constate au travers de ce critère que les modèles de LC et BFK alternatif donnent des résultats de qualité, cependant, ce n'est pas le cas pour le modèle de CBD, cette faiblesse sera à prendre en compte lors de comparaison des résultats de l'étude. Les éventuels effets systématiques mentionnés devraient avoir des impacts non significatifs sur la problématique traitée :

- Les modèles de mortalité employés ne comprennent pas de paramètres de cohorte. Comme on le constate, l'absence de ces paramètres limite la capacité du modèle à saisir pleinement les effets de cohorte. Ainsi, l'intégration de ces paramètres dans nos modèles pourrait permettre une analyse plus approfondie des tendances spécifiques à certaines générations, ceci représentant une piste d'amélioration à l'étude ;
- La méthode de fermeture de tables adoptée s'avère être une approche robuste pour les âges élevés. Néanmoins, il est important de noter certaines lacunes au niveau de la jointure, suggérant que des ajustements pourraient être apportés afin de renforcer la cohérence de la modélisation, pour cela, d'autres approches de fermetures de tables pourraient être explorées.

6.3.3 Amélioration d'espérance de vie

Pour rappel, ce critère de prudence a pour vocation de s'assurer, à l'aide du Imp Ratio, que les améliorations bicentennaires projetées sont pertinentes au regard des améliorations historiques, mais également que les scénarios sont suffisamment prudents sans pour autant être trop agressifs au travers du Imp Ratio^{max}.

	Femmes	Hommes
LC	237%	197%
CBD	180%	194%
BFK alternatif	207%	202%
$\epsilon^{(1)}$	170%	

TABLE 6.3: Imp Ratio par modèle

Les trois modèles étudiés ont un Imp Ratio au-dessus du seuil $\epsilon^{(1)}$ et valident donc le premier critère de prudence, cela signifie que les améliorations bicentennaires projetées sont supérieures au quantile 90% des Imp Ratio historiques. On constate que le modèle de LC est globalement plus prudent que les modèles de CBD et BFK alternatif, ainsi, sous Solvabilité 2 et notamment dans le cadre de l'optimisation du Capital de Solvabilité Requis, on aura tendance à choisir un modèle offrant les scénarios bicentennaires les moins agressifs. A ce titre et en se basant uniquement sur ce critère, on aura tendance à favoriser le modèle de CBD.

	Femmes	Hommes
LC	24%	28%
CBD	19%	28%
BFK alternatif	21%	29%
$\epsilon_{inf}^{(2)}$	20%	
$\epsilon_{sup}^{(2)}$	100%	

TABLE 6.4: Imp Ratio^{max} par modèle

Les modèles de LC et BFK alternatif ont un Imp Ratio^{max} compris entre les seuils $\epsilon_{inf}^{(2)}$ et $\epsilon_{sup}^{(2)}$ et valident donc le second critère de prudence, cela signifie que les améliorations bicentennaires projetées sont supérieures au quantile 90% des Imp Ratio^{max} historiques et inférieures aux améliorations maximales historiques. En revanche, le modèle de CBD est en-dessous du seuil $\epsilon_{inf}^{(2)}$ pour les femmes, indiquant une potentielle sous-prudence des scénarios bicentennaires pour ce modèle.

Finalement, seuls les modèles de LC et BFK alternatif répondent avec succès aux critères d'amélioration d'espérance de vie. On aura tendance à privilégier le modèle de BFK alternatif étant donné qu'il est le moins agressif tout en satisfaisant le critère de prudence.

6.3.4 Backtesting

Pour rappel, le *Backtesting* permet de vérifier la qualité de prédiction du modèle en s'assurant que l'écart entre la mortalité historique et la mortalité centrale projetée est compris entre les ratios de mortalités bicentennaires projetées (longévité et mortalité) rapportés à la mortalité centrale projetée. En d'autres termes, on considère que le modèle n'est pas adapté si l'erreur de prédiction est supérieure aux scénarios bicentennaires :

	SMR _{Longevity/Central}	SMR _{Historique/Central}	SMR _{Mortality/Central}
LC	96%	102%	104%
CBD	96%	109%	104%
BFK alternatif	91%	95%	111%

TABLE 6.5: *Backtesting* pour les femmes par modèle

Pour les femmes, les SMR historiques sont compris entre les SMR de longévité et mortalité bicentennaires à l'exception du modèle de CBD, pour lequel le SMR de mortalité bicentenaire est inférieur à celui historique. Cela confirme le constat effectué précédemment sur la *heatmap* des SMR centrés montrant un mauvais ajustement du modèle pour les femmes. Ainsi, on peut considérer l'erreur d'ajustement sur le scénario central trop importante pour considérer les scénarios bicentennaires comme étant pertinents.

	SMR _{Longevity/Central}	SMR _{Historique/Central}	SMR _{Mortality/Central}
LC	96%	94%	104%
CBD	96%	99%	104%
BFK alternatif	89%	90%	112%

TABLE 6.6: *Backtesting* pour les hommes par modèle

Pour les hommes, il s'agit du modèle de LC qui ne répond pas aux critères du *Backtesting*, en effet, le SMR de longévité bicentenaire est inférieur à celui historique. On peut également mentionner le modèle de BFK alternatif pour lequel le SMR historique est relativement éloigné de 100% et proche de celui de longévité bicentenaire.

En résumé, le seul modèle qui passe avec succès le *backtesting* pour les deux genres est celui de BFK alternatif. Quant au modèle de LC, il demeure pertinent étant donné que les SMR historiques sont proches de 100% et que lorsque le *backtesting* échoue, l'écart reste modéré. En revanche, ce critère met de nouveau en évidence la non-robustesse du modèle de CBD dans l'estimation de la mortalité pour les femmes. Enfin, il convient d'interpréter avec précaution ce critère de prudence, en effet, la profondeur d'historique de calibrage devient relativement faible en raison de la segmentation de l'intervalle temporel initial, ce qui pourrait dégrader la qualité des calibrages.

6.4 Impact de la COVID-19 sur l'estimation de la mortalité prospective

Dans cette section et pour chacun des modèles, on étudiera les effets de l'intégration de la COVID-19 dans les données de calibrage, soit 2020 dans notre cas. On peut d'ores et déjà identifier trois effets de la prise en compte de la pandémie :

- Calibrage : L'impact de l'intégration d'une année de pandémie sur le calibrage des paramètres des modèles ;
- Tendances : Le niveau des paramètres temporels intégrés relativement aux tendances historiques et les impacts sur l'évaluation prospective de ces paramètres ;
- Volatilité : L'impact sur la volatilité du paramètre de tendance, définissant le niveau de diffusion des scénarios stochastiques.

L'impact de l'intégration d'une année de pandémie sur le calibrage des modèles est considéré comme négligeable relativement aux autres effets. Premièrement, en ce qui concerne les paramètres liés à l'âge, l'ajout d'une année atypique dans l'historique constitue une fraction trop limitée pour entraîner une modification significative des calibrages. Deuxièmement, en ce qui concerne les paramètres temporels, leur modification est généralement marginale, se manifestant principalement dans les ajustements de fin de calibrage. Par conséquent, notre analyse se concentrera uniquement sur les effets liés aux composantes de tendance et de volatilité.

6.4.1 Lee-Carter

Le modèle de LC donne une expression du logarithme du taux de mortalité, avec α_x le niveau de base de la mortalité par âge, β_x la mesure pour chaque âge de la sensibilité aux évolutions temporelles et κ_t l'indicateur des tendances temporelles. Pour ces trois paramètres, on va observer et commenter leur niveau en fonction de l'âge et du temps, et ce, pour les deux genres :

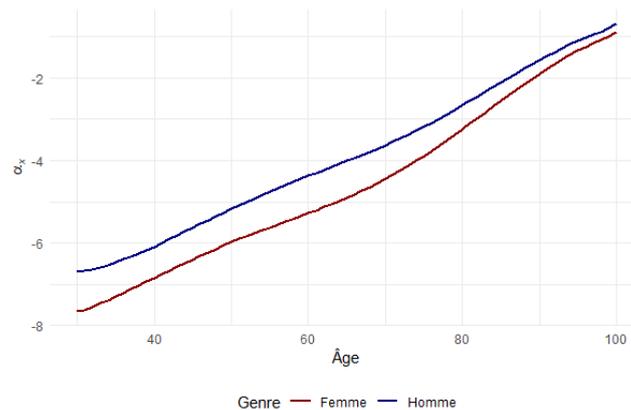


FIGURE 6.7: Paramètre α_x par genre pour le modèle de LC

Le paramètre α présente une tendance linéaire croissante, traduisant une hausse de la mortalité en fonction de l'âge. Son niveau est supérieur chez les hommes, ce constat est en accord avec les observations historiques constatées.

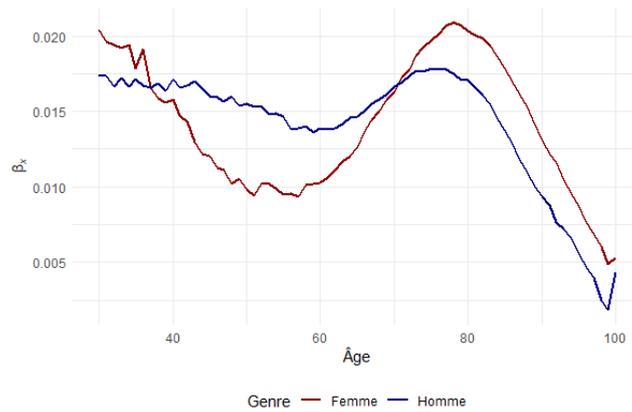


FIGURE 6.8: Paramètre β_x par genre pour le modèle de LC

Le paramètre β montre un niveau relativement stable entre 30 et 80 ans puis une forte décroissance au-delà, cela signifie que les tendances aux améliorations temporelles sont moindres aux âges élevés. Ensuite, la structure par âge est similaire pour les hommes et les femmes, avec des variations plus prononcées pour ces dernières, cela indique que, bien que les femmes et les hommes répondent d'une façon globalement similaire aux changements temporels en termes de structure d'âge, les femmes sont plus sensibles à ces changements.

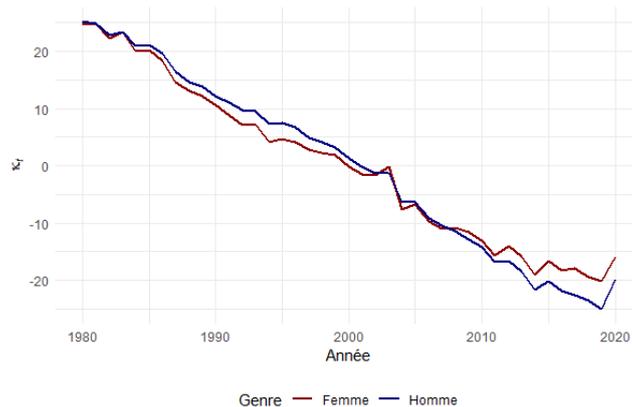


FIGURE 6.9: Paramètre κ_t par genre pour le modèle de LC

Le paramètre κ , en valeur très proche pour les femmes et les hommes, décroît de manière plutôt linéaire. Cette baisse régulière illustre une amélioration constante des taux de mortalité au cours du temps, ce qui corrobore les observations énoncées en première partie de ce mémoire. Toutefois, une inflexion notable de cette tendance est observée en 2020, reflétant l'augmentation de la mortalité générale liée à la COVID-19.

Cette augmentation aura deux impacts sur les projections de mortalité futures, elle influencera non seulement la tendance projetée mais aussi la volatilité des scénarios stochastiques :

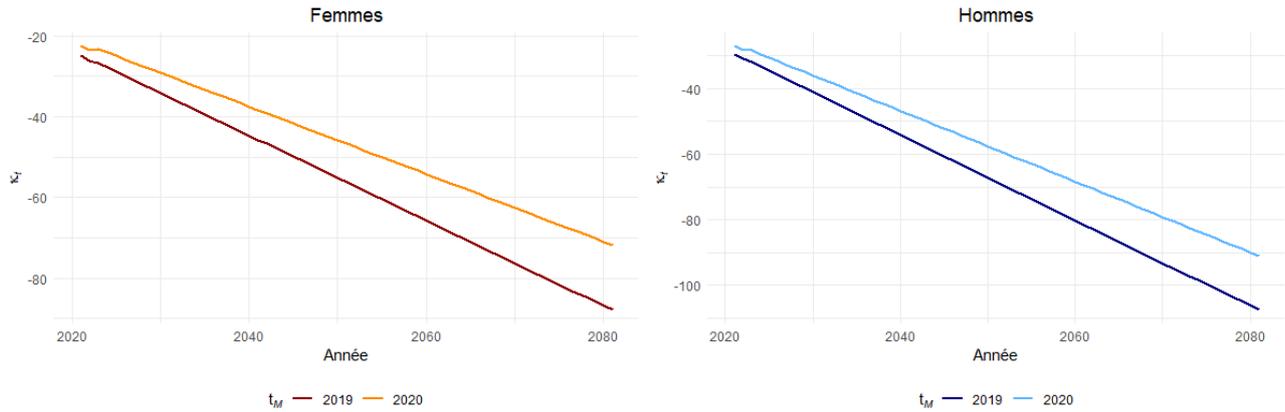


FIGURE 6.10: Projection du paramètre κ_t par genre pour le modèle de LC

Pour rappel, le paramètre κ est projeté à partir de la dernière observation et selon une régression linéaire pondérée, de paramètre de pondération $h = 10$ pour le scénario central. Ainsi selon cette méthode, le calibrage des κ impacte la projection de la tendance de la manière suivante :

- La projection de la tendance démarrant au dernier point calibré, soit t_M , la hausse soudaine de κ_{t_M} lorsqu'on intègre 2020 à notre calibrage implique un point de départ de la projection plus élevé, cela s'observe par une légère translation à la hausse des courbes ;
- La projection de la tendance étant basée sur les tendances historiques avec sur-pondération des années plus récentes, lorsqu'on intègre un point nettement plus élevé que la tendance initiale, cela a pour effet de redresser la pente et s'observe sur les graphiques avec une décroissance du paramètre moins prononcée.

L'impact sur la volatilité des scénarios stochastiques suit une logique similaire à celui constaté sur la tendance. Pour rappel, la volatilité correspond à une moyenne pondérée (avec $h = 10$ ou 30) des écarts au carré entre l'erreur d'estimation de κ_t compte tenu des informations disponibles à $t - 1$ et l'erreur d'estimation moyenne.

Jusqu'en 2019, les κ se sont montrés stables et ainsi les écarts d'estimations, et donc la volatilité, étaient relativement faibles. Néanmoins, l'intégration de 2020 génère une hausse significative de la volatilité en raison de l'ajout d'un point extrême à l'historique. Cette hausse est d'autant plus prononcée que les années récentes bénéficient d'une pondération supérieure :

Année	Genre	$h = 10$	$h = 30$
2020	Femmes	234%	229%
2019	Femmes	167%	183%
2020	Hommes	250%	205%
2019	Hommes	130%	129%

TABLE 6.7: Impact sur $\sigma_{Historique,h}$ de l'intégration de 2020 dans le calibrage par genre pour le modèle de LC

L'intégration de l'année 2020 dans le calibrage exerce une influence significative sur la volatilité. Cet impact est en moyenne 59% plus conséquent pour $h = 10$, étant donné que la volatilité récente reçoit une pondération plus importante que pour $h = 30$. De plus, l'effet est plus prononcé chez les hommes, en cohérence avec les chiffres indiquant un impact de la COVID-19 plus marqué pour ces derniers. En chiffres, cela se traduit par une augmentation de la volatilité qui est en moyenne 132% plus élevée pour les hommes que pour les femmes.

6.4.2 Cairns-Blake-Dowd

Le modèle de CBD donne une expression du logit du quotient de mortalité, avec $\kappa_t^{(1)}$ l'indicateur des tendances temporelles sur l'ensemble de la population étudiée indépendamment des âges et $\kappa_t^{(2)}$ le paramètre de pente ayant pour vocation de capter les différences de tendance entre les âges. Pour ces deux paramètres, on va observer et commenter leur niveau en fonction du temps, et ce, pour les deux genres :

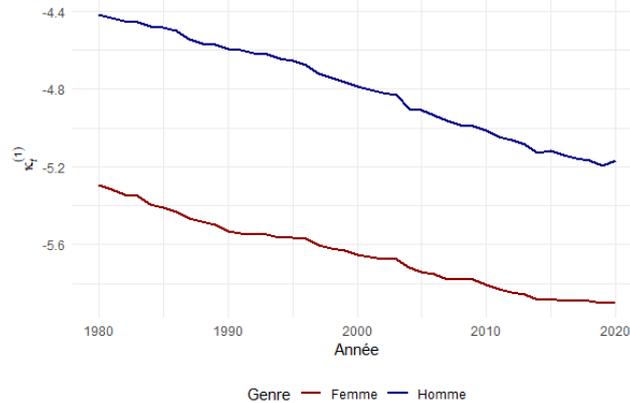


FIGURE 6.11: Paramètre $\kappa_t^{(1)}$ par genre pour le modèle de CBD

A l'instar du modèle de LC, le paramètre $\kappa_t^{(1)}$ présente une décroissance linéaire pour les femmes et les hommes, illustrant une amélioration continue des quotients de mortalité avec le temps. Une inflexion dans cette tendance est observée en 2020, illustrant l'impact de la COVID-19 sur l'augmentation de la mortalité générale. Cependant, cette augmentation est moins marquée que celle observée dans le modèle de LC, en partie grâce à l'intervention du paramètre $\kappa_t^{(2)}$. Néanmoins, cette hausse modérée aura malgré tout une incidence sur la tendance projetée et la volatilité des scénarios stochastiques.

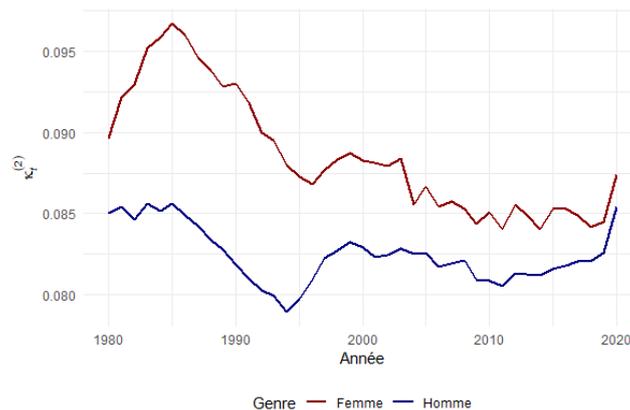


FIGURE 6.12: Paramètre $\kappa_t^{(2)}$ par genre pour le modèle de CBD

Le paramètre $\kappa_t^{(2)}$ représente la structure de la mortalité par âge, soit la proportion des phénomènes de mortalité spécifiques à certains âges et qui ne sont pas expliqués par une tendance générale. On constate une tendance décroissante du paramètre pour les femmes tandis qu'il demeure relativement stable chez les hommes. Cela suggère qu'au fil des années, les effets liés à l'âge sur la mortalité se sont atténués chez les femmes.

Pour l'année 2020, on note une augmentation tant chez les femmes que chez les hommes, ce qui est attribuable à la COVID-19 dont les effets sur la mortalité ont été fortement corrélés à l'âge.

L'incorporation de l'impact de la COVID-19 dans $\kappa^{(2)}$ est donc tout à fait appropriée puisqu'elle reflète un effet dépendant de l'âge.

Après avoir constaté l'impact de l'intégration de l'année 2020 à notre historique de calibrage, on oriente désormais notre analyse sur l'impact de ces évolutions sur la projection des paramètres :

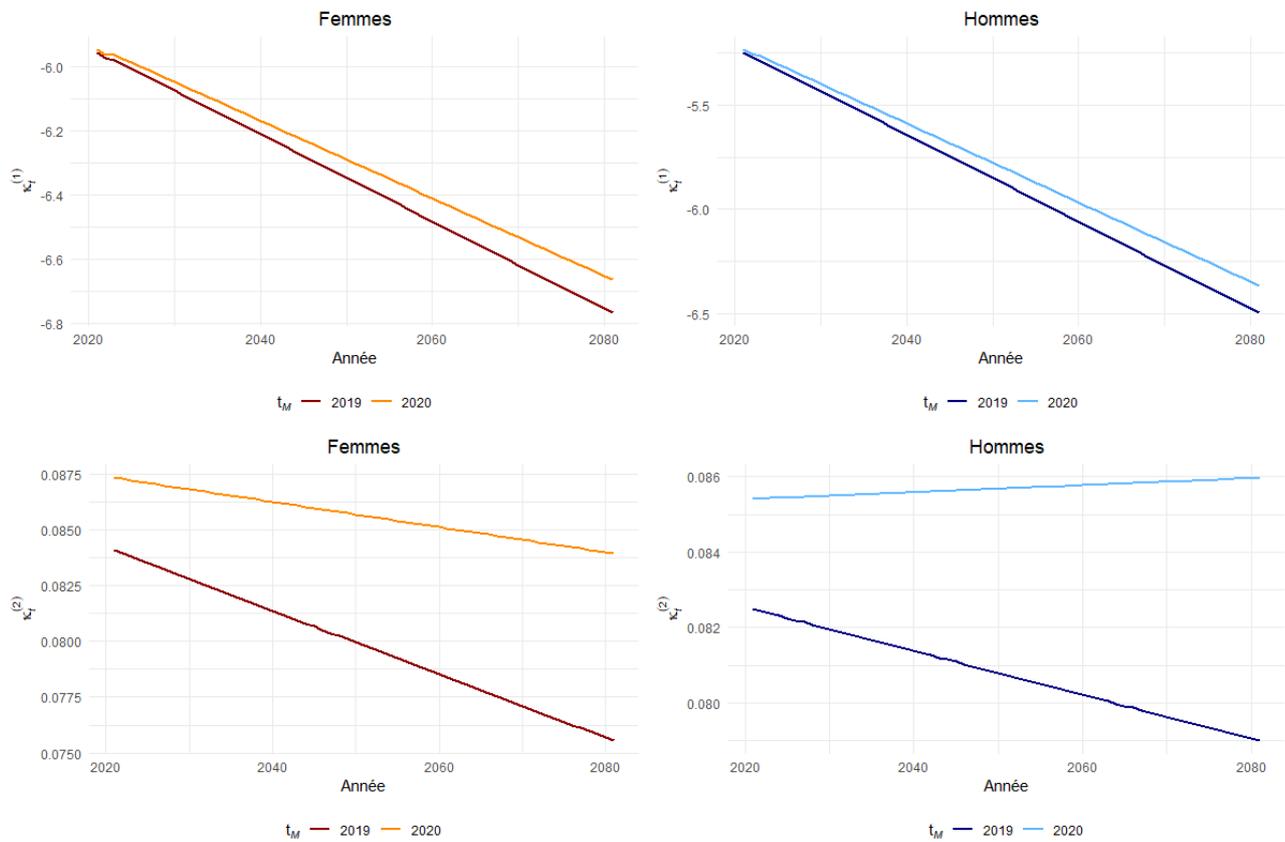


FIGURE 6.13: Projection des paramètres $\kappa_t^{(1)}$ et $\kappa_t^{(2)}$ par genre pour le modèle de CBD

Le paramètre $\kappa^{(1)}$ n'est que peu impacté par l'intégration de 2020, la nouvelle observation étant peu éloignée de la tendance initiale. Cependant, les impacts restent, dans une moindre mesure, identiques à ceux du modèle de LC :

- La prise en compte de la COVID-19 dans notre calibrage a impact relativement faible sur la translation de la tendance. En effet, la dernière observation, point de départ de la projection du paramètre, se situe à un niveau comparable, que l'on intègre ou pas l'année 2020 ;
- L'impact sur la pente est également limité, il reste cependant plus prononcé que le précédent en raison de la sur-pondération des dernières années. La nouvelle observation, se situant au-dessus de la tendance passée, entraîne un redressement de la pente, suggérant ainsi une réduction de l'ampleur des améliorations de mortalité.

L'analyse de l'impact sur le paramètre $\kappa^{(2)}$ est similaire et relativement simple car la méthode utilisée repose sur la projection de l'écart moyen historique. La nouvelle observation est nettement plus élevée que les observations antérieures, qui montraient une tendance à la baisse ou stable, ce qui crée un écart positif significatif. Cet écart élève l'écart moyen, entraînant ainsi un réhaussement de la projection du paramètre et donc une stabilisation de la proportion des phénomènes de mortalités spécifiques à certains âges.

On note donc que l'impact de la COVID-19 se manifeste essentiellement sur le paramètre $\kappa^{(2)}$. Ce paramètre n'étant pas projeté de manière stochastique, les scénarios bicentennaires ne devraient alors être que peu impactés. En effet, l'intégration de 2020 à nos données n'a qu'un effet marginal sur $\kappa^{(1)}$, ce qui devrait générer un surplus de volatilité modéré. De plus, la tendance de ce paramètre est remarquablement stable, ce qui implique une volatilité initiale faible :

Année	Genre	$h = 10$	$h = 30$
2020	Femmes	1.97%	1.96%
2019	Femmes	1.77%	1.77%
2020	Hommes	2.44%	2.21%
2019	Hommes	1.70%	1.73%

TABLE 6.8: Impact sur $\sigma_{Historique,h}$ de l'intégration de 2020 dans le calibrage par genre pour le modèle de CBD

Tout comme pour le modèle de LC, l'intégration de l'année 2020 dans le calibrage exerce une influence sur la volatilité, cet impact est en moyenne 46% plus conséquent pour $h = 10$. De la même façon, on constate un effet plus fort chez les hommes, avec une augmentation moyenne de la volatilité de 232%.

Cependant, les données de ce tableau indiquent que le paramètre $\kappa^{(1)}$ du modèle de CBD est moins affecté par la COVID-19, en particulier pour les femmes, avec une augmentation moyenne de la volatilité de 1%, comparativement à une hausse de 76% observée pour le modèle de LC.

6.4.3 Börger-Fleischer-Kuksin alternatif

Le modèle de BFK alternatif donne une expression du logit du quotient de mortalité. Le paramètre α_x représente le niveau de base de la mortalité par âge, $\kappa_t^{(1)}$ l'indicateur des tendances temporelles sur l'ensemble de la population, $\kappa_t^{(2)}$ le paramètre de pente permettant de capter les différences de tendance entre les âges et $\kappa_t^{(3)}$ le paramètre temporel spécifique aux âges avancés. Pour ces quatre paramètres, on va observer et commenter leur niveau en fonction de l'âge et du temps, et ce, pour les deux genres :

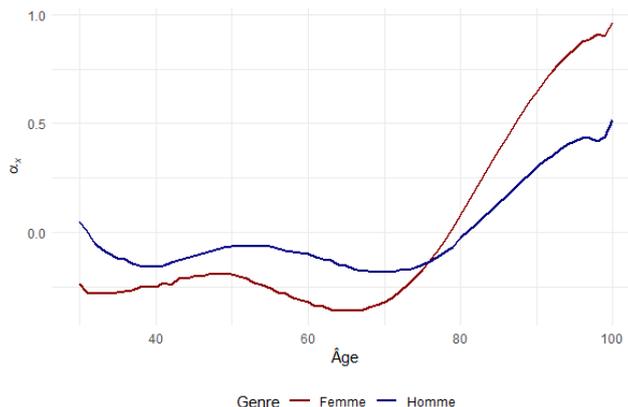


FIGURE 6.14: Paramètre α_x par genre pour le modèle de BFK alternatif

Le paramètre α demeure assez stable jusqu'à 70 ans puis croît fortement au-delà, reflétant ainsi la hausse de mortalité aux âges élevés. En complément, on note que les trajectoires de ce paramètre sont assez comparables pour les femmes et hommes avec une amplitude plus prononcées des variations pour les femmes, reflétant ainsi une plus forte sensibilité à l'âge sur la mortalité.

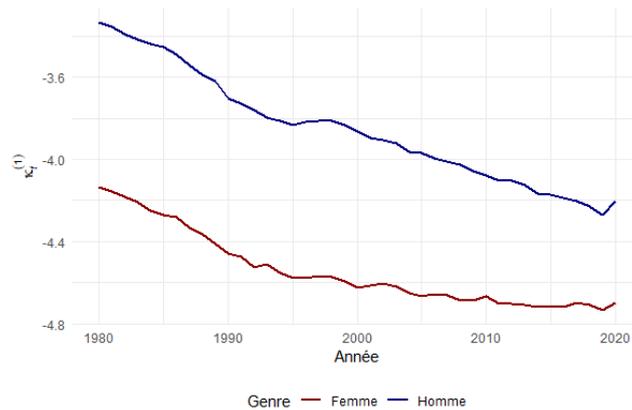


FIGURE 6.15: Paramètre $\kappa_t^{(1)}$ par genre pour le modèle de BFK alternatif

Tout comme dans les modèles précédents, le paramètre $\kappa^{(1)}$ affiche une tendance linéaire à la baisse pour les femmes et les hommes, ce qui illustre l'amélioration continue des quotients de mortalité dans le temps.

En 2020, on observe également une inflexion dans cette tendance, reflétant l'impact de la COVID-19 sur l'accroissement de la mortalité globale. Comme constaté dans le modèle de CBD, cette augmentation est moins marquée que celle observée dans le modèle de LC, en partie grâce à l'intervention des paramètres $\kappa^{(2)}$ et $\kappa^{(3)}$. En revanche, cette hausse modérée influencera aussi bien la tendance projetée que la volatilité des scénarios stochastiques.

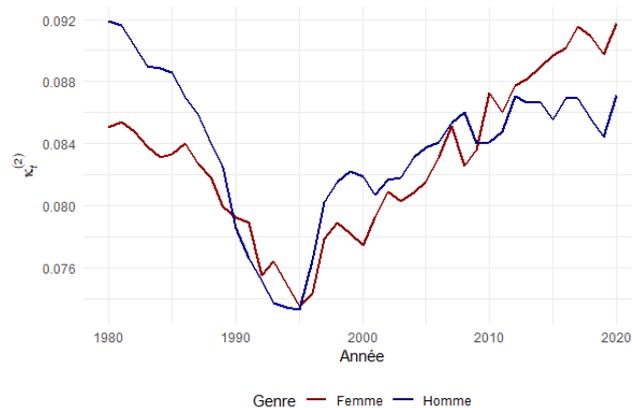


FIGURE 6.16: Paramètre $\kappa_t^{(2)}$ par genre pour le modèle de BFK alternatif

Le paramètre $\kappa^{(2)}$ illustre la structure de la mortalité par âge, soit la proportion des phénomènes de mortalité spécifiques à certains âges et qui ne sont pas expliqués par une tendance générale. Après une période de décroissance jusqu'en 1995, on constate une tendance croissante du paramètre pour les deux genres. Cette nouvelle tendance suggère que, avec le temps, les effets qui ne sont pas expliqués par la tendance générale se renforcent.

Pour l'année 2020, une hausse significative est perceptible tant chez les femmes que chez les hommes, un phénomène relié à la pandémie dont l'impact sur la mortalité ont été fortement corrélés à l'âge. Ainsi, l'intégration de ce phénomène au paramètre $\kappa^{(2)}$ est pertinente puisqu'elle représente précisément un effet lié à l'âge.

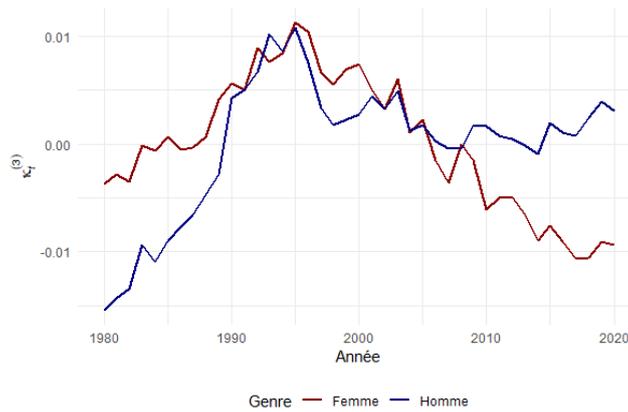


FIGURE 6.17: Paramètre $\kappa_t^{(3)}$ par genre pour le modèle de BFK alternatif

Le paramètre $\kappa^{(3)}$ représente les effets de mortalité supplémentaires aux âges élevés. Ce dernier évolue de façon symétrique à $\kappa^{(2)}$, soit une tendance haussière jusqu'à 1995 puis baissière pour les femmes et stable pour les hommes. Il convient de souligner que la valeur du paramètre est relativement proche de zéro, ce qui implique une contribution marginale de ce paramètre au niveau des quotients de mortalité. Ceci indique qu'il n'existe que peu d'effets temporels résiduels non capturés par $\kappa^{(1)}$ et $\kappa^{(2)}$.

La prise en compte de 2020 à l'historique de calibrage a un effet négatif et négligeable sur le paramètre, suggérant, de manière quelque peu contre-intuitive, une réduction des effets spécifiques aux âges avancés.

Suite à l'étude de l'impact de l'intégration de 2020 à notre historique de calibrage, on oriente désormais notre analyse sur les répercussions de ces changements sur la projection des paramètres :

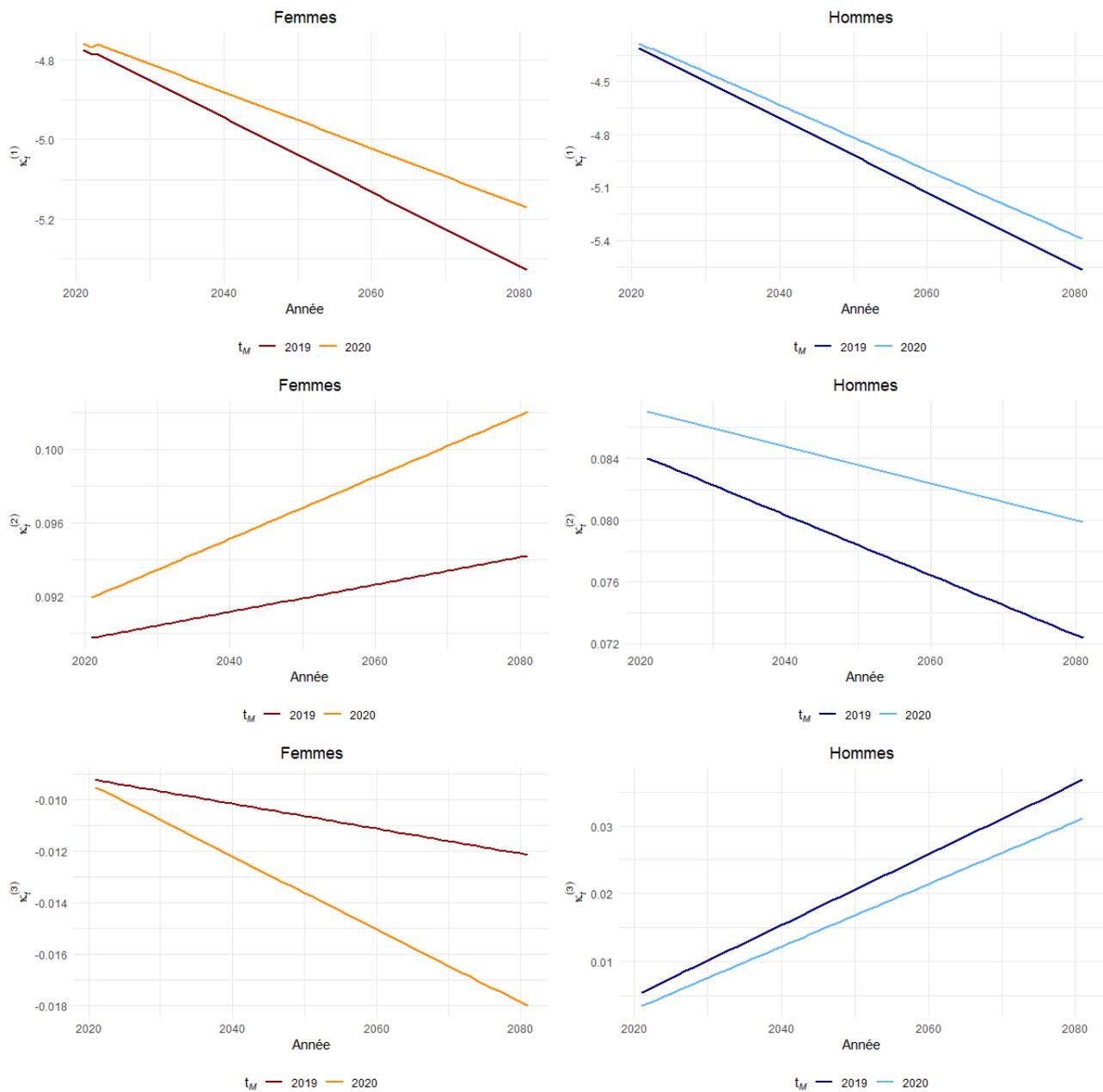


FIGURE 6.18: Projection des paramètres $\kappa_t^{(1)}$, $\kappa_t^{(2)}$ et $\kappa_t^{(3)}$ par genre pour le modèle de BFK alternatif

Similairement au modèle de CBD, le paramètre $\kappa^{(1)}$ subit un impact peu matériel suite à l'ajout de l'année COVID-19 en raison de la proximité entre la nouvelle observation et la tendance existante. Les effets restent en accord avec ceux identifiés dans le modèle de CBD, soit un impact marginal sur la translation de la tendance et une influence limitée sur le redressement de la pente.

Concernant le paramètre $\kappa^{(2)}$ (resp. $\kappa^{(3)}$), l'écart entre la nouvelle observation et la précédente étant supérieur (resp. inférieur) aux écarts historiques, cela entraîne une révision à la hausse (resp. baisse) de la moyenne et donc de la pentification du paramètre projeté.

Les effets de la COVID-19 se reflètent majoritairement dans les paramètres $\kappa^{(2)}$ et $\kappa^{(3)}$ et l'impact sur $\kappa^{(1)}$ demeure limité. De plus, la tendance historique est relativement stable, ainsi, pareillement au modèle de CBD, la volatilité ne devrait être que peu affectée :

Année	Genre	$h = 10$	$h = 30$
2020	Femmes	1.95%	3.65%
2019	Femmes	1.91%	3.62%
2020	Hommes	2.80%	3.49%
2019	Hommes	2.29%	3.39%

TABLE 6.9: Impact sur $\sigma_{Historique,h}$ de l'intégration de 2020 dans le calibrage par genre pour le modèle de BFK alternatif

Comme attendu et observé dans les modèles antérieurs, l'intégration de l'année 2020 élève le niveau de volatilité, avec un impact moyen de 548% plus conséquent pour $h = 10$. De façon homologue, l'effet est accentué chez les hommes avec une augmentation moyenne de la volatilité de 940%.

Il convient néanmoins de mentionner les effets d'échelle dans les résultats énoncés. D'une part, la hausse de la volatilité est négligeable sur les femmes, avec 1% en moyenne contre 13% pour les hommes rendant l'écart relatif conséquent. D'autre part, on constate peu d'impact sur la volatilité lorsque $h = 30$, avec une hausse moyenne de 2%, comparativement à $h = 10$ où l'on note une augmentation moyenne à 12%.

La COVID-19 a donc un impact notable sur l'évaluation de la mortalité prospective, avec des effets variables en fonction des modèles. Cet impact résulte de l'ajout d'une nouvelle observation dans les paramètres temporels, qui s'écarte sensiblement des tendances historiques. Cette observation ajoutée influence, à divers degrés selon les modèles et paramètres, les projections déterministes et stochastiques. Tout d'abord, l'influence se manifeste dans une faible mesure dans la translation du niveau de tendance. Notre approche consistant à projeter les paramètres temporels en partant de la dernière donnée disponible, l'ajout de 2020 modifie donc dans une certaine mesure ce point de départ.

Ensuite, cette nouvelle observation n'étant pas de même ordre que les tendances historiques, cela vient déformer la pentification des paramètres projetés. Cette déformation est plus prononcée pour le paramètre de tendance générale pour lequel les données récentes sont davantage pondérées.

Enfin, dans l'évaluation de la volatilité du paramètre de tendance générale, et suivant une logique homologue, le fort éloignement du nouveau point et de son estimation basée sur les données antérieures génère une hausse significative de la volatilité, en particulier lorsque le paramètre de pondération est faible.

Pour chacun des modèles étudiés, on a donc effectué le constat et l'analyse de l'impact de l'intégration de la COVID-19 sur les paramètres calibrés et projetés. On va désormais observer les conséquences de ces évolutions sur les scénarios bicentennaires de mortalité prospective, soit les *Reduction Rates*, et plus spécifiquement sur ceux de longévité :

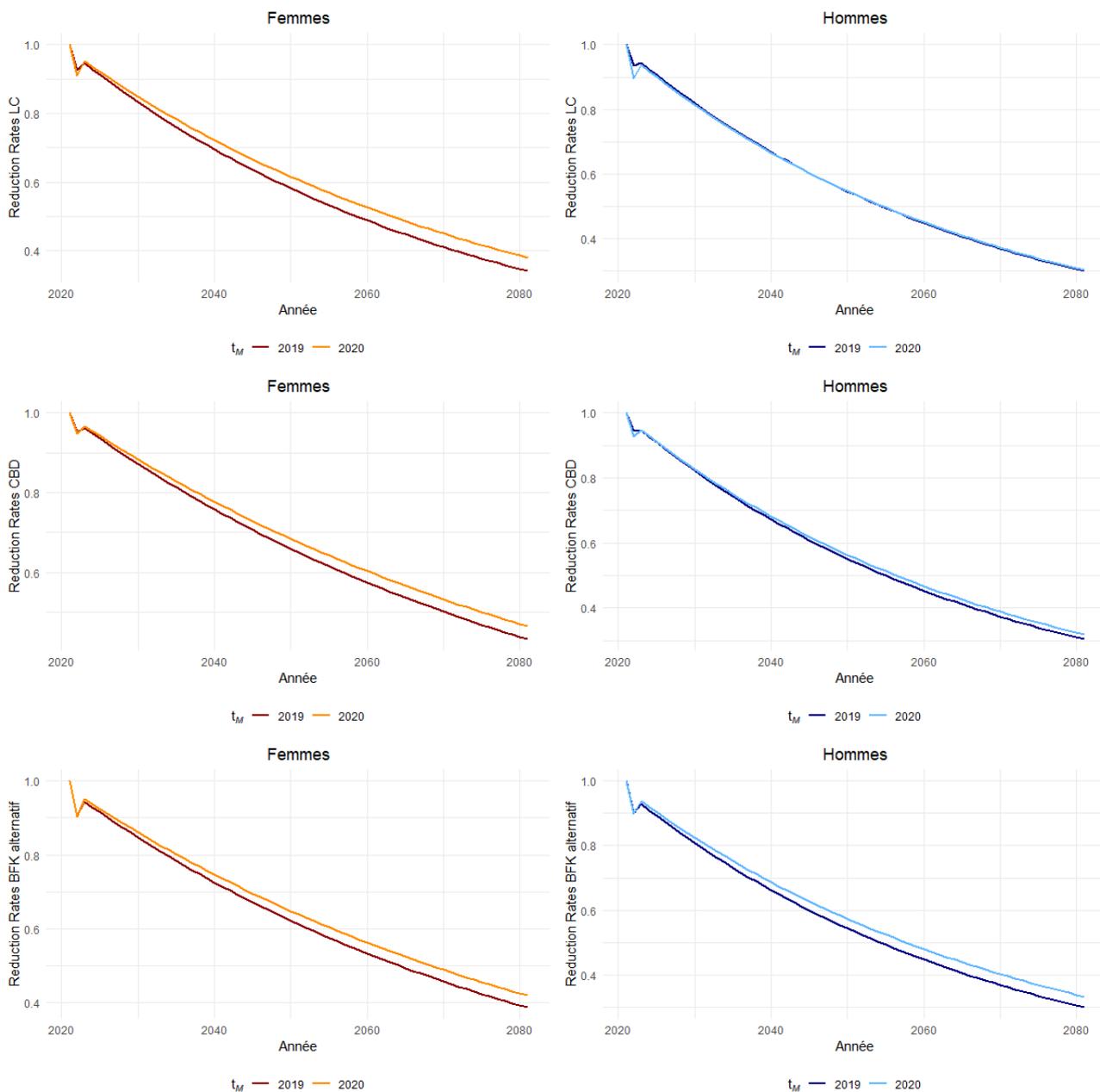


FIGURE 6.19: Comparaison des *Reduction Rates* bicentennaires de longévité moyens par genre pour les modèles de LC, CBD et BFK alternatif

Pour chaque genre et modèle, les *Reduction Rates* diminuent dans le temps, ce qui est en adéquation avec la décroissance des paramètres de tendance générale.

Un pic est observé en début de projection, il reflète l'événement bicentenaire affectant les tendances futures. Ce dernier se résorbe par la suite, notamment en raison du retour à une projection déterministe, de paramètre de pondération $h = 3$, qui dilue dans une certaine mesure, l'année choquée avec celles antérieures.

En ce qui concerne le niveau des améliorations à horizon de projection, pour les femmes, le modèle de CBD offre les améliorations moyennes les moins prononcées, tandis que le modèle de LC affiche les plus marquées. Pour les hommes, les améliorations sont approximativement identiques entre les modèles.

Cependant, ces graphiques ne permettent pas d'illustrer la capacité des modèles à absorber l'année COVID-19 dans les scénarios bicentennaires de longévité. En réalité, le scénario central est également influencé par l'intégration de l'année 2020 dans le calibrage au travers du redressement de la tendance générale projetée qui peut parfois compenser la hausse de volatilité, à l'image du modèle de LC pour les hommes :

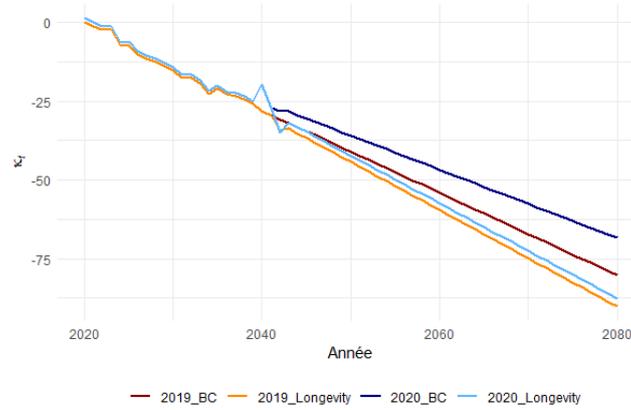


FIGURE 6.20: Illustration de la compensation entre hausse de volatilité et redressement de tendance centrale pour les hommes, pour le modèle de LC

Le graphique illustre l'équilibre entre l'augmentation de la volatilité, visible par la hausse de l'écart entre le scénario central et le scénario bicentenaire, et l'ajustement de la tendance générale, qui se s'observe par la comparaison des pentes des scénarios centraux .

Afin de mesurer la capacité d'absorption des modèles, il serait donc plus judicieux d'examiner l'évolution de l'écart entre le scénario bicentenaire et le scénario central. Soit, pour tout $t = t_M, \dots, t_{proj}$, avec une éventuelle pondération ω :

$$\text{Avg Gap Evol}_t = \frac{\sum_{i=x_m}^{x_M} RR_{i,t}^{\text{Longevity},2020} \times \omega_i}{\sum_{i=x_m}^{x_M} RR_{i,t}^{\text{Central},2020} \times \omega_i} \times \frac{\sum_{i=x_m}^{x_M} RR_{i,t}^{\text{Central},2019} \times \omega_i}{\sum_{i=x_m}^{x_M} RR_{i,t}^{\text{Longevity},2019} \times \omega_i}$$

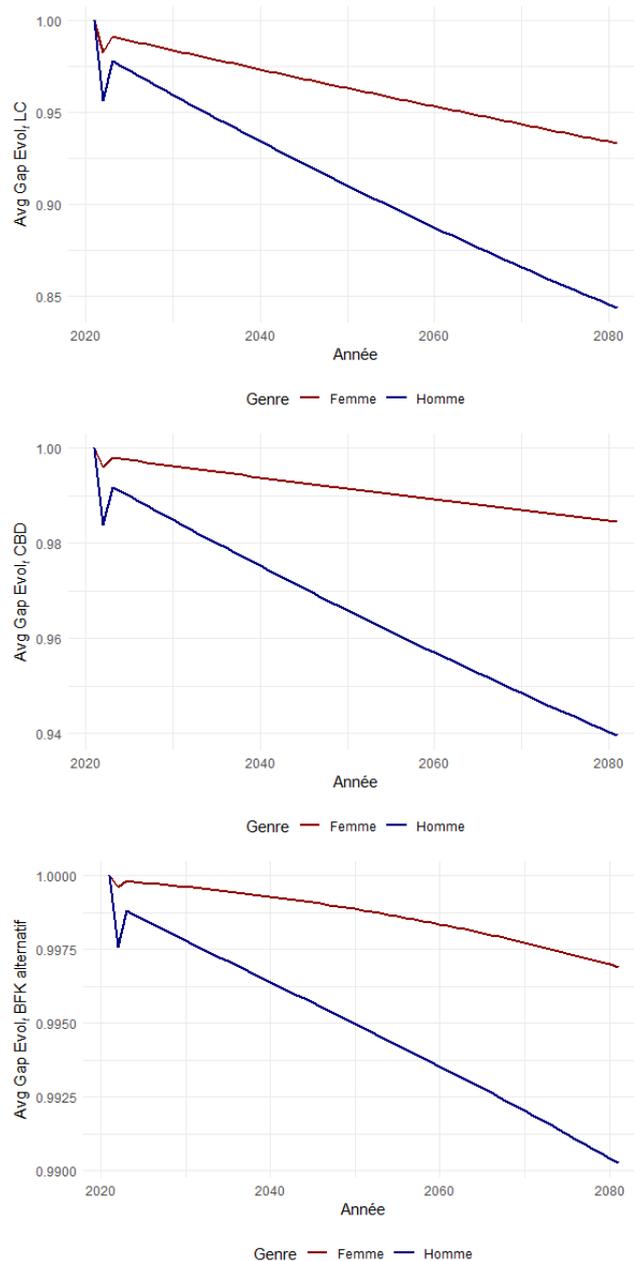


FIGURE 6.21: Avg Gap Evol_t par genre pour les modèles de LC, CBD et BFK alternatif

Les graphiques montrent l'évolution des écarts entre les scénarios bicentennaires et centraux concordent avec les observations précédentes. D'une part, l'impact plus marqué de la COVID-19 sur les hommes en comparaison aux femmes se manifeste par une décroissance plus importante de l'Avg Gap Evol pour ces derniers. D'autre part, l'augmentation de la volatilité est visible par l'accentuation du pic de début de projection correspondant à l'événement bicentenaire.

On peut définir la capacité d'absorption de la COVID-19 par le niveau de l'Avg Gap Evol à horizon de projection. Pour le modèle de LC, cet écart atteint 84% pour les hommes et 93% pour les femmes, témoignant de l'intégration de l'événement pandémique dans le paramètre de tendance générale. Concernant le modèle de CBD, le paramètre de tendance générale est moins affecté, ce qui se traduit par un écart à l'horizon de projection de 98% pour les femmes et 94% pour les hommes. Enfin, le modèle de BFK alternatif répond à sa fonction initiale, absorbant efficacement l'impact de la COVID-19 avec un écart avoisinant 100% pour les femmes et 99% pour les hommes.

Or, il est légitime de s'interroger sur la pertinence de considérer les événements pandémiques dans l'évaluation des tendances de mortalité bicentennaires. En effet, une pandémie ne correspond pas à un phénomène de tendance et s'en distingue nettement, car elle est typiquement de courte durée, à

l'opposé des tendances qui s'inscrivent dans le temps. Par conséquent, dans le prochain et dernier chapitre, l'approche sera d'examiner les méthodes envisageables pour évaluer la mortalité prospective bicentenaire en tenant compte des données COVID-19. Cette analyse se déclinera en trois axes :

- Robustesse du modèle de Börger-Fleischer-Kuksin alternatif : On expliquera, en se basant sur les critères définis et observations réalisées jusqu'à présent, pourquoi le modèle de BFK alternatif est le mieux adapté à l'évaluation de la mortalité prospective bicentenaire dans ce contexte ;
- Retraitement des données de mortalité : On étudiera les possibilités de retraitements des données affectées par la pandémie afin de mitiger l'impact sur les scénarios de mortalité bicentenaires ;
- Ajustement des paramètres et de la modélisation : Tout en s'assurant de leur pertinence, on effectuera des sensibilités sur les paramètres de modélisation afin de définir les paramètres offrant le meilleur compromis entre pertinence et absorption des impacts COVID-19. Suite à cela, on exposera les axes d'améliorations sur la modélisation.

Traitements envisageables pour la prise en compte de la COVID-19 pour les scénarios bicentennaires

7.1 Définition du portefeuille de mesure d'impact

Les scénarios bicentennaires de mortalité, sous forme de *Reduction Rates*, ne permettent pas de visualiser concrètement les impacts de la dérive de mortalité sur un portefeuille d'assurés. Notre étude étant principalement axée sur le risque de longévité, on va alors considérer un portefeuille de retraite comme référence afin d'effectuer l'ensemble des mesures d'impact sur le besoin en capital et spécifiquement pour la composante tendance (ou *Trend*) du risque de longévité.

L'assurance retraite vise à apporter un complément de revenu aux bénéficiaires lorsqu'ils parviennent à l'âge de la retraite. Elle s'appuie sur les versements effectués par l'assuré et/ou par son employeur pendant la période d'activité professionnelle. Ces versements sont investis dans différents supports financiers afin de produire des revenus, qui seront ensuite redistribués sous forme de rente ou de capital à l'âge de la retraite. Le montant versé, que ce soit sous forme de rente ou de capital, dépend en grande partie des sommes cotisées, de la durée de la cotisation et des termes spécifiques stipulés dans le contrat. Lorsque le bénéficiaire opte, au départ à la retraite, pour la conversion de son capital en rente, l'assureur fait face au risque de longévité puisque le calcul de la rente repose sur l'hypothèse de mortalité BE. Par nature, cette estimation n'anticipe pas les évolutions futures la mortalité atypiques, ainsi, en cas de sous-mortalité par rapport à la vision BE, les rentes versées au-delà de ce qui a été estimé par l'assureur sont à la charge de ce dernier.

Afin de mesurer ce surcoût, on utilisera le modèle interne d'AXA France, il s'agit d'un outil de projection des flux d'actif et de passif avec prise en compte des interaction Actif-Passif ayant pour vocation de représenter le plus fidèlement possible l'évolution de la compagnie, en alliant complexité et traçabilité. Ce dernier prend en compte les éléments suivants :

- L'évolution de l'actif, avec notamment la prise en compte de la réallocation des actifs ou la réalisation des PMVL¹ ;
- L'évolution du passif, avec notamment la prise en compte de la mortalité des assurés, leur comportement ou la revalorisation des contrats ;
- Les décisions du top management au travers du pilotage de la PB ou de l'allocation stratégique ;
- La réglementation avec par exemple le respect du minimum de PB au niveau compagnie.

1. Plus ou Moins Values Latentes

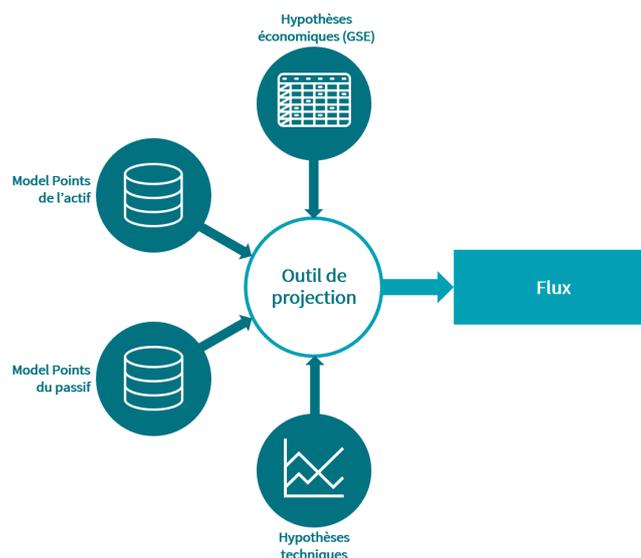


FIGURE 7.1: Structure simplifiée d'un modèle ALM

Les *Model Points*, que ce soit de l'actif ou du passif, sont des tables de données du portefeuille de la compagnie agrégées en groupes homogènes offrant le meilleur compromis entre agrégation et conservation de la qualité de l'information :

- Pour l'actif, ils représentent les actifs détenus par la compagnie, regroupés selon leurs caractéristiques (nature, type de taux, maturité, etc...);
- Pour le passif, ils représentent les contrats en portefeuille, selon les caractéristiques (produit, âge, TMG, etc...).

De sorte à limiter la complexité des calculs dans le modèle, on y intègre des hypothèses établies en amont de la projection. On peut distinguer les hypothèses économiques et techniques :

- Les hypothèses économiques ont un rôle majeur dans les calculs de l'assureur puisqu'elles permettent de prévoir les différentes situations économiques dans lesquelles l'assureur pourra potentiellement se trouver. Ces hypothèses sont généralement regroupées au sein d'un GSE², il s'agit d'un outil de simulation stochastique qui permet de diffuser pour chaque simulation et sur l'horizon de projection, les évolutions des facteurs de risques financiers auxquels l'assureur est exposé (courbes de taux, inflation, *spread*, etc...);
- Les hypothèses techniques couvrent l'ensemble des hypothèses qui ne sont pas économiques et sont très variées (mortalité, frais, rachats, etc...). Ces hypothèses sont établies grâce à des études statistiques, principalement sur les données du portefeuille de l'assureur.

Dans le modèle on dispose de deux hypothèses de mortalité, d'une part la mortalité attendue et d'autre part la mortalité observée. Dans un scénario central (ou BE), ces hypothèses se confondent puisqu'on suppose que notre meilleure estimation de la mortalité va se réaliser et donc qu'aucune dérive ne sera observée.

Dans le cas d'un scénario de longévité, on va remplacer la mortalité observée par la table sur laquelle on aura appliqué nos scénarios bicentennaires mortalité. Cela permettra de mesurer l'impact de la dérive sur les fonds propres et ainsi évaluer le niveau de capital de solvabilité à détenir pour la composante tendance du sous-module longévité.

Plus spécifiquement, pour le portefeuille de retraite servant de mesure d'impact, la table de mortalité BE est la TGHF05³, il s'agit de la table réglementaire en vigueur pour les opérations de rentes viagères. La table a été élaborée à partir de données de rentiers issues de 19 portefeuilles français, portant sur

2. Générateur de Scénario Économique

3. Table Générationnelle Hommes-Femmes 2005

700 000 rentes liquidées entre 1993 et 2005. En raison du volume limité de données, des modèles utilisant une référence externe de mortalité ont été privilégiés. Pour connaître davantage de détails sur la méthode de création de cette table, on renvoie le lecteur à Planchet (2007) [14].

7.2 Robustesse du modèle de Börger-Fleischer-Kuksin alternatif

Dans cette section, on va démontrer en quoi le modèle de BFK alternatif est approprié pour évaluer la mortalité prospective et particulièrement le contexte de cette étude. Pour ce faire, on récapitulera les observations faites jusqu'à présent, en commençant par la sélection du modèle, puis en comparant les modèles au travers des impacts constatés sur les paramètres et la volatilité des scénarios bicentennaires. Enfin, en utilisant le modèle interne d'AXA France appliqué au portefeuille de retraite, on évaluera l'effet des scénarios de longévité sur le capital de solvabilité requis de la composante tendance du sous-module de longévité.

Tout d'abord, concernant les critères statistiques, le modèle de BFK alternatif offre, avec le modèle de LC, les meilleurs résultats aux SMR. Les écarts d'estimations sont relativement faibles, avec des oscillations variant entre 10 et -20%, et nuls en moyenne. De plus, les tendances systématiques liées aux cohortes ou aux fermetures de tables sont minimales, résultant à des estimations de mortalité globalement cohérentes.

En revanche, pour le BIC qui est à minimiser, le modèle obtient le score le plus élevé des 3 modèles étudiés, indiquant un compromis entre qualité des ajustements et simplicité moins intéressant. Néanmoins, le BIC seul ne permet pas de rejeter un modèle mais plutôt d'affiner le choix entre les modèles lorsque les autres critères ne permettent pas de discrimination claire. Ce critère est donc à considérer après une évaluation complète de l'ensemble critères.

Ensuite, concernant les critères de prudence, le modèle de BFK alternatif est le seul modèle étudié passant avec succès les deux critères. Les améliorations d'espérance de vie des scénarios bicentennaires se situent au plus proche des seuils $\epsilon^{(1)}$ et $\epsilon_{inf}^{(2)}$ suggérant ainsi une prudence suffisante et non superflue, idéale pour l'optimisation du capital de solvabilité requis.

Pour le *Backtesting*, le modèle de BFK alternatif passe avec succès le test puisque les SMR historiques sont compris entre les SMR de longévité et mortalité bicentennaires pour les deux genres. Parallèlement, bien que le modèle de LC ne valide pas le test pour les hommes, ce dernier ne doit pas automatiquement être rejeté étant donné que le dépassement est relativement faible.

En poursuivant sur l'impact de l'intégration de l'année 2020 à l'historique de calibrage sur la projection de la tendance et sur la volatilité, on note que le modèle de BFK alternatif semble être le plus adapté à l'absorption de la pandémie.

En effet, la nouvelle observation sur le paramètre de tendance générale, soit $\kappa^{(1)}$, est relativement peu éloignée des tendances historiques étant donné que la dérive est absorbée par les autres paramètres temporels. Cette dérive sur le paramètre de tendance générale se traduit par un faible redressement de la courbe projetée indiquant des améliorations futures moins importantes que celles estimées jusqu'à présent. On peut comparer ce redressement de tendance à celui observé sur le modèle de LC, qui est nettement plus important en raison de l'unicité du paramètre temporel, impliquant nécessairement un impact de l'année pandémique entièrement absorbé par ce paramètre.

L'impact sur la tendance générale se répercutant sur la volatilité, les effets sont de même ordre, soit une hausse de la volatilité importante, de 54%, pour le modèle de LC et moindre pour les modèles de CBD et BFK alternatif. En comparant ces deux derniers modèles, on note que celui de CBD offre la volatilité la plus faible, à environ 1.9% contre 2.9% pour celui de BFK alternatif. Cependant, le modèle de BFK alternatif est nettement moins impacté par l'intégration de la nouvelle observation, avec une hausse de 7% de la volatilité en moyenne contre 23% pour le modèle de CBD, traduisant ainsi une meilleure capacité d'absorption de l'année COVID-19 pour le modèle proposé.

Enfin, en reprenant l'ensemble des constats effectués ci-dessus, on peut désormais constater l'impact cumulé sur les scénarios bicentennaires de longévité, soit les *Reduction Rates*. Tout d'abord, en termes d'absorption de l'année COVID-19, on retrouve bien cette capacité du modèle de BFK alternatif

dans l'évolution des écarts entre scénario bicentenaire et scénario central. En effet, cette évolution, matérialisée par l'Avg Gap Evol, est négligeable pour le modèle de BFK alternatif, avec un niveau supérieur à 99% à horizon de projection tandis qu'elle s'approche de 80% (resp. 93%) pour le modèle de LC (resp. CBD).

Avant toute chose, il convient de rappeler le phénomène contre-intuitif observé qui peut biaiser l'analyse des résultats. La trajectoire de tendance centrale étant réévaluée à la hausse suite à l'intégration de la COVID-19 et la diffusion des scénarios stochastiques s'observant autour de cette nouvelle trajectoire, il est possible que les *Reduction Rates* ne soient pas nécessairement plus agressifs suite à l'intégration de la pandémie. Ce phénomène s'observe dès lors que la hausse de volatilité ne suffit pas à compenser celle sur la tendance centrale.

Afin d'avoir une vision plus concrète de la mise en place de ces scénarios pour un assureur, on va désormais mesurer le niveau de capital de solvabilité requis, soit l'impact sur les fonds propres d'une dérive de longévité, pour le portefeuille de retraite d'AXA France pour chacun des modèles :

	LC		CBD		BFK alternatif	
	2019	2020	2019	2020	2019	2020
Fonds Propres Central	1 000 000					
Fonds Propres Longévité	822 939	832 142	851 496	855 385	836 045	846 555
SCR Longévité	177 061	167 858	148 504	144 615	163 955	153 445
Évolution 20/19	-5.2%		-2.6%		-6.4%	

TABLE 7.1: Évolution du SCR de Longévité pour la composante tendance par modèle

On constate le phénomène de réduction de l'agressivité des scénarios bicentennaires puisque le SCR de Longévité devient moins important suite à l'intégration de l'année 2020 à l'historique de calibrage. On voit que le modèle de BFK alternatif est celui bénéficiant le plus de cette diminution, signifiant ainsi que la hausse de volatilité est nettement plus faible que le redressement de tendance projetée.

Cette observation contre-intuitive pousse à s'interroger sur la pertinence des nouvelles trajectoires d'améliorations centrales et bicentennaires. La robustesse des nouvelles tendances pour le scénario central sera abordée dans les sections suivantes destinées aux différentes approches possibles de retraitements et ajustements sur ces dernières.

Afin de vérifier la prudence des nouveaux scénarios bicentennaires, et notamment dans le cadre de Solvabilité 2, on va réévaluer les critères sur les améliorations d'espérance de vie afin de s'assurer que les modèles qui étaient auparavant conformes restent valides après intégration de la pandémie au calibrage. Les autres critères ne seront pas recalculés car, par construction, on s'attend à ce qu'ils ne soient pas affectés ou que ce soit négligeable.

	Femmes		Hommes	
	2019	2020	2019	2020
LC	237%	213%	197%	200%
CBD	180%	165%	194%	189%
BFK alternatif	207%	191%	202%	187%
$\epsilon^{(1)}$	170%			

TABLE 7.2: Comparaison du Imp Ratio par modèle

On observe une baisse générale de l'Imp Ratio pour chaque modèle et genre. A l'exception du modèle CBD pour les femmes qui avait déjà montré des signes de non-robustesse, les modèles de LC et BFK alternatif semblent maintenir un niveau de prudence adéquat selon ce critère.

Concernant l'Imp Ratio^{max}, l'intégration de la COVID-19 aux données de calibrage n'impacte pas les conclusions du test. Pour le modèle de CBD, on conserve la sous-prudence pour les femmes et,

	Femmes		Hommes	
	2019	2020	2019	2020
LC	24%	22%	28%	28%
CBD	19%	17%	28%	27%
BFK alternatif	21%	20%	29%	27%
$\epsilon_{inf}^{(2)}$	20%			
$\epsilon_{sup}^{(2)}$	100%			

TABLE 7.3: Comparaison du Imp Ratio^{max} par modèle

pour les modèles de LC et BFK alternatif, le niveau de prudence est suffisant puisque les résultats sont compris entre les bornes $\epsilon_{inf}^{(2)}$ et $\epsilon_{sup}^{(2)}$.

Finalement, les analyses menées orientent naturellement vers le choix du modèle de BFK alternatif pour définir les scénarios bicentennaires de longévité avec prise en compte de la COVID-19 à l'historique de calibrage.

Cependant, une compagnie d'assurance pourrait préférer, d'une part, un modèle plus traditionnel et largement répandu dans le secteur et, d'autre part, l'utilisation d'un modèle unique afin de limiter les variations de capital de solvabilité requis dans le temps. Étant donné que le modèle de BFK alternatif assez peu utilisé en pratique, les sections suivantes se concentreront sur les approches alternatives pour mitiger l'impact de la pandémie sur les scénarios bicentennaires de mortalité. Cela inclura le retraitement des données de mortalité ainsi que l'ajustement des paramètres et de la modélisation.

7.3 Retraitement des données de mortalité

Dans cette section, on va explorer les opportunités de retraitement des données de mortalité liées à la COVID-19. L'approche consiste à conserver les données de 2020 tout en les ajustant, à l'opposé de retraitements totaux ou nuls, pour lesquels on inclurait l'année 2020 sans modification ou on l'omettrait complètement, ce qui reviendrait à reprendre des éléments implicitement évoqués dans la section précédentes.

Le retraitement partiel peut se diviser en deux catégories : le premier, dit "réel", repose sur des données concrètes et dans notre cas, le nombre de décès dus à la COVID-19. Le second, qualifié de "théorique", englobe diverses méthodes basées sur des techniques mathématiques permettant d'estimer le nombre de décès.

On décide d'opter sur le retraitement "théorique", en effet, le retraitement "réel" ne sera pas couvert par ce mémoire puisqu'on considère qu'il sera difficilement justifiable auprès du régulateur ou d'un auditeur. Comme évoqué dans la première partie de ce mémoire, la fiabilité et la granularité des données de mortalité de la pandémie sont souvent remises en question et nécessiteraient des hypothèses supplémentaires sur des bases déjà contestables. De plus, la classification des causes de décès durant cette période fut complexe puisque l'identification de la réelle source de décès d'une personne contaminée fut difficile, dégradant davantage la qualité des mesures.

Dans le contexte de Solvabilité 2, retraiter les décès liés à la pandémie est entièrement justifié puisque le cadre réglementaire inclut un sous-module catastrophe spécifique destiné à couvrir de tels événements, permettant ainsi d'éviter un double comptage et une prudence non nécessaire.

Pour effectuer ce retraitement, notre approche consiste à réévaluer la mortalité de 2020 à l'aide de la projection *Best Estimate* des modèles calibrés sur un historique s'arrêtant à 2019, soit pour chaque modèle et tout $x = x_m, \dots, x_M$:

$$D_{x,2020} = L_{x,2020} \times m_{x,2020}^{BE} = -L_{x,2020} \times \ln(1 - q_{x,2020}^{BE})$$

On obtient alors, pour chaque modèle et chaque âge, le nombre de décès retraits en 2020 suivant :

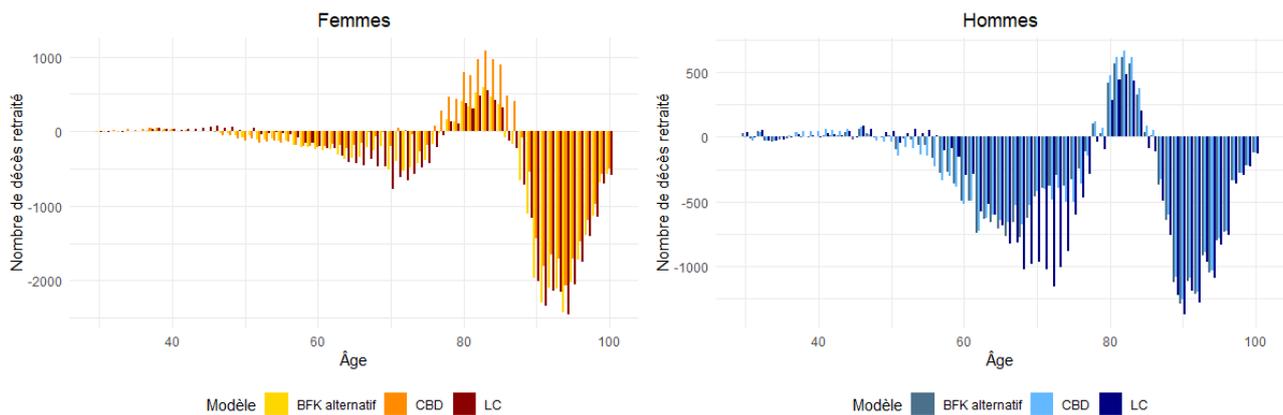


FIGURE 7.2: Impact du retraitement du nombre de décès en 2020 par modèle

Comme attendu, on note un retraitement significatif des décès aux âges élevés et, à l'image de l'impact de la pandémie, plus accentué pour les hommes. Ces retraitements s'élèvent à environ 48 000, 29 000 et 43 000 respectivement pour les modèles de LC, CBD et BFK alternatif, tandis que le nombre reporté en réalité s'approche de 65 000. A l'exception du modèle de CBD pour lequel les résultats sont incohérents, les différences de décès s'expliquent tout d'abord par le fait que nos retraitements se concentrent sur la tranche d'âges 30-100 ans tandis que les statistiques nationales englobent toutes les tranches d'âges. Un autre facteur d'écart réside dans le fait qu'on ajoute des décès attribués à la COVID-19 au voisinage de 80 ans dans chaque modèle, suggérant ainsi un faible impact de la pandémie à ces âges. Cependant, étant donné que la pandémie a affecté toutes les catégories d'âges avancés, cette observation s'attribue à la sur-estimation de la mortalité à la jointure lors de la fermeture de tables, une problématique déjà identifiée lors de l'analyse des SMR. En dehors de ce dernier effet, on peut considérer que la méthode de retraitement semble appropriée.

On dispose donc de nouvelles données pour lesquelles l'année 2020 a été retraits, on peut désormais calibrer nos modèles avec ces données et constater les évolutions sur l'évaluation des paramètres de tendance.

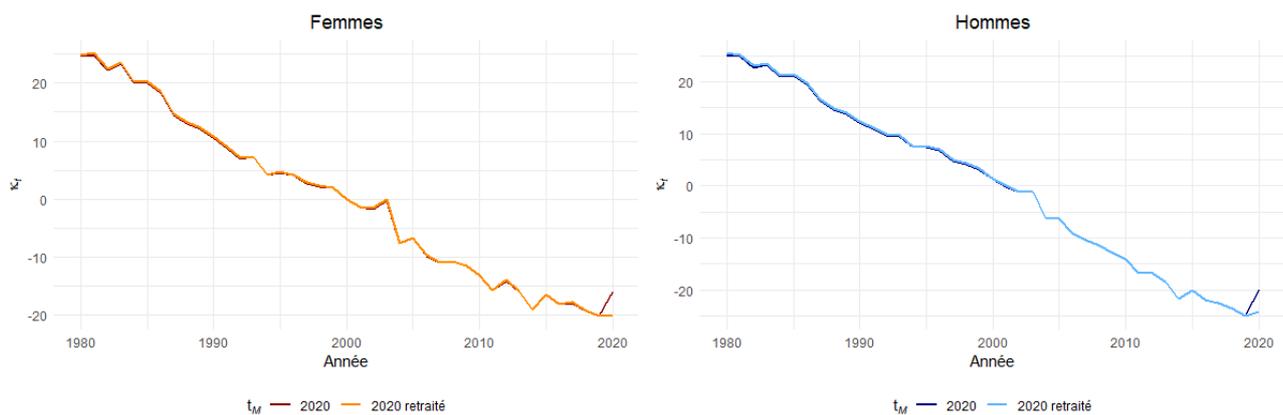


FIGURE 7.3: Impact du retraitement de 2020 sur le paramètre κ_t pour le modèle de LC

Pour les femmes et les hommes, l'observation retraits est davantage en phase avec les tendances passées. Par conséquent, la projection de ce paramètre montrera un redressement de la pente moins accentué et donc une volatilité moins impactée. A noter qu'on conserve une légère inflexion de la tendance, ce qui est pertinent puisque cela permet de conserver un certain degré d'augmentation de volatilité lié à la COVID-19.

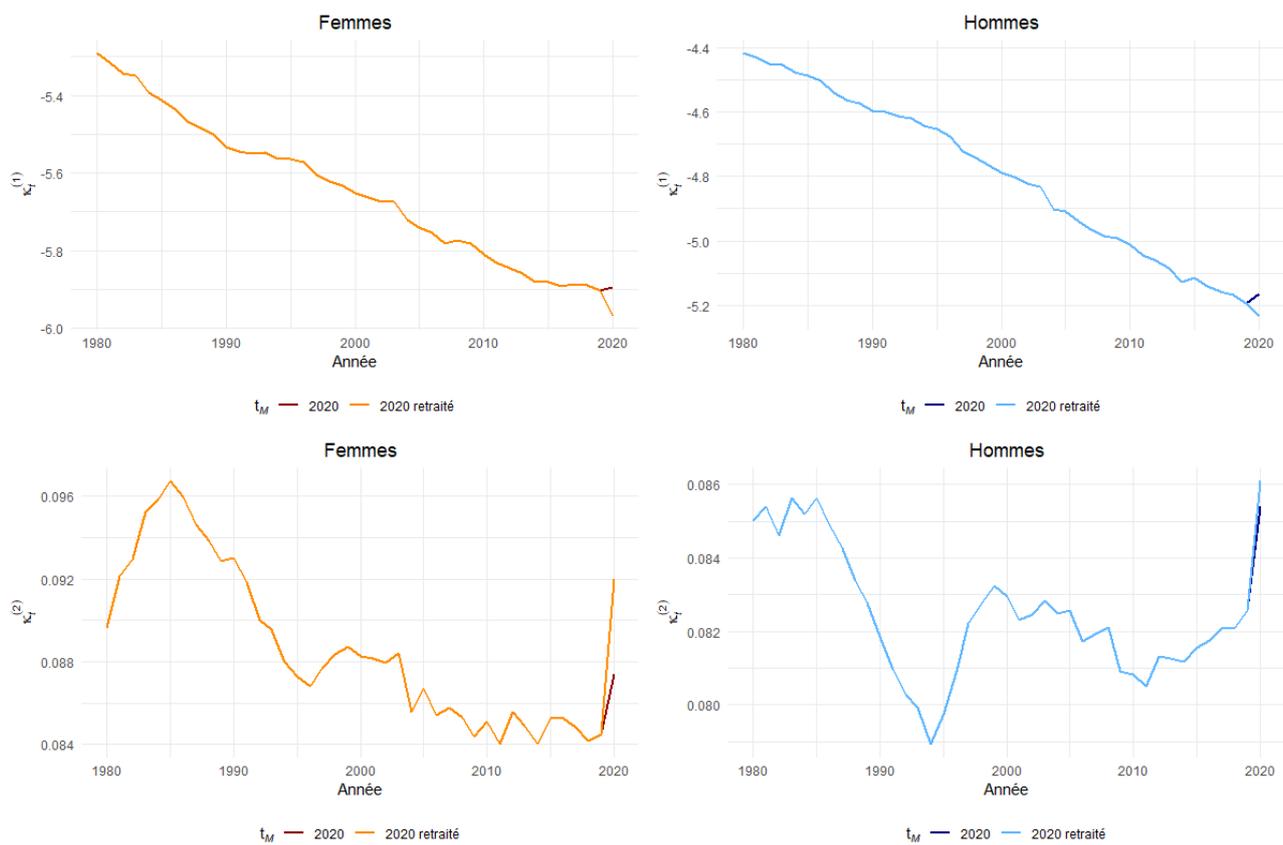


FIGURE 7.4: Impact du retraitement de 2020 sur les paramètres $\kappa_t^{(1)}$ et $\kappa_t^{(2)}$ pour le modèle de CBD

Pour les hommes, le constat est identique à celui effectué sur le modèle de LC, en revanche, pour les femmes le retraitement génère une observation retraitée de $\kappa_t^{(1)}$ symétrique à celle non retraitée par rapport à la tendance historique. Il est préférable de privilégier l'observation non retraitée puisqu'elle suggère une tendance aux améliorations amoindrie, ce qui est en ligne avec les observations de la pandémie. On décide alors d'appliquer le retraitement sur les hommes et de ne pas en tenir compte pour les femmes.



FIGURE 7.5: Impact du retraitement de 2020 sur les paramètres $\kappa_t^{(1)}$, $\kappa_t^{(2)}$ et $\kappa_t^{(3)}$ pour le modèle de BFK alternatif

Pour le modèle de BFK alternatif, les observations sont similaires à celles effectuées sur le modèle de CBD, soit un retraitement globalement cohérent pour les hommes et moins pertinent pour les femmes. En effet, l'observation de $\kappa_t^{(1)}$ retraité indique une nette amélioration des tendances de mortalité en 2020, ce qui est difficilement justifiable compte-tenu de la spécificité de cette année. Ainsi, à l'image du modèle de CBD, on ne retiendra les retraitements que pour les hommes.

Finalement, on décide d'appliquer le retraitement à tous des modèles et genres à l'exception des femmes pour les modèles de CBD et BFK alternatif. Contrairement à l'approche précédente, on ne va pas détailler l'évolution des projections des paramètres de tendance puisque la modification de la pente est simplement prévisible sans nécessiter d'illustration. En effet, lorsque le paramètre de l'observation ajustée est supérieur à celui de l'observation non ajustée, cela entraîne un redressement de la pente, et l'inverse mène à son abaissement. Cela permet maintenant de procéder à l'évaluation et à la comparaison des niveaux de volatilité résultant des différents retraitements :

Modèle	Genre	Année	$h = 10$	$h = 30$
LC	Femmes	2019	167%	183%
		2020	234%	229%
		2020 retraité	168%	190%
	Hommes	2019	130%	129%
		2020	250%	205%
		2020 retraité	155%	148%
CBD	Femmes	2019	1.77%	1.77%
		2020	1.97%	1.96%
		2020 retraité	2.02%	1.85%
	Hommes	2019	1.70%	1.73%
		2020	2.44%	2.21%
		2020 retraité	1.62%	1.69%
BFK alternatif	Femmes	2019	1.91%	3.62%
		2020	1.95%	3.65%
		2020 retraité	2.86%	3.65%
	Hommes	2019	2.29%	3.39%
		2020	2.80%	3.49%
		2020 retraité	2.22%	3.27%

TABLE 7.4: Synthèse des volatilités par pondération, année, genre et modèle

Ce tableau permet de situer la volatilité issue des données ajustées et de la mettre en perspective avec les niveaux de volatilités exposées précédemment. L'objectif est de garantir que la nouvelle volatilité est au minimum égale, sinon supérieure, à celle observée avant prise en compte de la pandémie, soit 2019. Cette approche de fixation de *floor* pour la volatilité découle des incertitudes sur la réévaluation de l'année 2020. Ainsi, la volatilité finalement retenue sera celle maximale entre 2019 et 2020 ajustée. Pour le modèle de LC, la volatilité résultante du calibrage incluant l'année 2020 ajustée est adoptée. Pour les modèles de CBD et BFK alternatif, le retraitement ne s'appliquant pas aux femmes, la volatilité de 2020 non ajustée est donc conservée. Pour les hommes, la volatilité pré-ajustement est privilégiée car elle surpasse celle post-ajustement.

La sélection des retraitements et des volatilités étant établie, il est désormais possible de procéder à la comparaison des scénarios bicentennaires, soit des *Reduction Rates* :

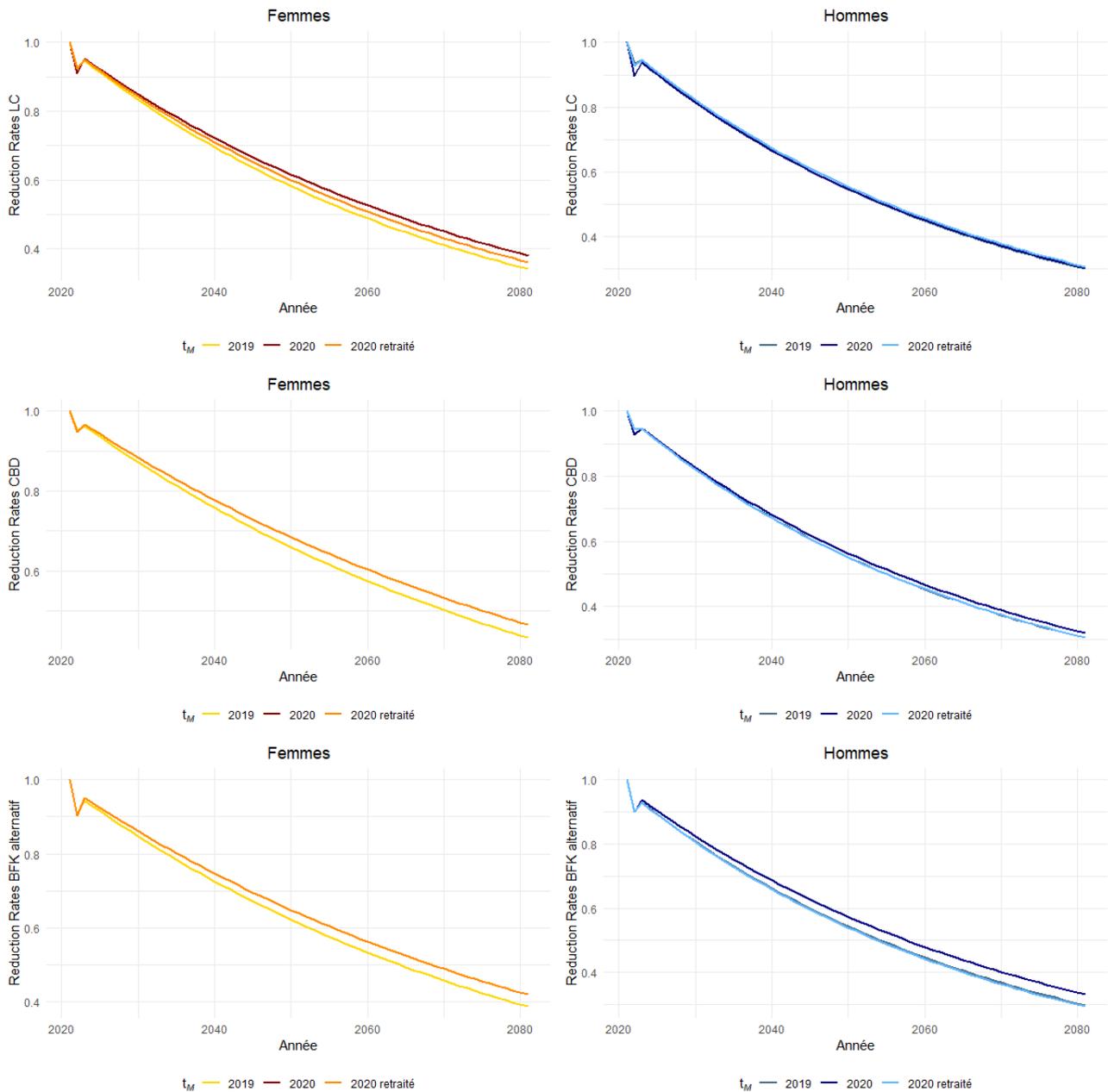


FIGURE 7.6: Comparaison des *Reduction Rates* bicentennaires de longévité moyens par genre pour les modèles de LC, CBD et BFK alternatif

Il est important de préciser que les analyses des graphiques concernant les femmes pour les modèles de CBD et BFK alternatif ne seront pas effectuées étant donné qu'aucun retraitement n'a été appliqué. Pour les autres genres et modèles, les éléments avancés dans le chapitre précédent sur les *Reduction Rates* de 2019 et 2020 sont également applicables à ceux de 2020 retraités, notamment sur la cohérence de la décroissance des scénarios dans le temps et l'observation du pic en début de projection.

Concernant le pic observé, en corrélation avec le niveau de volatilité, il se positionne entre celui de 2019 et 2020, atteignant le niveau de 2019 lorsqu'un *floor* est appliqué. Pour ce qui est des tendances, en prenant comme référence le niveau des *Reduction Rates* à horizon de projection, une stabilité est remarquée chez les hommes pour le modèle de LC, indiquant un équilibre entre la variation de volatilité et celle de la tendance centrale. Pour ce même modèle, mais pour les femmes, la hausse de volatilité est moins prononcée que la hausse de l'ajustement de la tendance, résultant en des scénarios plus

conservateurs avec des données arrêtées à 2019 et plus agressifs lorsqu'on tient compte de 2020 sans retraitement. Une situation homologue est notable chez les hommes pour les modèles de CBD et BFK alternatif, avec un faible écart entre 2019 et 2020 retraité, dû à l'application du *floor* sur la volatilité, entraînant un déséquilibre entre l'augmentation de la volatilité et l'ajustement de la tendance.

Pour mieux appréhender les effets de ces ajustements, on va désormais mesurer le niveau de capital de solvabilité requis, soit l'impact sur les fonds propres d'une dérive de longévité, pour le portefeuille de retraite d'AXA France pour chacun des modèles :

Modèle	Année	Fonds Propres Central	Fonds Propres Longévité	SCR Longévité	Évolution 20/19	Évolution 20/19 retr.
LC	2019	1 000 000	822 939	177 061	-5.2%	-3.1%
	2020		832 142	167 858		
	2020 retr.		828 503	171 497		
CBD	2019		851 496	148 504	-2.6%	-1.5%
	2020		855 385	144 615		
	2020 retr.		853 690	146 310		
BFK alternatif	2019		836 045	163 955	-6.4%	-4.6%
	2020		846 555	153 445		
	2020 retr.		843 561	156 439		

TABLE 7.5: Synthèse des SCR Longévité par année et modèle

Comme énoncé précédemment lors de la comparaison des *Reduction Rates*, les scénarios basés sur les données 2020 retraitées se positionnent entre les deux précédents, soit 2019 et 2020 non retraité. Dans le cas du modèle de LC, la diminution de SCR Longévité est majoritairement attribuable aux femmes, les 3 scénarios de ce modèle pour les hommes étant extrêmement proches. Concernant les modèles de CBD et BFK alternatif, la variation par rapport à 2020 non retraité est exclusivement expliquée par les hommes, les femmes n'étant pas concernées par ces ajustements.

En conclusion, les analyses révèlent que les scénarios résultants des modèles calibrés avec les observations de 2020 non ajustées présentent les niveaux de SCR Longévité les plus bas. Toutefois, ce scénario pourrait être le plus contesté par le régulateur ou un auditeur, car même si l'année atypique ne correspond pas directement à un événement de longévité, incorporer une telle année devrait, au minimum, conduire à des scénarios aussi prudents que ceux évalués sans pandémie. Par conséquent, cette approche pourrait nécessiter l'ajout d'un complément de volatilité, par exemple au travers du paramètre $\sigma_{\text{Additionnel}}$, pour pallier la baisse de prudence résultant du redressement de la tendance générale.

Les scénarios avec ajustement de l'année 2020 demeurent pertinents, évitant l'interruption de l'historique de calibrage dans le temps et conduisant à des évolutions de SCR Longévité modérées. Cependant, l'amélioration de cette approche réside dans la correction du biais de réplcation des erreurs passées, comme observé autour de 80 ans où une surévaluation de la mortalité mène à une sous-estimation des décès dus à la COVID-19, influençant alors tout le processus de projection. Rectifier ces imperfections pourrait se faire par l'affinement de la méthode de fermeture de tables ou par une révision de l'approche d'estimation des décès liés à la COVID-19. Face à ces incertitudes, considérer, à un moindre niveau, un surplus de volatilité $\sigma_{\text{Additionnel}}$ pourrait également s'avérer judicieux pour refléter les risques non mesurés.

Finalement, opter pour un retraitement total, soit l'exclusion des données de 2020, représente l'approche la plus simple à appliquer et la moins susceptible d'être contestée par le régulateur ou un auditeur. Cette méthode évite d'incorporer des données hypothétiques, contrairement aux retraitements discutés dans cette section, et ne nécessite donc pas l'ajout de volatilité supplémentaire. De plus, les scénarios basés sur un historique arrêté à 2019 se montrent les plus conservateurs, générant un SCR Longévité supérieur aux autres options sans pour autant s'éloigner significativement, minimisant ainsi l'impact de cette prudence accrue.

Après avoir démontré la capacité d’absorption des effets de la pandémie par le modèle de BFK alternatif et les options de retraitement des données de calibrage en proposant une méthode d’estimation des décès dus à la COVID-19, on va désormais, pour conclure ce mémoire, évoquer les opportunités d’ajustements des paramètres et de la modélisation afin de rendre les scénarios adaptés dans un contexte marqué par une pandémie.

7.4 Ajustement des paramètres et de la modélisation

Dans cette section finale, on explorera tout d’abord les ajustements possibles sur les paramètres de modélisation, en examinant la sensibilité des modèles aux paramètres de pondération. Ensuite, on conclura en abordant les pistes potentielles permettant d’affiner la modélisation qui n’auront pas été traitées dans cette étude.

Avant d’examiner les améliorations possibles du paramètre de pondération h , il est important de rappeler comment la pondération des observations varie en fonction de la valeur du paramètre :

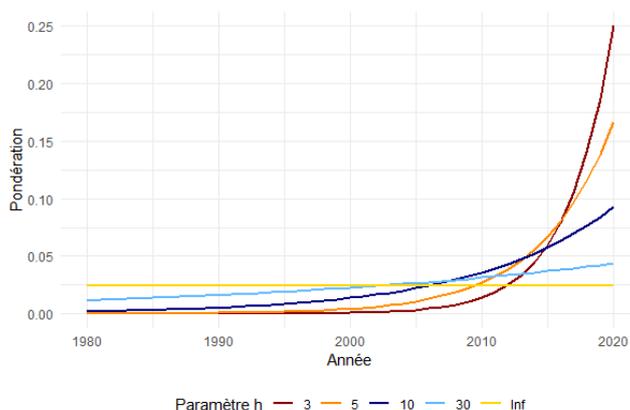


FIGURE 7.7: Comparaison de la pondération des observations en fonction du paramètre de pondération h

La sensibilité de la pondération des observations évolue peu pour des valeurs de h supérieures à 30, surtout en comparaison avec une absence de pondération, c’est-à-dire pour $h = \infty$. Il est donc raisonnable de plafonner le paramètre à ce niveau, d’autant plus que privilégier légèrement les observations récentes est judicieux. En effet, les données les plus récentes ont généralement tendance à être plus représentatives des tendances futures comparativement aux données anciennes. Par ailleurs, la pondération devient très sensible pour $h < 10$, attribuant jusqu’à 25% de la pondération totale à la dernière observation lorsque $h = 3$, tandis que cette proportion atteint environ 10% pour $h = 10$. Par la suite, on va donc examiner l’effet de l’ajustement du paramètre de pondération au travers des différentes phases de la modélisation. Cela inclut l’impact sur la projection de la tendance centrale, sur le calcul de la volatilité et enfin les effets sur la pondération après événement bicentenaire.

On va s’intéresser dans un premier temps au paramètre de pondération servant à évaluer la trajectoire des améliorations pour le scénario central. Le niveau de pondération étant fixé actuellement à $h = 10$, on peut s’interroger sur la pertinence de conserver un tel niveau au regard de l’impact de l’intégration de la COVID-19 sur la trajectoire des améliorations futures. On envisage alors d’augmenter ce paramètre en examinant spécifiquement l’intérêt de le porter à $h = 30$, en effet, diminuer le paramètre ne serait pas cohérent car cela amplifierait l’impact des événements ponctuels tels que ceux observés en 2020.

Il convient de préciser que la pondération à $h = 10$ reste pertinente lors que les données récentes ne sont par fortement impactées par un événement ponctuel non représentatif des tendances futures. On propose donc de réévaluer à la hausse le paramètre de pondération à $h = 30$ dès lors que les données récentes intègrent un événement que l’on ne souhaite pas répliquer, ainsi :

$$h = 10 \text{ si } |\text{Slo}_{10}^{\text{Sans événement}} - \text{Slo}_{10}^{\text{Avec événement}}| < |\text{Slo}_{10}^{\text{Sans événement}} - \text{Slo}_{30}^{\text{Avec événement}}|$$

$$h = 30 \text{ sinon}$$

En d'autres termes, on décide de prendre la pondération de la pente qui s'approche au plus de la pente sans événement indésirable, ce qui orientera fréquemment vers un choix de $h = 30$. Cette orientation restera valable tant que l'effet de l'observation atypique sur la pente reste significatif avec une pondération à $h = 10$, et ce, jusqu'à ce que le temps atténue suffisamment son influence pour que son impact sous $h = 10$ devienne négligeable, soit à l'intersection des courbes de pondération $h = 10$ et $h = 30$. Le changement de paramètre de pondération se répercute de la façon suivante sur la projection déterministe des paramètres de tendance générale de nos modèles, soit κ pour LC et $\kappa^{(1)}$ pour CBD et BFK alternatif :

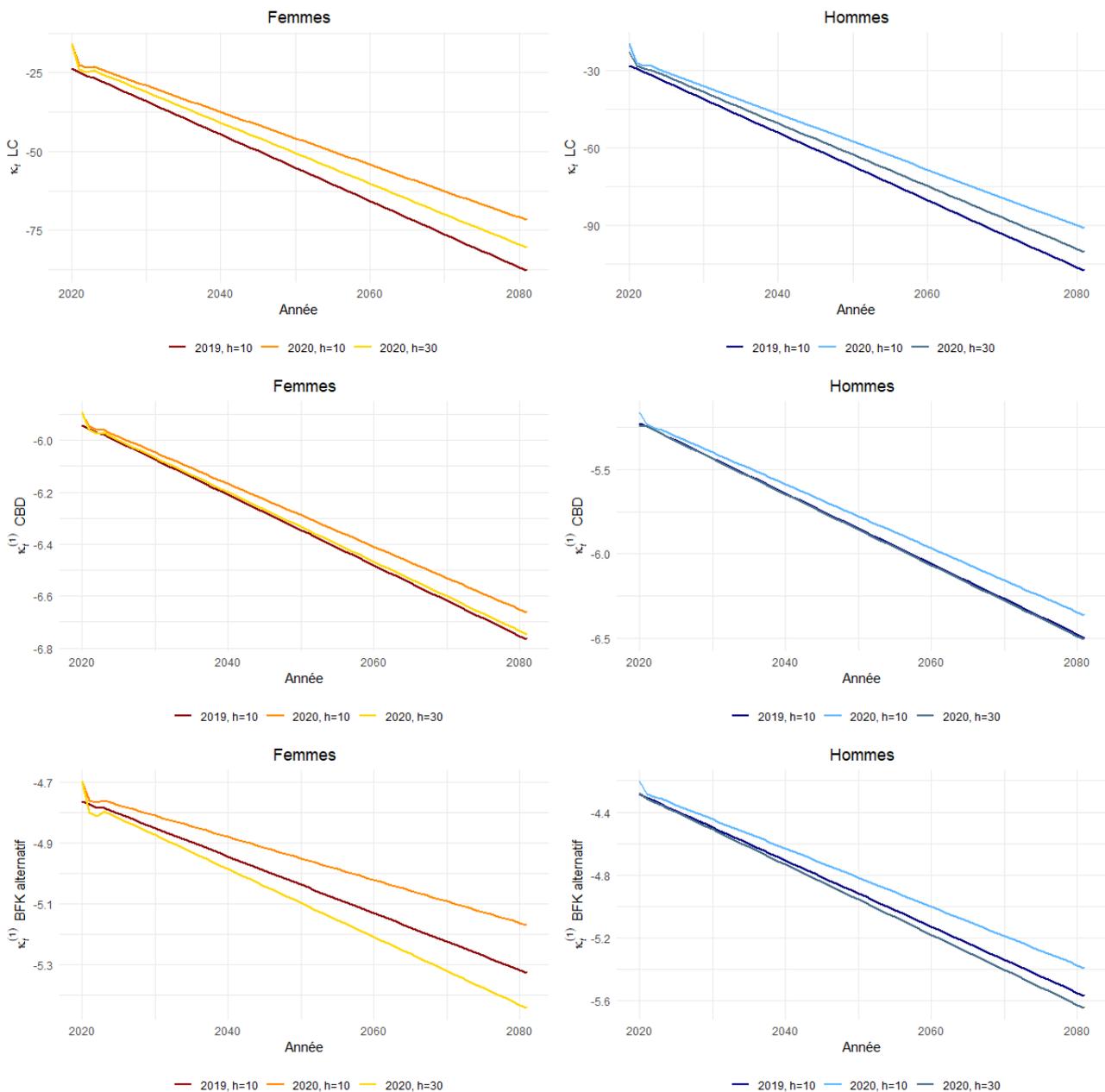


FIGURE 7.8: Projection des paramètres de tendance générale en fonction de la pondération, par genre pour les modèle de LC, CBD et BFK alternatif

Le changement de pondération permet de revenir à une tendance générale projetée pour le scénario central proche de celle avant pandémie. En effet, en ajustant la pondération de $h = 10$ à $h = 30$, on divise par 2 le poids attribué à l'année 2020, ce qui a pour effet de nettement réduire le redressement de la pente. Pour le modèle de BFK alternatif, cette modification permet même d'atteindre des niveaux d'amélioration de la mortalité supérieurs à ceux de 2019, grâce à une plus grande importance accordée aux données historiques, qui montraient des améliorations plus importantes que les observations récentes.

Concernant le paramètre de pondération utilisé dans le calcul de la volatilité, on retient actuellement la volatilité maximale entre une pondération à $h = 10$ et $h = 30$. On pourrait décider d'aligner le paramètre de pondération de la volatilité à celui utilisé pour la projection centrale présenté précédemment. Néanmoins, ces scénarios ayant pour vocation d'évaluer le capital à immobiliser pour faire face à un événement bicentenaire dans une norme prudentielle, on estime qu'il est préférable de conserver une marge de prudence face aux incertitudes et ainsi conserver la volatilité maximale entre les deux paramètres de pondérations. En revanche, il est envisageable de ne pas tenir compte de l'année COVID-19 dans le calcul de la volatilité puisqu'il s'agit d'une pandémie et que celle-ci doit être mesurée dans un module dédié. Par simplification et afin de conserver une prudence suffisante, on n'explorera pas cette option dans cette section.

Enfin, s'agissant du paramètre de pondération après événement bicentenaire, on le fixe délibérément à un faible niveau, $h = 3$, afin de s'assurer que l'événement bicentenaire influe de manière significative sur la trajectoire de tendance projetée, garantissant une prudence suffisante des scénarios résultants. Étant donné que les critères de sélection et de validation des modèles se situent près des limites d'acceptabilité, il serait difficilement envisageable d'augmenter ce paramètre. À l'inverse, le réduire pourrait introduire une couche de prudence excessive et négligerait les observations antérieures. Par exemple, passer de $h = 3$ à $h = 1$ passerait la pondération de l'observation bicentenaire de 25% à 50%, ne laissant que 50% d'influence au reste des observations. On choisit donc de ne pas modifier ce paramètre, la gestion de la prudence restant ajustable au travers du paramètre $\sigma_{\text{Additionnel}}$.

Finalement, sur l'ensemble des paramètres de pondération, on décide de ne modifier que celui sur la projection de la tendance centrale en réhaussant temporairement son niveau afin que l'observation impactée par le pandémie n'influe pas de manière significative sur la tendance projetée. Les autres paramètres restants inchangés afin de conserver un niveau de prudence adéquat, on va désormais mesurer l'impact de cet ajustement sur les *Reduction Rates* obtenus puis celui sur le montant de SCR Longévité résultant :

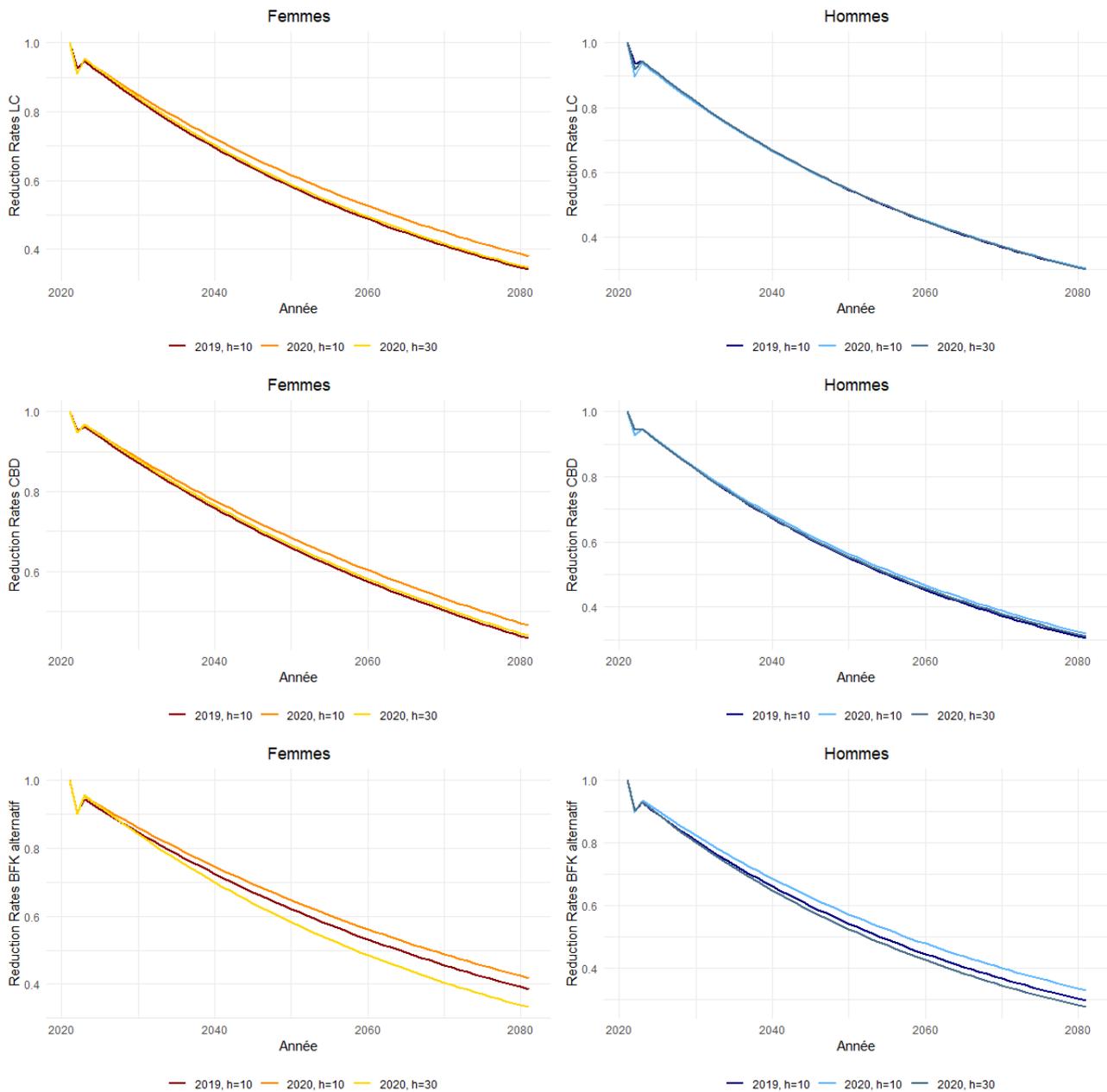


FIGURE 7.9: Comparaison des *Reduction Rates* bicentennaires de longévité moyens par genre pour les modèles de LC, CBD et BFK alternatif

Étant donné qu'aucun ajustement n'est apporté à la pondération de la volatilité, le niveau du pic du début de projection n'évolue pas par rapport à la projection de 2020 sans ajustement. Cependant, pour ce qui est des améliorations à horizon de projection, on retrouve la convergence entre la projection ajustée de 2020 et celle arrêtée à 2019, conformément aux observations précédentes sur la tendance générale. Plus spécifiquement, pour le modèle de BFK alternatif, les *Reductions Rates* se situent sous ceux des deux autres scénarios, indiquant une approche plus prudente.

Toutefois, il est relativement complexe d'évaluer graphiquement l'impact de ces ajustements, on va donc quantifier ces derniers à l'aide du modèle interne d'AXA France appliqué au portefeuille de Retraite :

Modèle	Année	Fonds Propres Central	Fonds Propres Longévité	SCR Longévité	Évolution 20/19	Évolution 20/19, $h = 30$
LC	2019, $h = 10$	1 000 000	822 939	177 061	-5.2%	-2.2%
	2020, $h = 10$		832 142	167 858		
	2020, $h = 30$		826 801	173 199		
CBD	2019, $h = 10$		851 496	148 504	-2.6%	-0.9%
	2020, $h = 10$		855 385	144 615		
	2020, $h = 30$		852 773	147 227		
BFK alternatif	2019, $h = 10$		836 045	163 955	-6.4%	+2.4%
	2020, $h = 10$		846 555	153 445		
	2020, $h = 30$		832 169	167 831		

TABLE 7.6: Synthèse des SCR Longévité par année, pondération et modèle

Les niveaux des SCR Longévité sont cohérents au regard des observations effectuées sur les *Reduction Rates*, illustrant un rapprochement entre les scénarios de 2020 avec une pondération ajustée et ceux excluant la pandémie, ce qui résulte en une diminution de la réduction de SCR. Pour le modèle de BFK alternatif, cette diminution se transforme en une augmentation suite au changement de pondération, indiquant ainsi une hausse de la volatilité supérieure au redressement de la tendance projetée.

Pour conclure sur les opportunités d'ajustement de pondération, on considère que le changement de pondération effectué sur la tendance central est pertinent et plus particulièrement dans le cadre d'une norme prudentielle, il permet d'obtenir des scénarios qui s'approchent davantage de ceux avant COVID-19. Pour le modèle de BFK alternatif en particulier, les ajustements aboutissent à des scénarios encore plus prudents que ceux d'avant pandémie, ce qui est cohérent étant donné les incertitudes introduites par l'année 2020. Il semble raisonnable de retenir des scénarios au moins aussi prudents que ceux d'avant pandémie afin que la hausse de volatilité soit principalement reflétée. Quant aux autres modèles, leur version ajustée présente une pertinence accrue, offrant des scénarios plus prudents par rapport à ceux non ajustés.

Après avoir exposé les bénéfices de l'ajustement de la pondération, on va désormais, pour clore cette section et ce mémoire, aborder les axes d'améliorations envisageables. Ces derniers, non explorés dans le cadre de cette étude, principalement pour des raisons de simplification, pourraient constituer la base des recherches futures pour affiner davantage les résultats :

- **Extension de l'étude à d'autres modèles et méthodes de fermetures de tables :** Les modèles examinés dans cette étude ne constituent qu'une sélection infime de l'ensemble des modèles disponibles dans la littérature. Or, chaque modèle réagit d'une façon unique à l'intégration d'une année atypique, cela soulève donc l'intérêt d'étendre l'analyse à d'autres modèles de référence. Dans notre cas, afin de permettre l'inter-comparabilité, on s'est limité aux modèles sans paramètre de cohorte et utilisé une unique méthode de fermeture de tables, ce qui a introduit certaines incertitudes, particulièrement lors du retraitement des décès liés à la COVID-19. En ce sens, explorer d'autres approches de fermetures de tables et intégrer des modèles avec paramètres de cohorte pourrait offrir des améliorations significatives pour pallier ces incertitudes ;
- **Revue de la méthode de définition des chocs bicentennaires :** La méthode actuelle, qui définit les chocs bicentennaires en se basant exclusivement sur la trajectoire bicentenaire sans considérer les impacts relatifs au scénario central, peut conduire à des résultats paradoxaux, comme une réduction de la prudence des scénarios bicentennaires après intégration de la pandémie. Cela est dû à un ajustement de la tendance centrale, qui affecte également les scénarios bicentennaires, supérieur à la hausse de volatilité. Une approche alternative, qui pourrait être

étudiée afin d'améliorer la définition des chocs, serait de définir des *Reduction Rates* en prenant en compte non seulement la trajectoire bicentenaire mais aussi l'écart par rapport au scénario central.

- **Mise à jour de l'étude avec des données plus récentes** : Dans le cadre de ce mémoire, les données de calibrage s'étendent au maximum à 2020, cependant, la COVID-19 a eu des impacts sur la mortalité sur les années suivantes et notamment en 2021. Par simplification et afin de mieux appréhender la problématique, on s'est volontairement restreints à la prise en compte et au retraitement d'une année. En intégrant les années complémentaires impactées par la pandémie, les constats et approches correctives devraient être semblables. Toutefois, en tenant compte des années additionnelles affectées par la COVID-19, on s'attend à des impacts supérieurs en raison de la hausse du nombre d'années atypiques successives. En guise de complément, il serait donc pertinent de répliquer l'étude en intégrant ces nouvelles données et affiner les ajustements nécessaires.
- **Paramétrage de $\sigma_{Additionnel}$** : Pour des raisons de simplification et afin de faciliter l'interprétation des résultats, ce mémoire exclut volontairement le paramètre de volatilité additionnelle, soit un $\sigma_{Additionnel}$ nul. Ce paramètre a pour vocation d'intégrer toute prudence souhaitée faces aux incertitudes de modélisation résiduelles. Il permettrait notamment de couvrir les lacunes de modélisation, telles qu'évoquées ci-dessus, mais également les écarts de trajectoires d'améliorations entre la population assurée et la population nationale servant de base de calibrage. Une étude dédiée à l'évaluation de ce paramètre et à son impact pourrait s'avérer pertinente pour traiter le sujet exhaustivement.

Dans le contexte de Solvabilité 2, le risque de longévité représente le risque de sur estimation des probabilités de décès des assurés. Les modèles permettant l'évaluation de ce risque se basent sur les données historiques et sont ainsi impactés par l'intégration d'années exceptionnelles au calibrage telles que celles marquées par la pandémie. Ce mémoire propose diverses approches permettant l'intégration des données COVID-19 dans la modélisation afin d'assurer des estimations précises malgré la présence de données perturbées.

L'inclusion des données relatives à la COVID-19 dans le calibrage a modifié les scénarios de longévité bicentennaires. L'impact sur les paramètres liés à l'âge est négligeable, il reste néanmoins significatif pour les paramètres temporels en raison de la déviation notable de la nouvelle donnée en comparaison aux tendances historiques. La projection dans le temps du paramètre de tendance démarrant à partir de cette dernière observation, cela entraîne une translation du paramètre projeté.

Ensuite, la méthode de projection des paramètres temporels consistant en une régression linéaire pondérée, la prise en compte de l'observation liée à la pandémie a pour effet d'ajuster sensiblement à la hausse la pente du paramètre projeté, suggérant une tendance plus faible aux améliorations futures.

Enfin, l'intégration de la pandémie augmente le niveau de volatilité, mesurée par l'erreur quadratique moyenne pondérée, s'expliquant par une forte pondération apportée à une observation nettement éloignée des tendances historiques.

La combinaison de ces trois effets conduit à une réduction de la prudence des scénarios, ce constat contre-intuitif provient d'un rehaussement de la tendance centrale plus prononcé que l'augmentation de volatilité, entraînant des améliorations bicentennaires moins importantes.

Au regard des impacts constatés et de la pertinence limitée de considérer une pandémie dans le calibrage de scénarios de tendance de longévité, le choix s'est tout d'abord porté sur le modèle le plus apte à absorber les conséquences de la pandémie sur la modélisation. Le modèle alternatif de Börger-Fleischer-Kuksin proposé dans ce mémoire s'est révélé être le plus efficace à ce titre, validant l'ensemble des critères de validation. La présence des paramètres $\kappa^{(2)}$ et $\kappa^{(3)}$ permet de n'affecter qu'une portion négligeable des effets de la pandémie au paramètre de tendance générale, résultant en une hausse de volatilité de 7% et un taux d'absorption de 99%. En comparaison, le modèle de Lee-Carter (resp. Cairns-Blake-Dowd) subit une hausse de volatilité 54% (resp. 23%) et un taux d'absorption de 80% (resp. 93%).

La réduction de prudence des scénarios bicentennaires en réponse à l'intégration des données de pandémie est difficilement recevable au vu de la particularité de l'année intégrée. Un retraitement théorique des décès COVID-19 a donc été étudié par la suite, en supposant un nombre de décès en 2020 égal à celui estimé par nos modèles pour cette même année. Pour certains modèles ou genres, les calibrages obtenus après retraitement manquent de cohérence et sont donc écartés. Ces incohérences résultent principalement des lacunes dans la méthode de fermeture de tables, où une sur-estimation de la mortalité au point de jonction entraîne une sous-estimation des décès dus à la COVID-19. Les retraitements retenus améliorent la pertinence des scénarios bicentennaires, cependant, les incertitudes apportées par

les hypothèses utilisées nécessiterait l'ajout d'un complément de volatilité. Étant donné la complexité ajoutée par ce complément, une omission totale des années affectées par la pandémie est suggérée. Cette approche simplificatrice éliminerait le besoin d'introduire un surplus de volatilité permettant de couvrir les hypothèses de retraitements.

Enfin, afin d'affiner la modélisation, la dernière approche proposée consistait à réévaluer l'adéquation et la robustesse des paramètres de pondération dans le contexte de la pandémie. Ainsi, on a opté pour l'adoption d'un paramètre de pondération déterministe de remplacement dès lors qu'un événement influençant à tort la tendance s'intègre aux données de calibrage. Concernant les pondérations de la volatilité et de la tendance après événement bicentenaire, on suggère de les maintenir inchangées, estimant que le niveau de prudence actuel des scénarios bicentenaires est adéquat. L'ajustement du paramètre de pondération déterministe ramène les améliorations de mortalité à des niveaux similaires à ceux observés avant la pandémie et offre des impacts sur le SCR nettement plus cohérents, en particulier pour le modèle de Börger-Fleischer-Kuksin alternatif.

Ce mémoire, portant sur l'intégration des données de mortalité COVID-19 dans la définition des scénarios bicentenaires de longévité, se concentre sur un cadre théorique spécifique. Cela implique donc que certains aspects n'ont pas été explorés et pourraient constituer le fondement de recherches futures. L'étude pourrait notamment être étendue en incluant d'autres modèles ou méthodes de fermeture de tables afin de capter les effets de cohortes ou de réduire la surestimation des décès à la jointure. Par ailleurs, la méthode de définition des chocs ne tenant pas explicitement compte de l'écart relatif entre le scénario central et bicentenaire, une approche alternative serait d'intégrer cet écart, offrant ainsi une meilleure représentation des variations dans la trajectoire centrale. L'étude portant sur l'année 2020 seule ne traite pas la pandémie dans son exhaustivité, néanmoins, les observations et approches identifiées devraient rester applicables. En revanche, les impacts devraient être amplifiés, suggérant ainsi la nécessité d'adapter les analyses futures. Enfin, la dernière piste de complément d'étude concerne le paramétrage de la volatilité additionnelle, permettant de refléter tout risque non capté dans la modélisation actuelle, notamment les différences potentielles entre la population assurée et celle utilisée pour le calibrage.

- [1] ACPR (site consulté en Avril 2023). *Réglementation générale des assurances*. Site de l'ACPR.
- [2] HMD (site consulté en Juillet 2023). *Human Mortality Database*. Site de la HMD.
- [3] Santé Publique France (site consulté en Août 2023). *Coronavirus (COVID-19)*. Site de Santé Publique France.
- [4] Vermet, F. (2021). *Modèles de durée*. EURIA, cours magistral non publié.
- [5] Planchet, F. (site consulté en Juillet 2023). *Modèles de durée, applications actuarielles*. Site de Ressources Actuarielles.
- [6] AXA GRM Life (2022). *Risques biométriques en assurance vie Santé*. ISUP, cours magistral non publié.
- [7] Fromenteau, M., Petauton, P. (2017). *Théorie et pratique de l'assurance-vie*. Édition Dunod.
- [8] Lee, R. D., Carter, L. R. (1992). *Modeling and forecasting U.S. mortality*. Journal of the American Statistical Association.
- [9] Cairns, A. J. G., Blake, D., Dowd, K. (2006). *A two-factor model for stochastic mortality with parameter uncertainty : Theory and calibration*. Journal of Risk and Insurance.
- [10] Börger, M., Fleischer, D., Kuksin, I. (2014). *Modeling the mortality trend under modern solvency regimes*. ASTIN Bulletin : The Journal of the IAA.
- [11] Dobson, A. J. (1990). *An Introduction to Generalized Linear Models*. Chapman and Hall/CRC.
- [12] Denuit, M., Gorniaux, M. (2005). *Closing and projecting lifetables using log-linear models*. Bulletin de l'Association Suisse des Actuaires, sous presse.
- [13] Quashie, N., Denuit, M. (2005). *Modèles d'extrapolation de la mortalité aux grands âges*. UCL, document de travail.
- [14] Planchet, F. (2007). *Tables TGH / TGF 05 : construction*. Conférence de l'Institut des Actuaires, Support de présentation.
- [15] AXA GRM Life (2022). *SII Life technical doc Longevity-Mortality FY22*. AXA Group, documentation interne.

1	Paramètre κ_t par genre pour le modèle de Lee-Carter	5
2	Illustration de l'effet compensatoire pour le modèle de Lee-Carter	6
3	Nombre de décès retraités en 2020 par modèle	7
4	Parameter κ_t by gender for the Lee-Carter model	11
5	Illustration of the compensatory effect for the Lee-Carter model	12
6	Number of deaths suppressed in 2020 by model	13
1.1	Comparaison entre un cycle de production classique et inversé	21
1.2	Piliers de Solvabilité 2	22
1.3	Bilan sous Solvabilité 2	23
1.4	Approche <i>Bottom-Up</i> pour le calcul du SCR	26
1.5	Calcul du SCR d'un sous-module	27
2.1	Évolution de l'espérance de vie à la naissance en France	30
2.2	Causes de décès en 2017 en France	31
2.3	Nombre de cas COVID-19 en France	32
2.4	Nombre de décès COVID-19 en France	33
6.1	Évolution historique de l'espérance de vie à 30, 60 et 80 ans	56
6.2	Distribution des taux de mortalité par âge en 1979, 1999 et 2019	57
6.3	Moyenne sur 5 ans glissants des taux d'amélioration de mortalité par âge en 1980, 2000 et 2020	57
6.4	<i>Heatmap</i> des $SMR_{x,t}$ centrés pour le modèle de LC	60
6.5	<i>Heatmap</i> des $SMR_{x,t}$ centrés pour le modèle de CBD	61
6.6	<i>Heatmap</i> des $SMR_{x,t}$ centrés pour le modèle de BFK alternatif	62
6.7	Paramètre α_x par genre pour le modèle de LC	65
6.8	Paramètre β_x par genre pour le modèle de LC	66
6.9	Paramètre κ_t par genre pour le modèle de LC	66
6.10	Projection du paramètre κ_t par genre pour le modèle de LC	67
6.11	Paramètre $\kappa_t^{(1)}$ par genre pour le modèle de CBD	68
6.12	Paramètre $\kappa_t^{(2)}$ par genre pour le modèle de CBD	68
6.13	Projection des paramètres $\kappa_t^{(1)}$ et $\kappa_t^{(2)}$ par genre pour le modèle de CBD	69
6.14	Paramètre α_x par genre pour le modèle de BFK alternatif	70
6.15	Paramètre $\kappa_t^{(1)}$ par genre pour le modèle de BFK alternatif	71
6.16	Paramètre $\kappa_t^{(2)}$ par genre pour le modèle de BFK alternatif	71
6.17	Paramètre $\kappa_t^{(3)}$ par genre pour le modèle de BFK alternatif	72
6.18	Projection des paramètres $\kappa_t^{(1)}$, $\kappa_t^{(2)}$ et $\kappa_t^{(3)}$ par genre pour le modèle de BFK alternatif	73

6.19	Comparaison des <i>Reduction Rates</i> bicentennaires de longévité moyens par genre pour les modèles de LC, CBD et BFK alternatif	75
6.20	Illustration de la compensation entre hausse de volatilité et redressement de tendance centrale pour les hommes, pour le modèle de LC	76
6.21	Avg Gap $Evol_t$ par genre pour les modèles de LC, CBD et BFK alternatif	77
7.1	Structure simplifiée d'un modèle ALM	80
7.2	Impact du retraitement du nombre de décès en 2020 par modèle	84
7.3	Impact du retraitement de 2020 sur le paramètre κ_t pour le modèle de LC	84
7.4	Impact du retraitement de 2020 sur les paramètres $\kappa_t^{(1)}$ et $\kappa_t^{(2)}$ pour le modèle de CBD	85
7.5	Impact du retraitement de 2020 sur les paramètres $\kappa_t^{(1)}$, $\kappa_t^{(2)}$ et $\kappa_t^{(3)}$ pour le modèle de BFK alternatif	86
7.6	Comparaison des <i>Reduction Rates</i> bicentennaires de longévité moyens par genre pour les modèles de LC, CBD et BFK alternatif	88
7.7	Comparaison de la pondération des observations en fonction du paramètre de pondération h	90
7.8	Projection des paramètres de tendance générale en fonction de la pondération, par genre pour les modèle de LC, CBD et BFK alternatif	91
7.9	Comparaison des <i>Reduction Rates</i> bicentennaires de longévité moyens par genre pour les modèles de LC, CBD et BFK alternatif	93

1	Impact de l'intégration de la COVID-19 sur le SCR de Longévité par modèle	5
2	Impact du retraitement du nombre de décès sur le SCR de Longévité par modèle	7
3	Impact de l'ajustement de la pondération sur le SCR de Longévité par modèle	7
4	Impact of COVID-19 integration on Longevity SCR by model	11
5	Impact of adjusting for the number of deaths on Longevity SCR by model	13
6	Impact of weighting adjustment on Longevity SCR by model	13
3.1	Extension des notations à la dimension temporelle	38
6.1	Corrélation historique des améliorations d'espérance de vie entre femmes et hommes par tranche d'âges	59
6.2	BIC pour les femmes par genre	59
6.3	Imp Ratio par modèle	63
6.4	Imp Ratio ^{max} par modèle	63
6.5	<i>Backtesting</i> pour les femmes par modèle	64
6.6	<i>Backtesting</i> pour les hommes par modèle	64
6.7	Impact sur $\sigma_{Historique,h}$ de l'intégration de 2020 dans le calibrage par genre pour le modèle de LC	67
6.8	Impact sur $\sigma_{Historique,h}$ de l'intégration de 2020 dans le calibrage par genre pour le modèle de CBD	70
6.9	Impact sur $\sigma_{Historique,h}$ de l'intégration de 2020 dans le calibrage par genre pour le modèle de BFK alternatif	74
7.1	Évolution du SCR de Longévité pour la composante tendance par modèle	82
7.2	Comparaison du Imp Ratio par modèle	82
7.3	Comparaison du Imp Ratio ^{max} par modèle	83
7.4	Synthèse des volatilités par pondération, année, genre et modèle	87
7.5	Synthèse des SCR Longévité par année et modèle	89
7.6	Synthèse des SCR Longévité par année, pondération et modèle	94