

Université de Sherbrooke
Faculté d'administration

Quelques résultats sur la modélisation et la couverture
de portefeuilles exposés au risque de longévité

par Frédéric Fontaine

En vue de l'obtention du grade de
Maîtrise en Finance - Maître ès sciences (M.Sc.)

Ce Mémoire est présenté à

Alain BÉLANGER (directeur)

Christian ROBERT

Guy BELLEMARE

Mise en contexte

Les discussions concernant le vieillissement de la population prennent de plus en plus de place dans l'actualité. Ce vieillissement, jumelé à la hausse de l'espérance de vie amène certaines inquiétudes face à la viabilité de certains fonds de retraite. D'autant plus que depuis plusieurs années, les espérances de vie ont été sous-estimées. En parallèle, la démographie des pays développés présente des caractéristiques non-conventionnelles. Plus précisément, au Canada, on observe que le taux de naissance est en deçà du seuil de renouvellement de la population. Ce faible taux, qui perdure depuis plusieurs années, a eu pour effet d'inverser la pyramide des âges. Conséquemment, il est évident qu'éventuellement, il y aura plus de retraités que de gens actifs sur le marché du travail. De plus, il est observé que l'espérance de vie augmente à un rythme important. Considérant que certains régimes de retraites offrent des produits à prestations déterminées, il est évident qu'il y a un risque financier lié à ces régimes qui pourraient éventuellement faire défaut. De plus, ce risque est difficilement assumable par les nouveaux travailleurs qui font leur entrée sur le marché du travail. Cette situation incite les participants du marché financier à trouver une solution à ce problème grandissant. C'est dans cette optique que la présente recherche vise à trouver une solution afin de protéger les parties prenantes face à ces nouveaux risques.

Remerciements

Tout d'abord, je tiens à remercier mon directeur de recherche, Alain Bélanger, pour son temps, sa disponibilité et surtout son soutien tout au long de mon parcours depuis la fin de mon Baccalauréat. Sa grande expérience et son éthique de travail m'ont appris à être plus rigoureux et je lui en serai toujours reconnaissant. Il est, depuis mon premier cours avec lui, un modèle d'excellence et un exemple à suivre.

Je tiens également à remercier Mme Isabelle Archambault ainsi que son équipe pour leur soutien moral tout au long de ma Maîtrise. Isabelle m'a toujours rappelé que le travail fini par rapporter. Elle s'est toujours battue pour que je puisse avoir un emploi afin de m'éviter un stress financier lors de mon parcours au deuxième cycle. Également, elle a pris les bonnes décisions aux bons moments. Je ne pourrai jamais la remercier suffisamment pour tout ce qu'elle a fait pour moi.

Je tiens aussi à remercier Addenda Capital ainsi que Christian Robert pour le stage offert afin de compléter le mémoire ainsi que leurs judicieux conseils. Leur environnement de travail m'a permis d'être le plus efficace possible.

Un remerciement spécial aux membres de ma famille qui ont été présents tout au long de mes études pour me soutenir du mieux qu'ils le pouvaient. Leur présence a été essentielle à ma réussite.

Table des matières

1	Introduction	3
2	Objectifs du mémoire et contributions	6
3	Études antérieures	7
3.1	Historique	7
3.2	Approches possibles	8
3.2.1	Les anticipations	8
3.2.2	La modélisation stochastique	9
3.2.3	Modèles explicatifs	12
3.3	Utilisation pratique des modèles	13
3.4	Théories utilisées	14
3.4.1	Tests d'hypothèses	14
3.4.2	Sélection d'un modèle	16
3.4.3	Algèbre linéaire	17
3.4.4	Analyse par composantes principales	19
3.4.5	Mesures de risque	20
4	Données	23
5	Cadre méthodologique	27
5.1	Modèle utilisé	27
5.2	Calibration du modèle	29
5.2.1	Paramètres attendus	29
5.3	Fenêtre de temps sélectionnée	30
5.3.1	Résultats de la fenêtre choisie	32
5.4	Résultats de la calibration	33
5.5	Décomposition des résidus	36
5.5.1	Analyse des résidus	37
5.6	Validité du modèle	40

5.7	Analyse par composantes principales	41
5.8	Prévisions hors-échantillon	41
5.9	Sensibilités	44
5.10	Lissage des données	44
5.11	Ajustement pour le risque	45
6	Portefeuille simple cohorte	47
6.1	Portefeuille de base	47
6.1.1	Obligation de longévité	48
6.2	Calibration pour le risque	49
6.3	Le portefeuille et ses couvertures	50
6.3.1	Les S-forwards/S-Swaps	53
6.3.2	Les caps de longévité	55
6.4	Résultats et discussion	56
7	Portefeuille multiples cohortes	60
7.1	Mise en situation	60
7.2	Résultats et discussion	61
	Conclusion	65
	Bibliographie	67

1. Introduction

Depuis plusieurs années, le vieillissement de la population est devenu une préoccupation de plus en plus grande pour les fonds de retraite. En effet, l'augmentation de l'espérance de vie de la population au-delà des prévisions actuariales, jumelé à de faibles taux d'intérêt affectent les liquidités de ces fonds. Si la situation venait à empirer, c'est-à-dire que les taux de survie augmentaient davantage au cours des années à venir, ces régimes de retraite pourraient se retrouver avec des difficultés à rester solvable et ainsi manquer à leurs engagements envers leurs rentiers. Cette situation pourrait créer de graves problèmes économiques considérant que les rentiers comptent sur ce revenu afin de subvenir à leurs besoins lors de leur retraite. Les statistiques démographiques démontrent l'ampleur du problème. Au Canada, l'espérance de vie à la naissance est passée de 69 ans en 1970 à 79 ans en 2009. En 2015, cette espérance était estimée à 81 ans. De ce fait, il est important d'apporter des changements au niveau social afin de protéger les parties prenantes exposées aux risques financiers que présente cette situation. Faute de changement, les futurs retraités devront accumuler plus de capital lors de leur passage sur le marché du travail pour atténuer le risque de leur propre longévité.

Booth et Tickle (2008) [8] définissent le risque de longévité comme le risque financier dû à une sous-estimation de l'espérance de vie comparative-ment à la durée de vie réelle. Chaque personne est soumise à ce risque lors de sa retraite, puisque les fonds accumulés lors de leur passage sur le marché du travail sont limités. Conséquemment, cette personne, ne sachant pas le nombre d'années lui restant à vivre, court un risque financier lié à une sous-estimation de ce nombre. Dans le cas d'un portefeuille de rentes viagères, une sous-estimation de l'espérance de vie du portefeuille pourrait entraîner d'importants problèmes de liquidité puisque le passif actuariel est estimé en fonction des probabilités de survie. Actuellement, le risque de longévité,

contrairement à d'autres risques présents sur les marchés financiers, est difficile à diversifier pour les compagnies d'assurance. Cependant, il est observable sur les marchés OTC (over-the-counter) que de plus en plus de contrats en lien avec le taux de survie sont transigés ce qui semble être le début d'un changement.

Le risque de longévité est une importante source de risque au niveau des rentes viagères et des fonds de retraite. Conséquemment, il est important de bien modéliser l'espérance de vie future. Or, Dahl (2005) [11] a montré que les prévisions d'espérance de vie estimées à partir de modèles déterministes ont été sous-estimées par le passé. De ce fait, ils sous-estiment les dépenses futures et ainsi surestiment leur solvabilité future. Bien qu'utile, les modèles déterministes ne permettent pas la génération de chocs aléatoires. Afin de pallier à cette faiblesse, les firmes d'assurances ont accès à d'autres outils tels que les modèles stochastiques. À l'aide de ceux-ci, les effets et les conséquences d'un choc important peuvent être pris en compte dans l'évolution de la variable modélisée (Lee-Carter 1992) [18].

Afin de couvrir les portefeuilles de rentes, ce mémoire s'inspire de deux produits, existant dans le marché des produits dérivés, qui sont principalement utilisés pour protéger des portefeuilles d'actions ou d'obligations. Ces produits sont les options d'achat et les forwards. Les contrats forwards ont une valeur nulle à l'initiation alors que, les options d'achat ont une valeur positive à l'initiation. Un modèle couramment cité dans la littérature et utilisé en pratique au niveau de la tarification d'options est celui de Black-Scholes-Merton (BSM) (1973) [6]. Ce mémoire utilise une variante de ce modèle afin de tarifier les options qu'il propose. Cette variante est utilisée afin de déterminer s'il est possible de couvrir un portefeuille de rentes viagères contre le risque de longévité à l'aide d'un produit simple et standard. Il se penche principalement sur l'efficacité d'une couverture dans le cas d'un portefeuille constitué d'une seule cohorte (simple cohorte) et dans le cas d'un portefeuille constitué de plusieurs cohortes (multi-cohortes).

Une cohorte est définie comme un ensemble de personnes étant nées au cours de la même année. Dans notre cas, nous observons les taux de mortalité annuels de cohortes pré-déterminées.

L'étude du portefeuille simple cohorte constitue une première étape afin de déterminer si les produits agissent efficacement. L'étude du portefeuille

multi-cohortes est un apport au niveau de la littérature scientifique existante à ce sujet. Au niveau pratique, il permet de rapprocher le portefeuille de la réalité des portefeuilles de rentes existants.

Le travail est structuré comme suit : La section 2 présente les objectifs détaillés et les contributions pratiques. La section 3 présente la revue de littérature qui introduit les différentes approches possibles au niveau de la prévision de taux de mortalité, soit les anticipations, la modélisation stochastique et les modèles explicatifs. Dans cette section, il sera aussi discuté de l'utilisation pratique de ces modèles. La section 4 présente les données utilisées. La section 5 présente le cadre méthodologique. La section 6 présente les résultats de la calibration du modèle ainsi que des simulations de portefeuille à une seule cohorte. La section 7 présente les résultats des simulations d'un portefeuille contenant plusieurs cohortes. La section 8 conclut ce mémoire.

2. Objectifs du mémoire et contributions

L'objectif principal du mémoire est d'étudier la couverture de portefeuilles étant exposés au risque de longévité comme un portefeuille de rentes viagères. Cet objectif se décompose en deux sous-objectifs. Le premier est de modéliser les taux de mortalité à l'aide d'un modèle permettant une tarification pour le risque. Le deuxième est de simuler des décès dans un portefeuille de rentes pour ensuite évaluer leur rentabilité ainsi que le risque associé à ce portefeuille.

Les contributions de ce mémoire sont de l'ordre scientifique et pratique. Au niveau scientifique, la contribution de ce mémoire est de se concentrer sur un portefeuille composé de différentes cohortes d'âges dont le processus stochastique des taux de mortalité est donné par un unique modèle. Puisqu'il existe peu de littérature à ce sujet, notre étude au niveau Canadien permet une avancée dans ce domaine. De plus, il ouvre la porte à l'étude du risque de base de la population. Au niveau pratique, la contribution principale est de justifier le besoin d'un marché central de produits dérivés sur la longévité au Canada. Ce qui permettrait la transparence au niveau des transactions et éviterait de surcharger une entreprise avec du risque de longévité. On remarque qu'au niveau mondial, les transactions sur la longévité sont en croissance depuis quelques années [4]. Ceci démontre l'intérêt et la volonté des participants du marché de se protéger contre ce risque. D'autre part, un tel marché attirerait sans doute des investisseurs souhaitant diversifier leur portefeuille avec des produits ayant peu de corrélation avec le marché des actions et celui des obligations. Au final, cela permettrait de transférer le risque qui est présentement assumé majoritairement par les fonds de retraite et les assureurs. En développant un marché

de longévité, la société profiterait alors d'une diversification de ce risque.

3. Études antérieures

Ce mémoire couvre un aspect spécifique de la finance et de la science actuarielle. Afin de dresser un portrait global du sujet, il est important de mentionner les études ainsi que les résultats les plus importants qui ont été publiés par le passé. Puisque la prévision de la mortalité est un sujet datant de plusieurs années, nous avons choisi de dresser un portrait historique des recherches effectuées, des modèles prévisionnels existants ainsi que de l'utilisation pratique de ceux-ci.

3.1 Historique

Les techniques de prévision des taux de mortalité ont un long historique. L'une des premières idées en lien avec ce sujet est la loi de la mortalité de Gompertz [13] qu'il a publié en 1825. Cette loi suggère que la force de la mortalité augmente de façon exponentielle avec l'âge. Malgré cette date lointaine, ce n'est que depuis tout récemment que les modèles de prévisions se sont raffinés. Le vieillissement de la population, observé depuis plusieurs années dans les pays développés, a amené ce sujet au coeur de l'actualité. Ceci a poussé plusieurs chercheurs à se pencher davantage sur les problèmes potentiels que les caractéristiques de la démographie actuelle pourraient apporter. L'une des premières étapes pour déterminer leurs ampleurs est d'être en mesure de prévoir ces taux adéquatement. Malgré le peu de recherche ayant été faites entre les années 1825 et 1970, quelques évidences historiques, au niveau des prévisions de taux de mortalité, ont été démontrées au cours de cette période. En effet, les prévisions de taux de mortalité sont plus précises lorsque : l'horizon de prévision est plus court, l'échantillon étudié est plus grand et que ces prévisions sont faites pour des âges intermédiaires comparativement aux âges les plus jeunes et aux plus vieux (Alder & De Beer, 2005) [1]. Comme les taux de mortalité sont définis à partir du nombre de décès observés, il semble naturel de vouloir prévoir ce nombre. Cependant, Booth & Tickle (2008) [8] ont montré qu'il est plus

précis de prévoir les taux de mortalité directement à partir des données historiques de la populations plutôt que de prévoir le nombre de décès et de calculer les taux de mortalité par la suite. Ceci semble raisonnable comme conclusion considérant que lorsque les décès sont prévus, il faut aussi estimer la population vivante au milieu de cette année. Ce qui constitue une deuxième source d'erreur. Source que nous évitons en modélisant les taux de mortalité directement. De plus, pour 2 populations différentes, le nombre de décès peut augmenter considérablement si l'une d'entre elle est beaucoup plus grande. Pour ces raisons, nous allons nous concentrer sur la prévision de taux de mortalité.

3.2 Approches possibles

Il existe 3 approches utilisées en pratique en ce qui concerne les prévisions de taux de mortalité. Nous expliciterons chacune d'entre elles en montrant les points forts et les points faibles tout en donnant quelques exemples d'utilisation de chacune d'entre elles en pratique.

3.2.1 Les anticipations

Cette approche, qui est la plus simple, prend la forme de l'avis d'un expert ou d'un comité d'expert. L'approche utilise certains arguments de ces experts pour déterminer l'espérance de vie dans un certain nombre d'années d'une cohorte d'un certain âge. Ensuite, les taux de mortalité futurs sont déterminés en conséquence de cette prévision tout en conservant un aspect raisonnable. Il va s'en dire que cette technique est l'une des plus utilisées par les agences statistiques (Waldron 2005) [28] à cause de sa simplicité. Comme les actuaires utilisent toujours cette approche de nos jours, il est primordial de lui accorder une certaine importance afin de trouver les points forts ainsi que ceux à améliorer. Depuis quelques années, les agences ainsi que les actuaires ont commencé à utiliser une technique hybride. C'est-à-dire qu'ils utilisent les modèles comme référent et ajustent les améliorations de taux de mortalité en fonction de leur jugement (Hollmann et al., 2000) [14]. L'un des avantages liés à ce type de modélisation est que l'on peut inclure les caractéristiques démographiques et contemporaines (ex: une épidémie) de manière qualitative qui pourraient être négligées par les modèles stochastiques. Le désavantage de cette méthodologie est qu'elle laisse place à un

biais en ce qui à trait aux estimations. Conséquemment, les améliorations des taux de mortalité sont souvent surestimées ce qui entraîne une sous-estimation du taux de survie et donc une surestimation de la solvabilité du portefeuille (Booth & Tickle, 2008) [8]. De plus, les modèles généralement utilisés sont des modèles de moyenne. À première vue, il semble évident que des techniques plus sophistiquées pourraient diminuer voire éliminer ce biais induit par la subjectivité de l'analyse et ainsi améliorer les prévisions de variation des taux de mortalité.

3.2.2 La modélisation stochastique

Les techniques issues de la branche des statistiques en mathématique peuvent aider à la prévision des taux de mortalité. En effet, elles sont utilisées afin d'exprimer les taux de mortalité futurs comme une fonction des taux passés. L'hypothèse de base de ce type de modèle est que la tendance des taux de mortalité dans le temps est constante et donc que celle future sera la même que celle dans le passé. Cependant, (Li 2015) [19] a montré qu'il est important de bien cibler le nombre d'années de données utilisées puisque cette hypothèse n'est pas respectée. Si la période retenue est trop longue, Li (2015) a démontré que cette hypothèse s'avère fautive. Ces modèles ont généralement de la difficulté à prévoir des taux de mortalités haussiers ce qui peut poser problème lors d'une épidémie. C'est pour cette raison que les anticipations des experts peuvent s'avérer utiles dans ce contexte.

Lee-Carter et ses extensions

L'un des modèles de base en matière de prévision de taux de mortalité est celui de Lee-Carter (1992). Il fût le premier à incorporer l'effet cohorte dans un modèle. Son modèle, qui suppose que les taux sont log-linéaires, est souvent considéré comme un standard en matière de prévision de taux de mortalité. Le modèle se décrit comme suit :

$$m_x(t) = e^{a_x + b_x k_t + \epsilon_x(t)} \quad (3.1)$$

Où $m_x(t)$ est le taux central de mortalité à l'âge x et au temps t , a_x est la moyenne temporelle du log taux de mortalité à l'âge x , b_x est le facteur de changement de la mortalité dans le temps pour une cohorte d'âge x , k_t représente le niveau général de la mortalité au temps t et $\epsilon_x(t)$ est la

portion résiduelle du modèle. L'estimation des paramètres b et k se fait par la décomposition en valeur singulière avec les contraintes suivantes: la somme des paramètres b_x estimés est de 1 et la somme des k_t estimés est de 0 (Lee Carter 1992) [18] (Zahra Boutaggount 2017) [35].

Les extensions du modèle de Lee-Carter sont nombreuses. La première est celle de Wilmoth (1993) [32]. Il a proposé d'estimer les paramètres du modèle de Lee-Carter en utilisant la méthode des moindres carrés pondérés. Cette technique offre une bonne modélisation pour les taux de mortalité japonais. D'autres méthodes d'estimation des paramètres furent proposées entre autre par Wilmoth (1993) et Alho (2000). Ceux-ci ont proposé de les estimer à l'aide du maximum de vraisemblance. Cette approche se base sur l'approximation du nombre de décès observés qui lui, selon l'hypothèse du modèle, suit une loi de Poisson. Le modèle est le suivant :

$$D_x(t) \sim \text{Poisson}(m_x(t)E_x(t))$$

Où

$$m_x(t) = e^{a_x + b_x k_t}$$

Et $E_x(t)$ représente l'exposition au risque de mortalité. Elle peut être estimée à l'aide des carrés de Lexis. La méthodologie de cette estimation est présentée dans l'essai de Joanis (2016) [15]. La force de ce modèle est l'hypothèse d'un taux de mortalité suivant un processus de Poisson inhomogène. Ce type de processus est généralement utilisé afin d'ajouter du réalisme aux processus de Poisson homogènes. Les intensités d'occurrence sont variables dans le cas du processus inhomogène. Dans le cas de la mortalité, il est clair que la probabilité de décès augmente avec l'âge, il est donc plus juste d'utiliser les processus de Poisson inhomogène. L'utilité de ce type de processus est qu'il permet de laisser tomber l'hypothèse implicite de la variance constante dans le temps.

Une autre estimation possible des paramètres fût proposée par Booth et al (2002) [7]. Cette estimation se base sur l'incorporation de termes additionnels dans la décomposition en valeur singulière de la calibration de Lee-Carter. Cela a pour effet d'augmenter l'effet aléatoire dans les résidus (Booth et al. 2002) [7]. Par contre, cela aura aussi pour effet de faire augmenter la variance des prévisions.

Critiques du modèle de Lee-Carter

Le modèle de Lee-Carter, bien qu'il soit considéré comme un standard, a fait l'objet de critiques notamment au niveau des hypothèses sur lesquelles il repose. L'hypothèse la plus critiquée est celle stipulant que les variations de taux de mortalité demeurent constantes à travers le temps. Quelques études l'ont ouvertement critiquée (Lee 2000 et Gutterman and Vanderhoof 1999). Deux solutions ont été proposées par Lee(2000) [17]. L'une d'entre elles est de réduire la fenêtre de temps utilisée pour calibrer le modèle. Ceci permet de mieux représenter la courbe contemporaine des taux de mortalité. La deuxième est de poser a_x égal à la valeur du dernier log taux de mortalité observé. Selon Lee (2000) [17], en utilisant cette technique, on s'assure que la première année de prévision se rapprochera des dernières observations.

Après Lee-Carter

L'avancée amenée par le modèle de Lee-Carter a permis à plusieurs autres auteurs de tester différents modèles incluant l'effet cohorte. Wang & al(2015) [29] font partie de ceux qui ont introduit l'interdépendance des taux de mortalité entre les pays. Ils ont utilisé le même modèle que Lee-Carter et l'ont appliqué à la population de multiples pays simultanément. D'autres modélisations ont été faites pour estimer l'espérance de vie d'un individu. White (2002) [31] a montré qu'il est possible d'obtenir un bon estimé de l'espérance de vie d'un individu à l'aide d'une régression linéaire. Le désavantage de cette technique est la difficulté de tarifier un produit financier dont les paiements se basent sur les taux de mortalité prévus. Du côté de Lee-Carter, la difficulté repose sur l'hypothèse que la tendance des log taux de mortalité est constante à travers le temps. Caractéristique que plusieurs populations n'ont pas (Yue & al., 2008) [34]. Mitchell & al. (2013) [23] explicitent le fait que le modèle surestime la dépendance entre les groupes d'âge en raison de sa construction. Ils proposent alors une modélisation des log-améliorations des taux de mortalité. Cette avancée s'avère importante car les modèles qui ont suivis ont été basés sur cette idée. D'autres types de modèles ont été proposés, tel que celui de Bell (1997) [5] qui a combiné une droite sur les log-taux de mortalité avec un mouvement Brownien s'est avéré au moins aussi efficace que le modèle de Lee-Carter. Cependant, les prévisions à long terme de ce modèle peuvent entraîner des contradictions, telles que des taux de mortalité plus élevés pour des cohortes plus jeunes, ce qui

n'est pas souhaitable. Par conséquent, nous croyons raisonnable d'utiliser un modèle qui se base sur les variations de taux de mortalité puisqu'ils nous permettent d'obtenir des prévisions à chaque année et ces variations sont relativement stationnaires. Nous évitons donc les problèmes liés à la modélisation de Lee-Carter.

Les modèles âge-période-cohorte

Un sous-type de modèle stochastique qui existe est celui des modèles considérant les facteurs suivants: l'âge, la période de temps étudiée ainsi que l'effet cohorte. Les modèles âge-période-cohorte (APC) sont devenus populaires au cours des dernières années. Cependant, ces modèles sont très exigeants en terme du nombre de données requises pour en faire la calibration. Si l'on veut étudier les taux de mortalité pour une cohorte de 0 à 100 ans, alors un siècle de données ne constituera qu'une seule cohorte. Puisque cette quantité de données est accessible pour quelques pays européens, il a été possible pour Tableau et al. (2001) [26] de calibrer un modèle de ce type. Le manque de données peut être un problème majeur lorsque l'on veut employer ce type de modèle. Un autre problème, qui est amené par ces modèles à trois facteurs, est que n'importe quel de ceux-ci est linéairement dépendant des 2 autres (Tableau 2001) [26]. Il est aussi possible de considérer des cohortes de 1 an. Ce sous-ensemble de la famille de modèles des APC se nomme les modèles âge-période (AP). Ils sont beaucoup moins exigeants en terme de données et offrent généralement une bonne modélisation des taux de mortalité. Les modèles de Wills-Sherris (2011) [33] et Fung et al (2015) [12] en font partie. Ce sont ceux que nous utilisons dans ce mémoire.

3.2.3 Modèles explicatifs

Les modèles, de type explicatif, se basent sur la structure ou sur l'effet causal des causes de la mortalité. Plus explicitement, à défaut de modéliser le taux de mortalité à partir de ses variations, cette méthode tentera d'établir une relation entre celui-ci et les causes de décès. Selon Booth et Tickle (2008) [8], cette méthode tarde à être complètement développée car les relations de cause à effet ne sont pas parfaitement comprises. Ce qui rend les prévisions, à long terme, effectuées à l'aide de cette méthode, peu fiables. Un

modèle qui est digne de mention par contre est celui de Vartiainen et al (1994)[27]. Ces auteurs ont démontré que la diminution des maladies cardiaques, expliquée par 3 facteurs déterminés au départ, explique en grande partie l'amélioration des taux de mortalité en Finlande au cours des années 1970. Pour ce faire, ils ont utilisé un modèle de type logit. Malgré la démonstration de son efficacité explicative, les auteurs mentionnent que ces deux diminutions pourraient être expliquées par d'autres facteurs exogènes que les trois proposés. De plus, dans le cas d'une personne ayant plusieurs maladies, il n'est pas évident de déterminer la cause exact du décès. Cette difficulté peut induire certains biais dans les données. Pour ces raisons, nous croyons qu'il est préférable de ne pas poursuivre l'étude ce type de modèle dans le cadre de notre travail.

3.3 Utilisation pratique des modèles

Les modèles se sont ensuite orientés vers la tarification de produits étant exposés au risque de longévité. L'exemple le plus simple de ce type de produit est la police d'assurance vie. Puisque le prix de ce type d'assurance varie en fonction de l'espérance de vie de l'assuré, il va de soi que le prix demandé est fonction des taux de mortalité. Ces produits sont moins dispendieux aux assurés ayant une plus grande espérance de vie puisqu'ils devront payer la prime annuelle pendant un plus grand nombre d'années. Un autre exemple de tarification utilisant les taux de mortalité est celui d'une rente viagère. L'acheteur d'une rente reçoit un montant d'argent fixe à la fin de chacune des années et ce jusqu'à celle du décès. Celles-ci ont une exposition inverse à celle des assurances vie. En effet, plus l'assuré a une espérance de vie élevée plus la rente annuelle payée par la compagnie d'assurance sera payée longtemps. Les modèles cités précédemment, sont difficiles à calibrer afin de tenir compte d'une prime de risque. Une des propositions pour remédier à ce problème est venue de Wang en 2000 [30]. Il a proposé une classe d'opérateurs de distorsion modifiant la fonction décumulative des pertes. Celle-ci est utilisée pour déterminer la perte espérée d'un produit financier. Cette méthodologie fût énormément utilisée dans d'autres articles. Cependant, Pelsser (2008) [24] a démontré que le cadre proposé pour la tarification n'est pas universel ce qui remet en question sa fiabilité. De plus, il existe des possibilités d'arbitrage dans ce cadre. Ces deux raisons nous obligent à le laisser tomber.

Un des modèles qui semble utile pour la tarification des produits de longévité est celui de Wills-Sherris (2010) [33]. Celui-ci se rapproche du modèle de Black-Scholes-Merton (1973) [6] qui est très utile dans la tarification de produits dérivés. Le modèle de Wills-Sherris suppose que l'intensité des taux de mortalité est donnée par une fonction en lien avec l'âge de la cohorte que l'on souhaite étudier ainsi que le temps qui s'écoule. Ils utilisent aussi la matrice de variance-covariance des résidus de leur modèle afin de tenir compte des interdépendances entre les âges lors de la simulation de scénarios de taux de mortalité futurs. Ce modèle a pour avantage d'être facilement adaptable pour tarifier des produits financiers. Ce modèle est celui que nous avons retenu à cause de sa simplicité. Simplicité qui est très importante puisque nous cherchons à établir les bases d'un produit standard et donc le modèle retenu se doit d'être compréhensible pour la masse des agents qui, nous l'espérons, les transigerons. De plus, sa parenté avec le modèle de Black-Schole-Merton permet d'anticiper certaines faiblesses potentielles et de les corriger.

3.4 Théories utilisées

Plusieurs techniques statistiques sont employées dans les prochaines sections. Il est donc primordial d'explicitier certains concepts avant de les utiliser. Cette section sera divisée en 2 parties. La première se consacre aux différents tests statistiques et la deuxième explicite certains concepts d'algèbre linéaire ainsi qu'une technique nommée l'analyse par composantes principales.

3.4.1 Tests d'hypothèses

Les tests d'hypothèses sont l'un des aspects les plus importants dans la modélisation statistique. Ces tests sont composés de deux hypothèses. La première est celle que l'on nomme hypothèse nulle (H_0). La deuxième est l'hypothèse alternative (H_1). Ces tests ont tous un seuil de signification. Ce seuil est le point que l'on considère acceptable de commettre une erreur de type 1. C'est-à-dire la probabilité du rejet de H_0 alors que l'hypothèse est vraie. Typiquement, ce seuil de signification est de 1%, 5% ou 10%. Au cours de la modélisation statistique, plusieurs tests sont effectués afin de déterminer si les paramètres sont significatifs et si les résidus obtenus par

la calibration du modèle ont les propriétés souhaitées. Ces propriétés sont les suivantes :

1. Absence de tendance
2. Absence d'autocorrélation entre les résidus
3. Homoscédasticité (constance de la variance)
4. Normalité

Quatre tests différents sont utilisés pour confirmer ou infirmer les points ci-haut. Il s'agit des tests:

1. Augmented Dickey-Fuller
2. Breusch-Godfrey
3. Arch
4. Shapiro-Wilk

Le test Augmented Dickey-Fuller est le test qui vérifie la présence d'une tendance dans une série chronologique. L'hypothèse nulle de ce test est la confirmation de la présence d'une racine unitaire dans les données. Les hypothèses du test sont :

H_0 : La série est intégrée d'ordre 1 (Tendance)

H_1 : La série est intégrée d'ordre 0 (Stationnarité)

Il est à noter que la statistique calculée ne suit pas une loi normale. Cela s'explique par le fait que le théorème limite centrale ne s'applique pas sur les séries qui ne sont pas stationnaires. Le nom de cette distribution est simplement la table de Dickey-Fuller. Cette table est spécifique à ce test et fût développée par ces deux auteurs.

Le test de Breusch-Godfrey est le test qui vérifie l'autocorrélation de l'ensemble de données. L'hypothèse de base de ce test est l'absence d'autocorrélation et l'alternative est la présence d'autocorrélation. Ce test permet de déterminer s'il y a une structure dans l'ensemble de données autre que de la tendance. Pour fin d'interprétations justes, il doit être effectué sur des

séries chronologiques stationnaires. Ce test, sous H_0 , suit une loi χ^2 . Les hypothèses de ce test sont :

H_0 : La série n'est pas autocorrélée

H_1 : La série est autocorrélée

Le test Arch vérifie si la variance d'une série donnée est constante dans le temps. Ce test est l'un des plus importants considérant que l'on veut éviter de sous-estimer l'occurrence d'évènements extrêmes. En effet, l'hypothèse d'une variance constante alors qu'elle ne l'est pas a pour effet de sous-estimer ces évènements qui pourraient entraîner d'importantes pertes dans un portefeuille. Ce qui entraîne donc une surestimation de la viabilité du portefeuille. Le test utilise le carré de l'ensemble des résidus échantillonaux. Puisque par définition, l'espérance de cette série est 0, alors théoriquement $E[\epsilon_t^2] = \sigma^2$. Où ϵ_t est la variable des résidus. Les hypothèses du test sont :

H_0 : La série est homoscedastique (Variance constante)

H_1 : La série est hétéroscedastique (Variance variable dans le temps)

Le test de Shapiro-Wilk vérifie s'il est raisonnable de supposer la normalité d'une série donnée. Ce test est le plus approprié dans le cas des taux de mortalité car l'échantillon est de petite taille, soit moins que 2000. Les hypothèses du test sont :

H_0 : La série est distribuée normalement

H_1 : La série n'est pas distribuée normalement

Notons la totalité des tests d'hypothèses effectués dans cette étude ont pour seuil de signification 5%.

3.4.2 Sélection d'un modèle

Afin de sélectionner le meilleur modèle parmi un ensemble de ceux-ci, nous utiliserons deux critères de sélection soit le Akaike information criterion (AIC) et le Bayesian information criterion (BIC). Dans chacun des critères, on retrouve la même idée de base. C'est-à-dire qu'ils bonifient un modèle pour la vraisemblance de la normalité des résidus. Par contre, ils le pénalisent pour le nombre de paramètres à estimer. En résumé, à résidus égaux, le modèle choisi est celui qui a le moins de paramètres à estimer.

Puisque cela a pour effet de diminuer la variance des prévisions. La formule du AIC est la suivante :

$$AIC = 2k - 2 \ln(\mathcal{L})$$

Avec

$$\ln \mathcal{L} = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2$$

Où σ est l'écart-type des résidus échantillonnaires et \mathcal{L} représente la fonction de vraisemblance des résidus. Le BIC, quant à lui, est donné par la formule suivante :

$$BIC = -2 \ln(\mathcal{L}) + k \ln(n)$$

On remarque que le critère BIC pénalise davantage un modèle ayant un plus grand nombre de paramètres à estimer. Il est aussi à noter que ces mesures ne constituent pas un test d'hypothèse, mais seulement d'un calcul pour sélectionner le modèle optimal. Le choix du modèle sera celui où le critère AIC ou BIC est minimisé. Il peut arriver que dans un ensemble de plus de 3 modèles, les critères proposent deux modèles distincts. Il faut alors choisir le modèle ayant le moins de paramètres à estimer. Dans le cadre de notre étude, nous comparons les 2 critères afin d'utiliser le maximum d'informations disponibles.

3.4.3 Algèbre linéaire

Dans cette section, nous introduisons quelques définitions ainsi que quelques propriétés des matrices qui sont utiles à l'étude. Tout d'abord, nous dénotons une matrice A par la majuscule et les entrées de cette matrice par a_{ij} pour l'entrée située à la i -ème ligne et j -ième colonne. Une matrice A est symétrique si $a_{ij} = a_{ji}$ et ce $\forall i, j$. Un exemple typique de ce type de matrice est celui des matrices de variance-covariance. Elles sont fréquemment utilisées et sont certainement symétriques.

Les matrices transposées sont un autre type de matrices qui sont utilisées dans la suite de l'étude. Soit a_{ij} les entrées de la matrice A . La transposée de A , A^T , est donnée par la matrice où les entrées sont a_{ji} . À la différence de la matrice symétrique, l'entrée a_{ij} n'a pas forcément besoin d'être égale à l'entrée a_{ji} .

Les valeurs propres d'une matrice A sont les valeurs, λ_i , telles que :

$$Ax = \Lambda x \quad (3.2)$$

$$\iff (A - \Lambda)x = 0 \quad (3.3)$$

Pour x un vecteur différent de 0. Λ est la matrice diagonale contenant les valeurs propres, λ_i , de la matrice A . De plus, x est un vecteur propre de la matrice A . Il est à noter qu'une matrice peut contenir plusieurs valeurs propres et qu'il existe autant de vecteurs propres distincts et indépendants que de valeurs propres distinctes.

La trace d'une matrice est donnée par la somme des éléments sur sa diagonale. Conséquemment, la trace se calcule seulement pour les matrices carrées. Nous notons cette mesure ainsi:

$$\sum_{i=1}^N a_{ii} = tr(A)$$

Un résultat utile et important, utilisant les valeurs propres d'une matrice, est le suivant:

$$tr(A) = tr(\Lambda)$$

Ce qui nous indique que nous pouvons classer les vecteurs propres en ordre d'importance et déterminer l'importance relative de chacun d'entre eux en effectuant le calcul suivant :

$$\frac{\lambda_i}{tr(\Lambda)}$$

L'utilisation de ce résultat est de réduire le nombre de variables utilisées dans un modèle. La technique à employer est explicitée dans la section suivante.

Ces valeurs et vecteurs propres permettent aussi de calculer la décomposition spectrale de la matrice A . Elle permet de simplifier certains calculs matriciels. Elle est donnée par l'égalité suivante :

$$A = x\Lambda x^T$$

Où x est la matrice contenant les vecteurs propres associés à ces valeurs λ_i . Cette égalité est vraie seulement si chacune des valeurs propres sont distinctes dans le cas d'une matrice non-symétrique. Dans le cas où la matrice est symétrique, les valeurs propres n'ont pas à être distinctes afin d'admettre

cette décomposition. Une autre façon de décomposer une matrice A est de trouver une matrice D qui est telle que :

$$A = DD^T$$

La technique pour trouver cette matrice se nomme la décomposition de Cholesky. Cette décomposition permet de simuler des variables normales ayant une certaine structure de corrélation. La matrice D peut être obtenue de quelques différentes façons. L'une d'entre-elle est d'utiliser la décomposition spectrale comme outil. Tel que stipulé plus haut, nous calculons les vecteurs propres ainsi que les valeurs propres d'une matrice. Nous optons pour $e_i\sqrt{\Lambda}$ comme matrice D . Il est vérifiable que $A = DD^T$ puisque :

$$\begin{aligned} DD^T &= (x\sqrt{\Lambda})(x\sqrt{\Lambda})^T \\ &= x\sqrt{\Lambda}\sqrt{\Lambda}x^T \\ &= x\Lambda x^T \\ &= A \end{aligned}$$

Cette matrice est utilisée lorsque vient le temps de simuler des variables aléatoires normales ayant une structure de variance-covariance.

3.4.4 Analyse par composantes principales

Décomposition des matrices

Lorsque le nombre de variables est grand, il peut s'avérer difficile d'interpréter correctement la matrice de dispersion. L'objectif de l'analyse par composantes principales est de réduire ce nombre de variables à étudier en ne conservant que les plus importantes. L'algèbre linéaire est l'outil utilisé pour déterminer les variables qui ont la plus grande influence sur les variables étudiées. Pour ce faire, il faut d'abord déterminer les valeurs propres de la matrice de variance-covariance, Σ , des données étudiées. Nous dénotons la matrice contenant les valeurs propres de la matrice par V_i . Nous calculons aussi les vecteurs propres associés à ces valeurs dénotées par P_i . Il faut ensuite les classer en ordre décroissant et les placer sur la diagonale d'une matrice, \bar{V} , avec le nombre de lignes égal au nombre de valeurs propres. Par la suite, il faut placer les vecteurs correspondant, en colonnes, dans une matrice, \bar{P} , dans le même ordre que les valeurs propres correspondantes. Celle-ci est une matrice orthogonale Johnson, Wichern (2007) [16].

De ce fait, nous avons que $P^{-1} = P^T$. En utilisant le résultat précédent, nous sommes en mesure d'établir que :

$$\Sigma = \bar{P}\bar{V}\bar{P}^{-1}$$

et que

$$\Sigma = (\bar{P}\sqrt{\bar{V}}\bar{P}^{-1})^2$$

à l'aide de la décomposition spectrale de la matrice. Nous retrouvons alors une matrice qui satisfait l'équation de la décomposition de Cholesky puisque $\bar{P}\sqrt{\bar{V}}\bar{P}^{-1} = (\bar{P}\sqrt{\bar{V}}\bar{P}^{-1})^T$.

Cette relation est utilisée lorsqu'il faut simuler des scénarios taux de mortalité à l'aide du modèle de Wills-Sherris (2011)[33]. Puisqu'elle nous permet d'obtenir des variables aléatoires normales ayant la même structure de variance-covariance que celle que l'on désire simuler.

Composantes principales

L'objectif de l'analyse par composantes principales est de simplifier les interprétations des matrices de variance-covariance (Johnson-Wichern 2007)[16]. Cette technique repose seulement sur la décomposition de ces matrices et donc aucune hypothèse en lien avec la distribution des données n'est requise. De plus, s'il s'avère que l'échantillon suit une loi normale multivariée, alors il est possible d'obtenir certaines propriétés facilitant l'analyse (Johnson-Wichern 2007)[16]. Afin de déterminer les composantes, expliquant la variance de la variable Y à l'aide de l'ensemble de variables X_i , il faut décomposer la matrice de variance-covariance Σ_X à l'aide des valeurs propres et les vecteurs associés. Il est possible de démontrer que $\sum_{i=1}^N Var(X_i) = tr(\bar{V}) = \sum_{i=1}^N Var(Y_i)$ (Johnson-Wichern 2007)[16]. Par conséquent, la proportion de la variance expliquée par la $k^{i\text{ème}}$ variable est donnée par :

$$\frac{\lambda_i}{tr(\bar{V})}$$

Cette décomposition permet de réduire le nombre de variables à simuler ce qui diminue le nombre de calculs à effectuer.

3.4.5 Mesures de risque

Jusqu'à présent, nous avons énoncé certaines théories pouvant être utilisées pour modéliser le comportement de variables aléatoires. Dans le cadre de

l'étude, ces variables sont les taux centraux de mortalité. Les prévisions de ceux-ci sont utilisées dans la construction de portefeuilles d'annuités. Afin de mesurer le risque lié à ces portefeuilles, introduisons certaines mesures qui ont un lien avec les portefeuilles d'actifs financier.

L'article de Markowitz (1952) [21] est le pionnier dans l'utilisation d'une mesure de risque comme critère de décision dans le choix d'un portefeuille. Dans son cas, il utilise la variance des rendements comme mesure de risque. Il utilise cette mesure afin de construire son critère de décision qui se base sur la relation risque-rendement. Bien que l'objectif de base soit d'optimiser le choix d'un portefeuille, la variance est depuis utilisée comme une mesure de risque. Au cours des dernières années, les mesures de risques ont été grandement développées. Des balises ont été établies par Artzner et al.(1999) [2]. Dans leur article, ils définissent le risque et proposent certains axiomes qui sont encore utilisés aujourd'hui comme base des mesures de risque. Le risque, au sens de Artzner, est l'incertitude face à un changement entre 2 dates futures. Avant d'énoncer les axiomes, définissons deux variables aléatoires de pertes que nous dénotons par p_1 et p_2 . Donc, en cas de perte du portefeuille 1, la variable p_1 est positive et en cas de gain, elle est négative. La même logique s'impose pour la variable p_2 . Ces variables permettent de formaliser les propriétés de cohérence des mesures de risque qui ont été proposées par Artzner et al. Selon eux, pour qu'une mesure de risque R soit cohérente, il faut qu'elle soit :

1. homogène :

$$R(\lambda p_1) = \lambda R(p_1)$$

2. monotone :

$$p_1 \leq p_2 \implies R(p_1) \leq R(p_2)$$

3. invariable par translation :

$$R(p_1 + c) = R(p_1) + c$$

Pour une constante c .

4. sous-additive :

$$R(p_1 + p_2) \leq R(p_1) + R(p_2)$$

Définissons maintenant deux mesures de risque qui sont parmi les plus populaires aujourd'hui, soit la valeur à risque (VaR) et l'expected shortfall

(ES). La VaR défini un niveau de potentiel de perte, sur une certaine période de temps, considérant un certain niveau de confiance (α). La phrase résumant le mieux la VaR est : «Je suis confiant à $(1-\alpha)\%$ que nous ne perdrons pas plus de VaR\$ dans les T prochains jours». Mathématiquement, la VaR, d'une variable de rendement x , est donnée par l'équation suivante:

$$VaR = \inf(x : 1 - F_X(-x) \geq \alpha) \quad (3.4)$$

En supposant la normalité de la distribution des rendements, on en détermine la valeur par le calcul suivant:

$$VaR(X) = \mu(x) - \sigma(x)\Phi^{-1}(\alpha)$$

Où $\Phi(\alpha) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha} e^{-\frac{s^2}{2}} ds$

Aujourd'hui, elle est devenue l'une des mesures les plus importantes dans les institutions financières puisque les autorités réglementaires l'utilisent dans la définition du capital requis. Elle est donc très utile afin de comparer les différents portefeuilles au niveau de leurs pertes potentielles. Une faiblesse de la VaR est qu'elle n'est pas une mesure de risque cohérente. Elle ne respecte pas la sous-additivité lorsque les pertes ne sont pas normalement distribuées.

Au niveau de l'expected shortfall, elle représente la moyenne des pertes supérieures la VaR. On calcule sa valeur de la façon suivante :

$$ES = \mathbf{E}[X|X > VaR] \quad (3.5)$$

L'avantage d'utiliser cette mesure est qu'elle répond à tous les critères de cohérence de mesure de risque. En supposant que les rendements sont distribués normalement, l'expected shortfall est alors donnée par :

$$ES = \mu - \sigma \frac{e^{-\frac{(\Phi^{-1}(\alpha))^2}{2}}}{\sqrt{2\pi}(1-\alpha)} \quad (3.6)$$

Rappelons que l'hypothèse de normalité des rendements est très peu respectée dans le cas des rendements sur les actifs financiers. Il est bien connu que la kurtose de ces distributions est en général plus élevée que 3. Cependant, il est possible de modifier cette hypothèse de distribution des rendements par une student à t degrés de liberté afin d'obtenir une VaR et un ES qui reflèteront mieux les pertes potentielles. D'autres formules et ajustements seront alors nécessaires pour obtenir leurs valeurs. Ces mesures complètent l'ensemble des concepts théoriques qui seront utilisés dans le cadre de notre étude.

4. Données

Avant de présenter les données utilisées, définissons le taux central de mortalité pour la cohorte âgée de x années à l'année t , $\mathbf{m}_x(t)$. Ce taux est donné par :

$$\frac{\text{Nombre de décès lors de l'an } t \text{ des gens âgés de } (x) \text{ années}}{\text{Population, observée, âgée de } (x) \text{ au 30 juin de l'an } t} \quad (4.1)$$

Dans le cadre de notre étude, les données utilisées sont les taux centraux de mortalité chez les femmes canadiennes âgées de 20 à 110 ans pour les années 1921 à 2011. Elles sont disponibles via la base de données de mortalité humaine (Human Mortality Database) [37]. Cette base de données a été créée par le département de démographie de l'Université de Californie ainsi que l'Institut Max Planck (situé en Allemagne). Elle contient les taux de mortalité de 38 pays dont ceux du Canada. Comme le modèle retenu pour l'étude l'oblige, nous utilisons les données des cohortes discrétisées. C'est-à-dire que nous utilisons les taux centraux âge par âge. Pour le modèle de Lee-Carter, les données utilisées sont, en général, groupées par cohorte de 5 ans. Elles sont d'ailleurs disponibles aussi via le site de la base de données de mortalité humaine. Dénotons par $\mathbf{m}_x(t)$ le taux central de mortalité d'une cohorte, avec x , l'âge de la cohorte observée, et t , l'année d'observation du taux. Avec cette notation, la matrice contenant les données est de la forme suivante :

$$\begin{array}{c} \text{Âge} \backslash \text{Années} \\ \begin{array}{c} 20 \\ 21 \\ \vdots \\ 110 \end{array} \end{array} \begin{pmatrix} 1921 & 1922 & \dots & 2011 \\ \mathbf{m}_{20}(1921) & \mathbf{m}_{20}(1922) & \dots & \mathbf{m}_{20}(2011) \\ \mathbf{m}_{21}(1921) & \mathbf{m}_{21}(1922) & \ddots & \mathbf{m}_{21}(2011) \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{m}_{110}(1921) & \dots & \dots & \mathbf{m}_{110}(2011) \end{pmatrix}$$

Nous pouvons constater visuellement, à l'aide de la figure 4.1, que les données semblent avoir une tendance à la baisse pour chacun des âges lorsque nous les observons à travers le temps. Ce phénomène est mis en

évidence à la figure 4.2 où sont présentés les taux centraux de mortalité des femmes âgées de 65 ans à travers le temps.

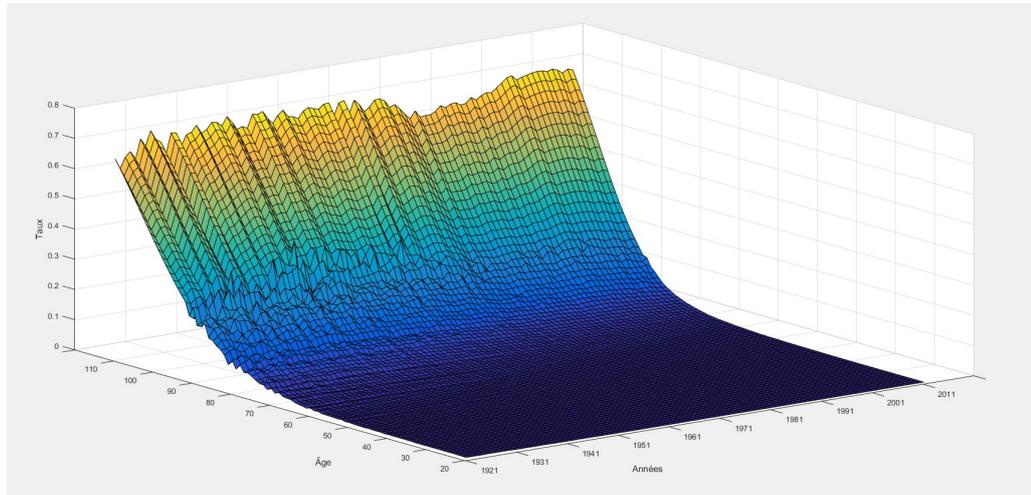


Figure 4.1: Taux de mortalité bruts des femmes âgées de 20 à 110 ans

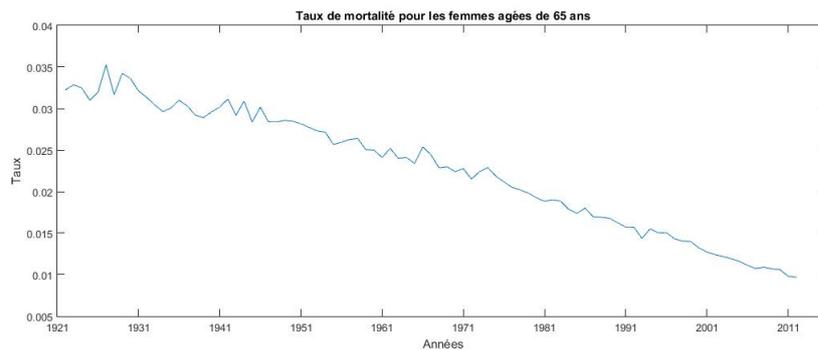


Figure 4.2: Taux de mortalité des femmes âgées de 65 ans par année

L'objectif du mémoire étant de tarifer des produits de longévité, nous devons donc utiliser une dynamique qui permet de suivre le taux de mortalité d'une cohorte à travers le temps. Un critère important en lien avec la dynamique retenue est la stationnarité des données. Puisque la série ne semble pas stationnaire, nous différencions les données afin de tenter de les rendre stationnaires à l'aide de l'équation suivante :

$$\Delta \mathbf{m}_x(t) = \mathbf{m}_{x+1}(t+1) - \mathbf{m}_x(t) \quad (4.2)$$

Cette équation est la dynamique qui nous intéresse puisqu'elle permet de suivre l'amélioration du taux de mortalité d'une cohorte d'un âge x déterminé à travers le temps. La matrice contenant ces variations a alors la forme

suivante :

$$\begin{array}{c}
 \text{\AA ge} \backslash \text{Ann\u00e9es} \\
 \begin{array}{c}
 20 \\
 21 \\
 \vdots \\
 109
 \end{array}
 \end{array}
 \begin{pmatrix}
 1921 & 1922 & \dots & 2010 \\
 \Delta \mathbf{m}_{20}(1921) & \Delta \mathbf{m}_{20}(1922) & \dots & \Delta \mathbf{m}_{20}(2010) \\
 \Delta \mathbf{m}_{21}(1921) & \Delta \mathbf{m}_{21}(1922) & \ddots & \Delta \mathbf{m}_{21}(2010) \\
 \vdots & \ddots & \ddots & \vdots \\
 \Delta \mathbf{m}_{109}(1921) & \dots & \dots & \Delta \mathbf{m}_{109}(2010)
 \end{pmatrix}$$

On pr\u00e9sente \u00e0 la figure 4.3 les donn\u00e9es diff\u00e9renci\u00e9es. Afin de mieux visualiser si les s\u00e9ries semblent stationnaires, nous isolons les diff\u00e9rences pour les femmes \u00e2g\u00e9es de 65 ans \u00e0 travers le temps. Celles-ci sont pr\u00e9sent\u00e9es \u00e0 la figure 4.4. Ces deux graphiques indiquent que les variations semblent stationnaires. Nous pouvons donc chercher un mod\u00e8le qui offre un bon fit des diff\u00e9rences de taux centraux de mortalit\u00e9.

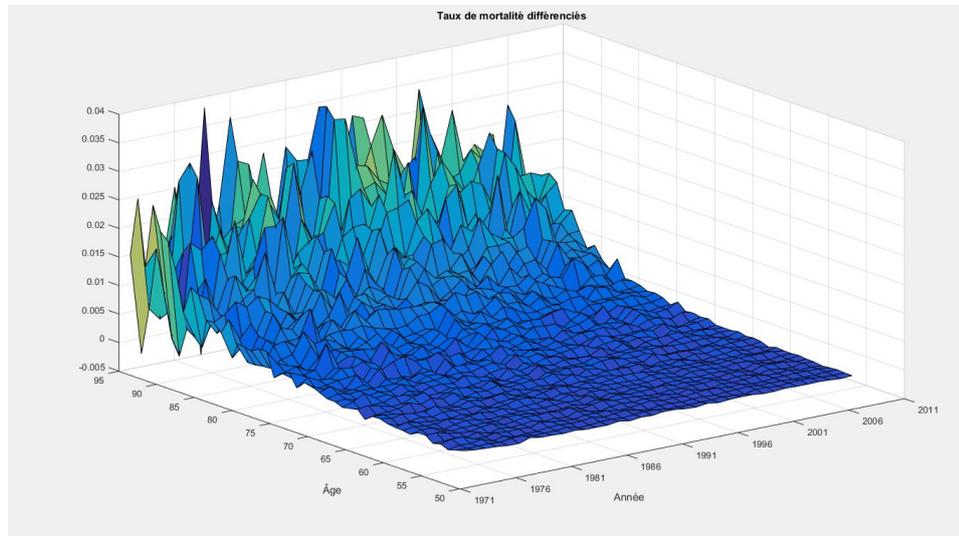


Figure 4.3: Taux de mortalit\u00e9 diff\u00e9renci\u00e9s

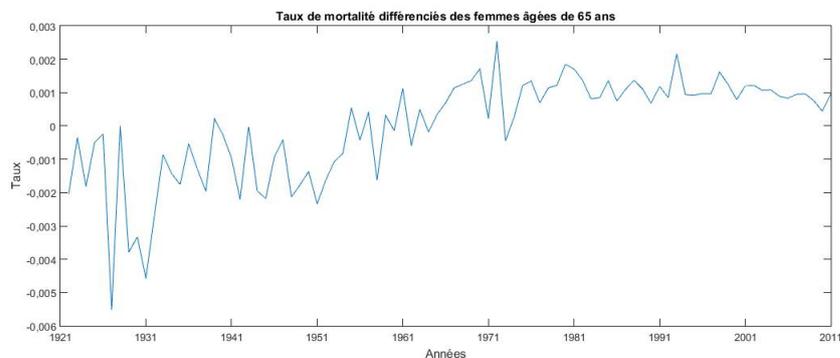


Figure 4.4: Taux de mortalit\u00e9 des femmes \u00e2g\u00e9es de 65 ans diff\u00e9renci\u00e9s

On constate généralement que la volatilité des taux de mortalité augmente avec l'âge Wills-Sherris (2008) [33]. Nous retrouvons aussi cette propriété au niveau des données canadiennes. Cette hausse de la volatilité s'explique par le calcul du taux central de mortalité. Nous savons que la population âgée de 90 ans est plus petite que celle âgée de 65 ans. Par conséquent, chaque décès de personnes âgées de 90 ans a un effet plus important que ceux de gens âgés de 65 ans par rapport à la population de chacun des âges.

5. Cadre méthodologique

Dans ce chapitre, nous présentons le cadre théorique du mémoire. Afin de faciliter la compréhension du modèle retenu, introduisons la définition formelle du taux central de mortalité telle qu'employée par Wills-Sherris (2008). Ce taux est donné par l'équation suivante:

$$m_x(t) = \frac{\text{Nombre de décès lors de l'an } t \text{ des gens âgés de } (x+t) \text{ années}}{\text{Population âgée de } (x+t) \text{ observée au 30 juin de l'an } t}$$

Nous modifions aussi la signification de t . Avec cette notation, t représente le nombre cardinal d'années après l'année 0 qui correspond à la première année d'observation de l'âge x . Cette modification a pour but de faciliter la compréhension de la signification des prévisions obtenues par le modèle. Elle entraîne aussi une modification de la relation 4.2 qui devient alors :

$$\Delta m_x(t) = m_x(t+1) - m_x(t) \quad (5.1)$$

Comme le modèle prévoit les variations des taux de mortalité futurs pour une cohorte prédéfinie, cette notation permet de garder un indice indiquant l'âge initial de cette cohorte.

5.1 Modèle utilisé

Le modèle retenu dans le cadre de ce mémoire en est un qui présente des caractéristiques similaires à celui de Black-Scholes-Merton [6]. Comme les faiblesses du modèle de BSM sont bien connues, il est alors plus facile d'identifier celles du modèle choisi. Il s'agit du modèle proposé par Will-Sherris en 2011 [33]. L'hypothèse de base est que les taux de mortalité suivent une distribution log-normale. Le modèle permet d'estimer les variations futures des taux de mortalité, d'une cohorte, d'une année à l'autre. Il est défini comme suit :

$$\begin{aligned} \frac{\Delta m_x(t)}{m_x(t)} &= (a(x+t) + b)dt + \sigma d\mathbb{W}_x(t) \\ \implies \Delta m_x(t) &= m_x(t)(a(x+t) + b)dt + m_x(t)\sigma d\mathbb{W}_x(t) \end{aligned}$$

Le modèle est basé sur 2 variables principales, x et t , qui sont données en années. La valeur de x est constante lors de l'étape des prévisions et fixée au début de celle-ci. C'est-à-dire que si nous voulions prévoir l'évolution des taux de mortalité des gens appartenant à la cohorte des 65 ans, nous fixerions $x = 65$. La fonction en t est un facteur considérant le nombre d'années écoulées depuis le début de l'ensemble des données choisies. Dans notre cas, 1975 est l'année de référence. Ce qui veut dire que $t=0$ correspond à cette année, $t=1$ correspond à l'année 1976 et ainsi de suite. Les composantes a , b et σ sont des paramètres que nous estimons à l'aide de la méthode du maximum de vraisemblance.

La variable $\mathbb{W}_x(t)$ représente un mouvement brownien ayant une matrice de variance-covariance Σ en x et en t . La matrice $d\mathbb{W}_x(t)$ peut s'exprimer comme le produit de 2 matrices. Soit $dZ(t)$ et D deux matrices de la même dimension que $d\mathbb{W}_x(t)$. $dZ(t)$ représentent une matrice diagonale de variables aléatoires normales standard. D représente la matrice correspondant à la décomposition de Cholesky de la matrice de variance-covariance des résidus échantillonnaux. Cette décomposition est faite en utilisant la méthodologie présentée dans la section 3.4.3. Soulignons que les portions déterministes ainsi qu'aléatoires du processus stochastique du modèle sont proportionnelles au dernier taux de mortalité observé. Cela signifie que plus l'âge ou le temps augmente, plus la volatilité des taux de mortalité prévus augmente et conséquemment celle des taux de mortalité prévus aussi. Ce qui indique que les caractéristiques des prévisions du modèle semblent être similaires à celles des données. Connaissant la variation des taux de mortalité, le calcul de la prédiction du taux de mortalité est donc:

$$m_x(t + 1) = m_x(t) + \Delta m_x(t) \quad (5.2)$$

Lors de l'étape de l'estimation des prévisions, $m_x(0)$ est donné par le dernier taux connu de la cohorte âgée de x années. Dans le cadre de notre étude, la dernière année correspond à 2011. Nous calculons ensuite la prévision de $\Delta m_x(0)$ et utilisons l'équation 5.2 pour obtenir la prévision du taux de mortalité de cette cohorte dans un an. Nous répétons la procédure pour le nombre d'années désirées en faisant varier t .

5.2 Calibration du modèle

La calibration du modèle se fait en deux étapes. La première est d'estimer les trois paramètres en supposant l'indépendance de ceux-ci. La deuxième est de décomposer la matrice de variance-covariance des résidus afin de tenir compte des différentes dépendances entre les cohortes. Tel que proposé par Wills-Sherris [33], l'estimation des paramètres du modèle est faite par maximum de vraisemblance.

L'hypothèse de base du modèle est $\Delta m_x(t) \sim \mathcal{N}(m_x(t)(a(x+t)+b), \sigma^2 m_x^2(t))$. Conséquemment, la fonction du log de vraisemblance est donnée par:

$$\ell(\Delta m_x(t)|a, b, \sigma) = - \sum_{\forall x,t} \ln(\sigma m_x(t) \sqrt{2\pi}) - \frac{1}{2} \sum_{\forall x,t} \left(\frac{\frac{\Delta m_x(t)}{m_x(t)} - (a(x+t) + b)}{\sigma} \right)^2$$

En calculant les dérivées partielles par rapport à chacune des variables, nous obtenons le système d'équations suivant :

$$\frac{\partial \ell}{\partial a} = \frac{1}{\sigma^2} \left(\sum_{\forall x,t} \frac{\Delta m_x(t)(x+t)}{m_x(t)} - [\hat{a} \sum_{\forall x,t} (x+t)^2 + \hat{b} \sum_{\forall x,t} (x+t)] \right) = 0 \quad (5.3)$$

$$\frac{\partial \ell}{\partial b} = \frac{1}{\sigma^2} \left(\sum_{\forall x,t} \frac{\Delta m_x(t)}{m_x(t)} - [\hat{a} \sum_{\forall x,t} (x+t) + \hat{b}(NT)] \right) = 0 \quad (5.4)$$

$$\frac{\partial \ell}{\partial \sigma} = \frac{(NT)}{\sigma} + \sigma^{-3} \sum_{\forall x,t} \left(\frac{\Delta m_x(t)}{m_x(t)} - (a(x+t) + b) \right)^2 = 0 \quad (5.5)$$

Où N est le nombre cardinal d'âges dans l'échantillon retenu pour la calibration. À titre d'exemple, lors d'une calibration du modèle en utilisant les données pour les femmes âgées de 20 à 110 ans, nous avons N=91 puisque $N = 110 - 20 + 1$ et T est le nombre d'années des données. Comme nous voulons maximiser le log-vraisemblance, nous devons résoudre le système d'équations pour $\frac{\partial \ell}{\partial a}$ et $\frac{\partial \ell}{\partial b} = 0$. Il est alors possible d'obtenir les estimés de \hat{a} et \hat{b} à l'aide des 2 premières équations. Avec ces 2 paramètres estimés, nous utilisons l'équation (3) pour estimer $\hat{\sigma}$.

5.2.1 Paramètres attendus

Il y a certaines attentes au niveau des valeurs des paramètres estimés. Celles-ci découlent des hypothèses du modèle. Nous anticipons que la valeur de \hat{a} sera très près de 0 puisque la distribution $\Delta m_x(t)$ est, par hypothèse, stationnaire pour une cohorte fixe. Elle doit cependant être différente de 0

puisque les améliorations dépendent de l'âge de la cohorte. Si cette valeur n'était pas statistiquement différente de 0, alors la conclusion serait que les améliorations sont une moyenne constante dans le temps. Ce qui n'est clairement pas le cas. De plus, les $\Delta m_x(t)$ doivent être positifs. C'est pour cette raison que nous croyons que \hat{b} doit être positif puisque dans le cas contraire, les taux de mortalité seraient décroissant dans le temps pour une cohorte dont l'âge augmente. Cette hypothèse est, selon nous, la plus importante à respecter afin que le modèle ait une certaine crédibilité.

5.3 Fenêtre de temps sélectionnée

Lors de l'étude préliminaire des données, il a été soulevé que certaines données pourraient potentiellement biaiser les résultats. Ceci se constate visuellement lorsqu'on observe le graphique 5.1. Celui-ci montre l'évolution des taux de mortalité des femmes âgées de 20 à 23 ans de 1921 à 1975. Nous expliquons ces hausses notoires, qui sont encadrées dans le graphique, par des événements historiques. À titre d'exemple, la deuxième guerre mondiale semble être un de ces événements qui ont eu pour effet d'augmenter anormalement la mortalité tant chez les hommes que chez les femmes. En faisant plus de recherches, le site de Santé Canada nous informe que lors de ces années, il y a eu une pandémie de la polio [36]. Ce qui explique la hausse des taux chez les jeunes femmes. De plus, il semble y avoir 2 tendances différentes au niveau des données. Nous nous retrouvons donc dans la même situation qu'exprimée par Li et al. (2015) [19]. À cause de ce changement de structure, il apparaît évident que certaines données plus anciennes nuisent à la précision du modèle. En effet, le changement de structure présent dans les données pourrait amener le modèle à surestimer l'amélioration des taux de mortalité. Li et al. (2015) [19] ont démontré que les interprétations faites à partir des modèles sont sensibles au nombre d'années considérées. Par exemple, ils ont démontré que l'espérance de vie augmente de manière significative lorsque le nombre de données diminue et vice-versa. Pour cette raison, nous croyons qu'il est important de calculer la fenêtre de temps optimale, qui sera utilisée afin de calibrer le modèle, dans le but d'en faire des interprétations qui sont justes. L'objectif est d'utiliser les données qui représentent le mieux les variations futures des taux de mortalité. Afin d'établir cette période, nous utilisons la méthodologie que Li et al. (2015) [19] ont employée.

Dénotons la matrice contenant les données présentées à la section précédente par \mathcal{J} . Cette matrice contient les taux de mortalité des années 1921 à 2011 pour les âges de 20 à 110 ans. Nous la séparons en deux parties. L'une sert de partie d'historique afin de calibrer différents modèles (H_t). Elle contient N éléments. La deuxième sert d'ensemble de réalisations hors-échantillon (R_t) et contient M éléments. Nous avons donc que $|N| + |M| = |\mathcal{J}|$. L'année de division est arbitraire, mais afin d'obtenir quelques statistiques élémentaires, nous optons pour un ensemble de 10 années de données post-calibration. La matrice des données est donc sous la forme suivante :

$$\left(\begin{array}{ccc|ccc} \mathbf{m}_{20}(1920) & \dots & \mathbf{m}_{20}(2001) & \mathbf{m}_{20}(2002) & \dots & \mathbf{m}_{20}(2011) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{m}_{110}(1920) & \dots & \mathbf{m}_{110}(2001) & \mathbf{m}_{110}(2002) & \dots & \mathbf{m}_{110}(2011) \end{array} \right)$$

In sample
Out of sample

Les années 2002 à 2011 sont utilisées pour obtenir des résidus prévisionnels. Pour ce faire, la première étape est de calibrer le modèle de Wills-Sherris (2011) en utilisant les 10 dernières années de l'ensemble historique de données. Dans cet exemple, la calibration est faite à l'aide des données 1992 à 2001. À partir du modèle, nous calculons les prévisions des taux centraux de mortalité des 10 années qui suivent 2001 pour la cohorte des femmes âgées de 65 ans. Cette cohorte plus particulière a été choisie puisqu'il s'agit de celle qui nous intéresse dans la suite de l'étude. Nous dénotons les prévisions du modèle par P_t . Nous pouvons maintenant calculer les erreurs de prévisions grâce à l'ensemble post-calibration. Ce qui nous permet de calculer la moyenne des résidus au carré, dénotée par MRC, pour cet ensemble de 10 éléments. Nous répétons ensuite la procédure avec une calibration utilisant les 11 dernières années de la partie historique. Ainsi de suite jusqu'à ce que nous ayons calibré le modèle en utilisant la totalité de l'échantillon historique. La fenêtre optimale choisie, est celle où la MRC est minimale. La formule utilisée afin de calculer la MRC est la suivante:

$$MRC = \sum_{i=1}^{10} (R_t - P_t)^2$$

Où P_t est la prévision au temps t.

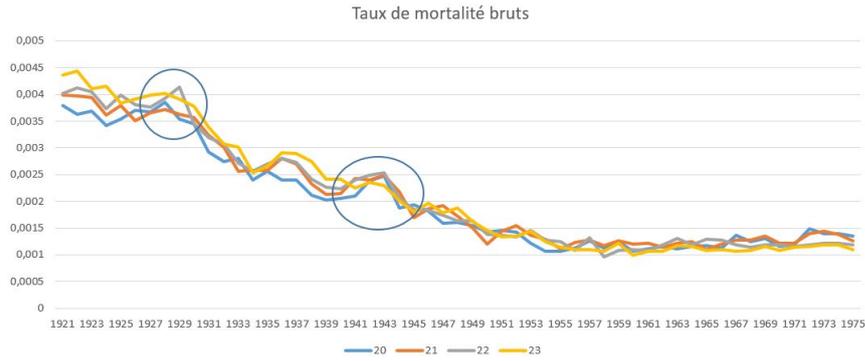


Figure 5.1: Taux de mortalité des femmes âgées de 20 à 23 ans

5.3.1 Résultats de la fenêtre choisie

Avant de donner les résultats, nous posons une hypothèse en lien avec le nombre minimal d'années nécessaires à la calibration du modèle. Puisque nous voulons effectuer des prévisions sur une très longue période, nous estimons que le nombre minimal requis d'années est de 25. L'argument principal pour le choix de ce nombre est la volonté d'éviter d'effectuer des prévisions sur une période de 40 ans en utilisant seulement 10 années de données comme base pour la calibration du modèle. Le tableau 5.1 contient les résultats obtenus. Puisque la somme est minimisée à 37 ans, nous choisissons donc les dernières 37 années de taux de mortalité pour calibrer le modèle.

Tableau 5.1: Somme de la MRC du modèle selon le nombre d'années

Nombre d'années de données	MRC des erreurs post calibration
25	1.24686E-07
26	1.23444E-07
27	1.23136E-07
28	1.22302E-07
29	1.22407E-07
30	1.22431E-07
31	1.23141E-07
32	1.22678E-07
33	1.22254E-07
34	1.21974E-07
35	1.22344E-07
36	1.2201E-07
37	1.21804E-07
38	1.2366E-07
39	1.22629E-07
40	1.21959E-07
41	1.22759E-07
42	1.22944E-07
43	1.23043E-07
44	1.22528E-07
45	1.23514E-07
46	1.22721E-07
47	1.23621E-07
48	1.24183E-07
49	1.24902E-07

5.4 Résultats de la calibration

Nous avons calibré le modèle, selon la procédure décrite plus tôt, pour 3 ensembles différents de cohortes et leurs taux de mortalité associés lors des années 1975 à 2011 inclusivement. Ces ensembles sont ceux des femmes 20 à 95 ans, 50 à 95 ans et 65 à 95 ans. Cette procédure permet de tester les calibrations du modèle et d'observer les variations des paramètres estimés selon les échantillons utilisés. Nous optons pour ces cohortes car elles

semblent être des choix logiques. Elles permettent aussi de voir l'influence qu'ont les cohortes des femmes âgées de 20 à 49 ans et de 20 à 64 ans sur les paramètres du modèle. Les paramètres estimés sont les suivants :

cohortes	\hat{a}	\hat{b}	$\hat{\sigma}$
(1) 20-95	0.000924464	0.008021946	0.09028358
(2) 50-95	0.000619717	0.034612454	0.04685208
(3) 65-95	0.000571767	0.041027404	0.038469237

Tableau 5.2: Estimations des paramètres

Comme les modèles (2) et (3) se ressemblent beaucoup, il est nécessaire de se baser sur les critères AIC/BIC pour faire un choix optimal. Nous calculons aussi les critères pour le modèle (1). Cependant le paramètre \hat{b} du modèle (1) ne semble pas représentatif de la réalité des cohortes âgées de plus de 50 ans puisqu'il est beaucoup plus petit par rapport aux deux autres calibrations. On remarque aussi que ce paramètre augmente considérablement lorsque les cohortes utilisées pour la calibration sont plus âgées. Ceci s'explique par le retrait des taux centraux de mortalité des femmes plus jeunes. Puisque la variabilité de ces taux est beaucoup plus faible pour l'ensemble des gens âgés entre 20 et 49 ans, nous concluons qu'ils font diminuer la moyenne des améliorations de taux de mortalité en pourcentage. De plus, comme les améliorations sont près de 0, il est normal d'observer des améliorations négatives parfois pour ces tranches d'âges. Par contre, ce phénomène ne devrait pas survenir fréquemment pour les tranches d'âges de 50 ans et plus.

Afin de choisir le meilleur modèle, nous nous basons, rappelons-le, sur les critères AIC et BIC de chacun d'entre-eux ainsi que sur la qualité des prévisions. Le fait d'avoir besoin de prévisions qui ont un certain sens pendant 45 années oblige à mettre l'emphasis sur leurs qualités. C'est-à-dire qu'il n'est pas souhaitable d'obtenir des prévisions de taux de mortalité négatives après un certain temps ou qu'elles soient supérieures à 1. Nous avons calculé les AIC et BIC en utilisant le maximum de vraisemblance des prévisions intra-échantillon. Selon les résultats présentés au tableau 5.3, nous choisissons le modèle qui est basé sur les taux de mortalités associés aux gens de plus de 50 ans puisqu'il a le plus petit AIC. Ce choix est aussi

en concordance avec les modélisations faites par Wills-Sherris(2011) [33] et Fung et al. (2015) [12]. Nous utilisons ce modèle pour le reste de l'étude.

cohortes	AIC	BIC
20-95	-5292.312974	-5270.752248
50-95	-5311.501132	-5289.940406
65-95	-3964.148983	-3944.210118

Tableau 5.3: Mesures AIC/BIC des différentes calibrations

La matrice de variance-covariance asymptotique du modèle (2) est donnée à la table 5.4. Elle permet d'estimer la variance des paramètres et conséquemment le niveau de confiance de ceux-ci. Cette matrice est donnée par:

$$\begin{bmatrix} var(\hat{a}) & cov(\hat{a}, \hat{b}) & cov(\hat{a}, \hat{\sigma}) \\ cov(\hat{a}, \hat{b}) & var(\hat{b}) & cov(\hat{b}, \hat{\sigma}) \\ cov(\hat{a}, \hat{\sigma}) & cov(\hat{b}, \hat{\sigma}) & var(\hat{\sigma}) \end{bmatrix} = \begin{bmatrix} -\frac{\partial^2 L}{\partial \hat{a}^2} & -\frac{\partial^2 L}{\partial \hat{a} \partial \hat{b}} & -\frac{\partial^2 L}{\partial \hat{a} \partial \hat{\sigma}} \\ -\frac{\partial^2 L}{\partial \hat{a} \partial \hat{b}} & -\frac{\partial^2 L}{\partial \hat{b}^2} & -\frac{\partial^2 L}{\partial \hat{b} \partial \hat{\sigma}} \\ -\frac{\partial^2 L}{\partial \hat{a} \partial \hat{\sigma}} & -\frac{\partial^2 L}{\partial \hat{b} \partial \hat{\sigma}} & -\frac{\partial^2 L}{\partial \hat{\sigma}^2} \end{bmatrix}^{-1}$$

Puisque ces valeurs sont petites, il est possible de conclure en un haut niveau de confiance envers les paramètres estimés.

Param	\hat{a}	\hat{b}	$\hat{\sigma}$
\hat{a}	3.02E-12	-2.71E-10	0
\hat{b}	-2.71E-10	2.51E-08	0
σ	0	0	3.16E-13

Tableau 5.4: Matrice de variance-covariance asymptotique

Puisque les améliorations historiques des taux de mortalité sont positives la majeure partie du temps, pour une cohorte d'âge fixe à travers le temps, il est important de s'assurer que les améliorations prévues par le modèle ont cette même caractéristique. Ces prévisions d'amélioration, qui sont données par l'équation 5.6, ont cette même propriété la quasi-totalité du temps.

$$I(t) = 1 - \frac{\mathbf{m}_x(t+1)}{\mathbf{m}_x(t)} \quad (5.6)$$

Nous avons effectué cette vérification sur une période de 30 ans. À l'exception de quelques années, la totalité des prévisions respectent cette propriété.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
65	-0.01684	0.059702	0.070849	-0.06676	0.044071	0.021913	0.007796	0.010204	0.026447	0.069683	-0.02998	0.04046	0.049155	0.027769	0.035444
66	0.036594	-0.01173	0.059984	0.071549	-0.06732	0.044592	0.021528	0.007182	0.010435	0.026556	0.069694	-0.02998	0.040769	0.0489	0.027773
67	0.018925	0.035612	-0.01608	0.059546	0.071765	-0.06748	0.044993	0.021272	0.007346	0.010177	0.026769	0.069705	-0.02976	0.040295	0.048966
68	0.025926	0.018061	0.035914	-0.01536	0.060122	0.071487	-0.06778	0.04512	0.021903	0.006819	0.009817	0.02704	0.070039	-0.02983	0.040312
69	0.010855	0.02469	0.019326	0.035786	-0.01574	0.060147	0.071147	-0.06708	0.044947	0.021885	0.006392	0.010686	0.026608	0.070001	-0.02993
70	0.04523	0.009772	0.026017	0.019309	0.0354	-0.01564	0.060349	0.071187	-0.06716	0.045093	0.021644	0.005846	0.011145	0.026652	0.070343
71	0.006845	0.044618	0.010721	0.026333	0.019161	0.034204	-0.01478	0.060026	0.071708	-0.06764	0.045583	0.021509	0.005655	0.011457	0.026279
72	0.024691	0.005429	0.045478	0.010579	0.027134	0.018506	0.034901	-0.01536	0.060169	0.072087	-0.0683	0.045875	0.021483	0.00564	0.011088
73	0.023417	0.023547	0.005937	0.045411	0.010686	0.026992	0.018685	0.034742	-0.0146	0.060204	0.071731	-0.06736	0.045674	0.021277	0.005812
74	0.033369	0.022354	0.023856	0.006374	0.045432	0.009924	0.027478	0.018409	0.034645	-0.01406	0.059643	0.071736	-0.06726	0.045422	0.021297
75	0.00858	0.032988	0.022109	0.024439	0.005924	0.046167	0.009412	0.027398	0.018841	0.034472	-0.01398	0.059645	0.071633	-0.06765	0.045914
76	0.037487	0.007449	0.03392	0.022198	0.024202	0.005662	0.046892	0.008704	0.027086	0.019423	0.034582	-0.01386	0.059189	0.071518	-0.06743
77	0.038258	0.036493	0.008036	0.034431	0.022437	0.02394	0.005115	0.046925	0.009573	0.026534	0.01908	0.034916	-0.01397	0.059104	0.071652
78	0.030099	0.037344	0.037519	0.007781	0.034786	0.02241	0.024054	0.004754	0.047101	0.009277	0.026819	0.018958	0.034956	-0.01376	0.058749
79	0.051571	0.027959	0.039341	0.037823	0.007638	0.03449	0.022059	0.024708	0.003699	0.047864	0.008939	0.027119	0.019135	0.035002	-0.01428

Figure 5.2: Taux d'amélioration par l'âge et le temps

Nous expliquons les améliorations négatives par l'effet de vague présent dans les prévisions. Il est possible de le constater visuellement dans le graphique 5.3. Wills et Sherris (2008) ont aussi fait face à cet effet qui est créé par le modèle. Ils expliquent ce phénomène par les taux de mortalité plus élevés de certaines cohortes plus jeunes lors de la dernière année d'observation. Comme ces taux servent de point d'ancrage et que les prévisions sont en moyenne du même ordre, il est normal que les prévisions aient les mêmes propriétés que ce point de départ.

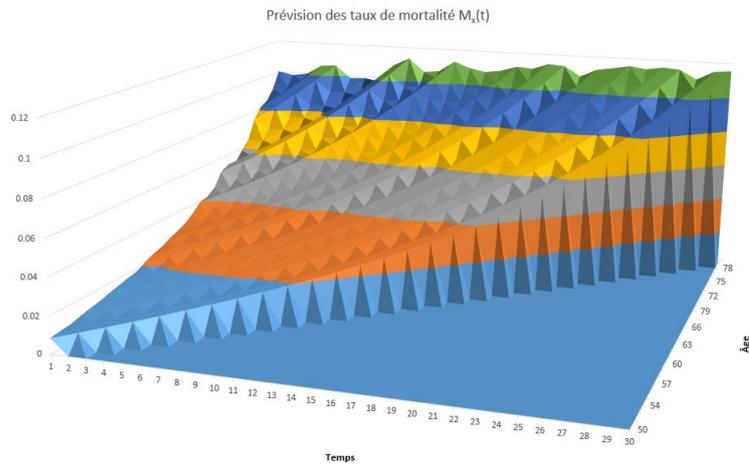


Figure 5.3: Prévision des taux par l'âge et le temps

5.5 Décomposition des résidus

Une fois la première étape de la calibration faite, nous devons calculer les résidus centrés réduits. Ceux-ci sont donnés sous la forme suivante :

$$r_x(t) = \frac{\Delta\mu}{\mu} - (\hat{a}(x+t) + \hat{b})$$

Cette matrice constitue la portion aléatoire du processus. Puisque l'objectif est de générer des scénarios cohérents de variation de taux de mortalité représentant l'évolution passée, il est nécessaire que ceux-ci aient la

même structure de variance-covariance que celle des résidus. Conséquemment, l'étape qui s'impose est la décomposition de la matrice de variance-covariance des résidus par âge. Les étapes effectuées lors de cette procédure sont les mêmes que celles décrites à la section 3.4.4.

5.5.1 Analyse des résidus

Cette section est consacrée à l'analyse des résidus obtenus précédemment. L'analyse est effectuée sur les résidus de la cohorte des femmes âgées de 65 ans puisque ce sont ceux que nous simulons dans les étapes de couverture d'un portefeuille de rentes. Les points vérifiés sont : l'absence de tendance, la normalité, l'homoscédasticité ainsi que l'absence d'autocorrélation pour ceux-ci. La figure 5.4 permet de constater visuellement que les résultats semblent prometteurs puisqu'il ne semble pas y avoir de tendance ni d'hétéroscédasticité pour les âges que nous observerons plus en détails. De plus, il ne semble pas y avoir de données aberrantes.

Le premier point vérifie la présence de la tendance. Pour ce faire, le test de Dickey-Fuller augmenté qui, on le rappelle, vérifie s'il existe une racine unitaire dans l'ensemble des données, est l'outil utilisé. Le résultat du test est présenté à la figure 5.5. Celui-ci confirme qu'il est raisonnable de conclure à la stationnarité des résidus. Du côté de la normalité, le test de Shapiro-Wilk indique qu'il est acceptable de supposer que les résidus générés par le modèle suivent une loi normale (figure 5.6). Le test Arch permet de supposer l'homoscédasticité des résidus (figure 5.7). Finalement, le test de Breush-Godfrey permet de supposer l'absence d'autocorrélation (5.8). Il est aussi possible de le constater visuellement à l'aide du corrélogramme présenté à la figure 5.9. Les tests ont été effectués pour chacun des ensembles des résidus (un par âge). Pour chacun d'entre-eux nous avons obtenus des résultats permettant de supposer que dans 90% des cas, les résidus ont toutes les caractéristiques souhaitées.

Figure 5.4: Résidus centrés réduits

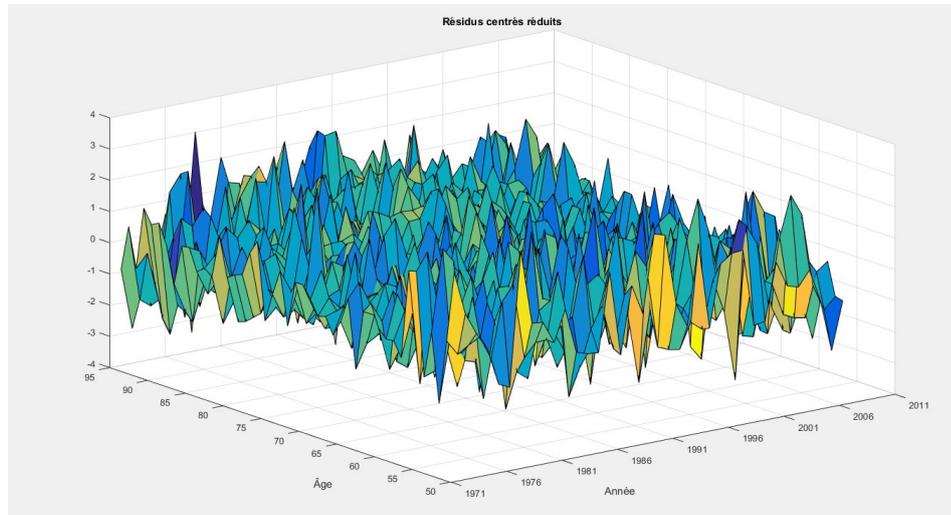


Figure 5.5: Test de la racine unitaire

Null Hypothesis: SERIES01 has a unit root
 Exogenous: Constant
 Lag Length: 0 (Automatic - based on SIC, maxlag=9)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-5.556296	0.0000
Test critical values:		
1% level	-3.626784	
5% level	-2.945842	
10% level	-2.611531	

*Mackinnon (1996) one-sided p-values.

Figure 5.6: Test de la normalité des résidus

Tests de normalité						
	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistique	ddl	Signification	Statistique	ddl	Signification
VAR00001	,090	37	,200	,967	37	,332

*. Il s'agit d'une borne inférieure de la signification réelle.
 a. Correction de signification de Lilliefors

Figure 5.7: Test le l'homoscédasticité de la variance

Dependent Variable: SERIES01
 Method: ML - ARCH (Marquardt) - Normal distribution
 Date: 07/11/16 Time: 12:11
 Sample: 1 37
 Included observations: 37
 Convergence achieved after 12 iterations
 Presample variance: backcast (parameter = 0.7)
 GARCH = C(1) + C(2)*RESID(-1)^2

Variable	Coefficient	Std. Error	z-Statistic	Prob.
Variance Equation				
C	0.379503	0.095393	3.978302	0.0001
RESID(-1)^2	-0.121352	0.114371	-1.061040	0.2887
R-squared	-1.066324	Mean dependent var		0.414884
Adjusted R-squared	-1.010477	S.D. dependent var		0.407316
S.E. of regression	0.577538	Akaike info criterion		1.839980
Sum squared resid	12.34138	Schwarz criterion		1.927057
Log likelihood	-32.03963	Hannan-Quinn criter.		1.870679
Durbin-Watson stat	0.844040			

Figure 5.8: Test de la présence d'autocorrélation

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	0.503318	Prob. F(10,26)	0.8721
Obs*R-squared	6.000924	Prob. Chi-Square(10)	0.8152

Figure 5.9: Test de la présence d'autocorrélation

Date: 07/11/16 Time: 12:11
 Sample: 1 37
 Included observations: 37

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 0.102	0.102	0.4151	0.519
		2 0.011	0.000	0.4198	0.811
		3 -0.055	-0.057	0.5497	0.908
		4 0.050	0.062	0.6584	0.956
		5 0.261	0.255	3.7382	0.588
		6 0.148	0.102	4.7644	0.574
		7 -0.053	-0.083	4.8983	0.672
		8 0.095	0.141	5.3485	0.720
		9 0.059	0.045	5.5304	0.786
		10 0.047	-0.056	5.6502	0.844
		11 0.006	-0.045	5.6519	0.896
		12 -0.076	-0.058	5.9887	0.917
		13 0.185	0.174	8.0391	0.841
		14 -0.064	-0.179	8.2963	0.873
		15 -0.112	-0.133	9.1236	0.871
		16 -0.227	-0.193	12.653	0.698

5.6 Validité du modèle

Afin de déterminer si le modèle représente bien les données en général, nous utilisons le test du Goodness-of-fit. Les hypothèses de ce test sont :

H_0 : Le modèle représente bien les données

H_1 : Le modèle ne représente pas bien les données

La statistique de ce test est donnée par :

$$\sum_{\forall x,t} \frac{(O(x,t) - E(x,t))^2}{E(x,t)} \quad (5.7)$$

où

$O(x,t)$ est la valeur observée à l'âge x au temps t

et

$E(x,t)$ est la valeur estimée à l'âge x au temps t par le modèle

Selon l'hypothèse du test, elle suit une loi χ^2 . Conséquemment, nous allons la comparer à la valeur critique du test qui est donnée par $\chi_D^2(\alpha)$. Où D est le nombre de degrés de liberté. Ce nombre est donné par:

$$D = \text{Nb. d'observations} - \text{Nb. de paramètres indépendants} - 1 \quad (5.8)$$

Dans le cas où cette statistique est plus petite que la valeur critique, il est acceptable de supposer que le modèle représente bien les données. Dans le cas qui nous intéresse, il y a 1620 données et 1225 vecteurs dépendants. La distribution chi-carré nécessaire au test statistique a donc 399 degrés de liberté. Nous avons obtenu les résultats qui sont présentés à la table 5.5. Comme la valeur de la statistique est plus petite que la valeur théorique, il est acceptable de supposer que le modèle offre une bonne représentation des données.

χ^2 Femme	χ^2 théorique
41.694	353.699

Tableau 5.5: Résultat du test chi-carré

5.7 Analyse par composantes principales

Dans cette section, nous effectuons l'analyse par composantes principales telle qu'expliquée à la section 3.4.4. Cette analyse constitue la deuxième étape de la calibration du modèle dont la finalité, le rappelle-t-on, est d'effectuer des simulations de taux de mortalité. Les résultats de l'analyse démontrent que cet ensemble est constitué de 35 vecteurs indépendants et que 15 suffisent à expliquer 85,02% de la variation des résidus. Le tableau 5.6 présente les résultats obtenus. Par souci d'exhaustivité, nous conservons les 35 vecteurs indépendants pour la génération de nombres aléatoires. Il est cependant acceptable, en pratique, d'en conserver un nombre inférieur pour cette étape.

nombre de vecteurs	% de la variation expliquée
1	14.1
2	25.12
5	47.26
10	70.70
15	85.02
25	97.51
35	1

Tableau 5.6: *Pourcentage de la variation observée dans les résidus en utilisant l'analyse en composantes principales*

5.8 Prévisions hors-échantillon

Puisque la précision est de mise pour les prévisions hors-échantillon, nous croyons légitime de les étudier afin d'avoir une idée de la magnitude de celles-ci. De plus, elle peut servir de comparatif avec les tables actuarielles générées par l'institut canadien des actuaires. Nous portons une attention plus particulière aux résidus générés par le modèle sur 40 ans de prévisions. Pour ce faire, nous calibrons le modèle en utilisant toujours une fenêtre de 37 ans ayant comme dernière année d'observation 1971. Ceci permet d'effectuer des prévisions sur une période de 40 ans et ainsi obtenir les résidus hors-échantillon sur cette même période. De plus, nous pouvons observer le comportement à très long terme des prévisions. Nous avons calibré le modèle en utilisant les données de 1935 à 1971. Les paramètres

estimés de celles-ci sont présentés au tableau 5.7. Les prévisions de taux de mortalité sont ensuite comparées aux $\mathbf{m}_x(t)$ correspondant. Dans notre cas, nous effectuons ces prévisions pour la cohorte âgée de 65 ans à $t = 0$ et calculons la différence entre l'évolution des taux espérés par le modèle et les taux réalisés. C'est-à-dire que la première prévision du taux central de mortalité donnée par $m_x(1)$ est comparée à $\mathbf{m}_{66}(1971)$. Idéalement, la distribution de ces résidus est de moyenne 0 et la variance est constante. En pratique nous ne nous attendons pas à ce que ce soit le cas. Cette hypothèse provient de la construction du modèle. Comme nous utilisons la prévision au temps t dans la prévision pour le temps $t+1$, il est raisonnable de croire que si la 1^{ière} est erronée alors la 2^{ième} le sera tout autant sinon plus puisqu'elle est proportionnelle à la première. Les résultats obtenus sont présentés au tableau 5.8.

\hat{a}	\hat{b}	$\hat{\sigma}$
0,00020383	0,065574269	0,084736807

Tableau 5.7: Paramètre du modèle de Will-Sherris 1935-1971

En regardant les résultats, on peut affirmer que le modèle effectue d'excellentes prévisions à long terme. La moyenne des erreurs, après 10 ans, est très près de 0. De plus, le niveau de précision des estimations est excellent. Cependant, sur la totalité de la période, le modèle sous-estime les taux de mortalité de 1,76% en moyenne sur la totalité de la période. Par contre, pour l'évaluation de swap ou cap de 20 ans, le modèle semble suffisamment précis pour être utilisé.

T	Réalisation	Prévision	Erreur
0=1971	0.0215		
1	0.02404	0.02319	0.00085
2	0.02447	0.02503	-0.00056
3	0.02762	0.02701	0.00061
4	0.02944	0.02916	0.00028
5	0.03173	0.03148	0.00025
6	0.03374	0.03399	-0.00025
7	0.03564	0.03671	-0.00107
8	0.03785	0.03966	-0.00181
9	0.04098	0.04285	-0.00187
10	0.04408	0.04631	-0.00223
11	0.04841	0.05005	-0.00164
12	0.0525	0.05411	-0.00161
13	0.0556	0.0585	-0.0029
14	0.06126	0.06327	-0.00201
15	0.06942	0.06844	0.00098
16	0.07225	0.07404	-0.00179
17	0.08018	0.08012	0.00006
18	0.08871	0.08671	0.002
19	0.0976	0.09387	0.00373
20	0.10321	0.10163	0.00158
21	0.11091	0.11005	0.00086
22	0.1275	0.1192	0.0083
23	0.13681	0.12913	0.00768
24	0.15442	0.13991	0.01451
25	0.16941	0.15163	0.01778
26	0.1962	0.16435	0.03185
27	0.20682	0.17818	0.02864
28	0.22763	0.1932	0.03443
29	0.24861	0.20953	0.03908
30	0.26364	0.22729	0.03635
31	0.29273	0.24659	0.04614
32	0.31912	0.26759	0.05153
33	0.34631	0.29043	0.05588
34	0.37539	0.31527	0.06012
35	0.39551	0.34231	0.0532
36	0.43293	0.37173	0.0612
37	0.46071	0.40376	0.05695
38	0.47966	0.43863	0.04103
39	0.52107	0.4766	0.04447
40	0.54856	0.51796	0.0306

Tableau 5.8: Erreur des prévisions hors-échantillon sur 40 ans

5.9 Sensibilités

Nous abordons maintenant l'effet qu'ont différentes variations des paramètres sur les taux de survies. Il est important de bien connaître les effets globaux de ces variations potentielles puisque les gestionnaires veulent connaître l'ampleur des gains et pertes potentielles. Comme les erreurs liées à la paramétrisation existent, il faut conséquemment tenir compte des effets de ces erreurs potentielles.

Débutons avec l'effet qu'a \hat{a} sur les taux de survie. Lorsqu'il augmente, il a pour effet d'augmenter les taux de mortalité et par conséquent, diminuer les taux de survie. Globalement, une sous-estimation de \hat{a} implique une sur-estimation du prix juste de l'actif financier. À l'inverse, une surestimation du paramètre aura pour effet de sous-estimer le prix de l'actif. En ce qui concerne \hat{b} , une augmentation de celui-ci a pour effet d'augmenter les taux de mortalité. Nous nous retrouvons donc avec la même logique que pour \hat{a} , c'est-à-dire qu'une sous-estimation de \hat{b} entraîne une surestimation du prix de l'actif. Au niveau de $\hat{\sigma}$ s'il augmente, les taux de mortalité seront plus volatiles dans le temps. Cependant au niveau de la tarification de produits de longévité un paramètre plus élevé aura pour effet d'augmenter le prix du produit. Ce qui est normal puisque la prime de risque liée au produit est liée à la volatilité de celui-ci. Ce qui nous indique qu'une sous-estimation de σ entraîne une sous-estimation du prix des actifs. Puisque la matrice de variance-covariance asymptotique des paramètres estimés, présentée au tableau 5.4, nous permet d'avoir un haut niveau de confiance envers nos estimés, nous considérons que le risque d'erreur lié à l'estimation des paramètres est faible.

5.10 Lissage des données

Nous abordons maintenant la calibration du modèle à l'aide de données lissées. Ce qui constitue une deuxième modélisation. La technique du lissage de données est surtout utilisée afin de réduire le bruit présent dans les données. Elle permet ainsi des estimations plus robustes. Ce qui se traduit généralement par une baisse de la volatilité des estimations. Il existe plusieurs techniques de lissage. Dans notre cas, nous employons celle du lissage cubique. Cette technique utilise une suite de polynômes du 3^{ième}

degré reliés un à un par des noeuds. Cette méthodologie permet d’assurer la continuité et la différenciabilité de la courbe, quelques conditions supplémentaires y sont imposées. Ces deux caractéristiques sont nécessaires afin de permettre l’obtention de taux instantané de mortalité et ce à n’importe quelle date (Joanis 2016) [15]. De plus, le lissage cubique s’avère être la technique utilisée par LLMA [20] ainsi que la plus répandue au niveau de l’ajustement des taux de mortalité. Conséquemment, le choix de cette technique permet de rapprocher la méthodologie employée dans cette étude de celle utilisée en pratique.

Une fois les données lissées, nous calibrons le modèle avec ces nouvelles données. Les résultats des deux calibrations effectuées sont présentées dans le tableau 5.9. Ces résultats ont de quoi rassurer certains investisseurs potentiels puisque malgré les chocs introduits dans les données, le modèle reste sensiblement le même à l’exception de la variance qui, comme attendu, réduit de presque la moitié.

cohortes	\hat{a}	\hat{b}	$\hat{\sigma}$
50-95	0.0006197	0.03461	0.04685
50-95 Lissées	0.000678	0.0298131	0.025397

Tableau 5.9: Comparaison des résultats avec/sans données lissées

5.11 Ajustement pour le risque

Puisque l’utilisation principale du modèle est de tarifier des produits financiers, il est important d’y inclure une composante afin de tenir compte de la prime de risque. Cet ajout constitue la troisième modélisation que nous abordons. Cette variable, λ , est ajoutée au modèle en employant les mêmes étapes qu’expliquées par Cairns (2004) [9]. λ a pour effet de diminuer la tendance déterministe et, par conséquent, augmenter les taux de survie prévus. Globalement, ce changement de tendance a pour effet d’augmenter le prix de produits financiers dérivés ayant le taux de survie comme sous-jacent. En y incorporant une prime pour le risque, cet ajustement de processus permet aussi de tarifier d’autres produits ayant les taux de mortalité comme sous-jacent tout en prévenant l’arbitrage. Nous déno-

tons dW^{RA} comme le bruit blanc ajusté pour le risque. Comme Wills-Sherris l'ont fait en 2011 [33], l'ajustement est donné sous la forme suivante:

$$\begin{aligned} dW_x^{RA}(t) &= \sum_{i=1}^N \delta_{xi}(dZi(t) - \lambda_i(t)) \\ \implies dW_x^{RA}(t) &= dW_x(t) - \sum_{i=1}^N \delta_{xi}\lambda_i(t) \end{aligned}$$

Le modèle ajusté devient alors :

$$\Delta m_x(t) = m_x(t) \left(a(x+t) + b - \sigma \sum_{i=1}^N \hat{\lambda}_i(t) \delta_{xi} \right) dt + m_x(t) \sigma dW_x(t) \quad (5.9)$$

Les $\hat{\lambda}_i(t)$ peuvent être déterminé à l'aide du prix d'autres actifs de mortalité/longévité dans le marché. Il est adapté de telle sorte que si l'on calcule le prix d'un actif financier à l'aide des taux de mortalité ajustés pour le risque, alors le prix demandé au marché est égal à celui de la moyenne des scénarios générés par le modèle. Il prend donc la forme d'un vecteur puisque les primes de risque dépendent de l'âge de la cohorte ciblée par le produit. Comme il n'existe pas de marché liquide pour les produits de longévité, il est plus difficile d'obtenir plusieurs valeurs de $\hat{\lambda}_i$. Dans notre cas, nous utilisons une composante agrégée moyenne tenant compte du risque. Cette hypothèse facilite la calibration du modèle au niveau de l'ajustement pour le risque puisqu'il y a qu'un seul λ a trouvé. L'équation du modèle ajusté pour cette prime devient alors :

$$\Delta m_x(t) = m_x(t) \left(a(x+t) + b - \sigma \lambda(t) \right) dt + m_x(t) \sigma dW_x(t) \quad (5.10)$$

Éventuellement, avec le développement d'un marché, plusieurs facteurs λ_x pourront être considérés. La formule 5.9 sera utilisée à ce moment afin de déterminer les variations de taux de mortalité en tenant compte de la prime de risque liée à l'âge.

6. Portefeuille simple cohorte

Dans ce chapitre, la discussion porte sur les simulations d'un portefeuille de rentes ainsi que des statistiques en lien avec la rentabilité de celui-ci. Il est aussi question de produits financiers qui pourraient être utilisés à des fins de couverture de portefeuille. La dernière section porte sur l'effet de ces produits sur les mesures de risque du portefeuille. Les mesures étudiées en lien avec le portefeuille sont la moyenne, l'écart-type, la kurtose, la valeur à risque à 99% ainsi que l'expected shortfall à 99%.

6.1 Portefeuille de base

Dans cette section, nous présentons les hypothèses de travail en lien avec le portefeuille simple cohorte. Ce type de portefeuille est le plus simple qu'il soit possible de construire. De ce fait, il est naturel de l'utiliser afin de bien comprendre les effets potentiels des produits de couverture.

Malgré sa simplicité, il est important de poser quelques hypothèses raisonnables permettant d'obtenir certains résultats de base. Ces résultats sont utilisés afin d'évaluer l'efficacité des produits de longévité. La première d'entre elles est que ce portefeuille est constitué uniquement de 4000 femmes âgées de 65 ans. Malgré le fait que cette hypothèse éloigne la situation de la réalité elle permet de conserver une certaine simplicité. Nous rappelons que cette étape constitue une prémice dans la couverture du risque de longévité et donc la simplicité facilite la compréhension des résultats. Il est habituel d'imposer un âge maximal dans la population. Nous supposons comme Fung et al (2015)[12] l'ont fait que les décès arrivent avant 110 ans. Le taux sans risque utilisé, qui est le même que celui utilisé par Fung et al. (2015) [12], est de 4% constant. Cette hypothèse simplificatrice nous permet de nous concentrer sur la dynamique des taux de mortalité. Comme l'utilisation des modèles de taux d'intérêt est bien comprise, il suf-

fira d'intégrer les taux adéquats lors de la tarification des produits dans un marché transigé. De plus, ces portefeuilles sont, en général, protégés contre le risque de taux d'intérêt. Évidemment, lors de l'inclusion de swap de longévité dans une stratégie de couverture, un gestionnaire de risque doit tenir compte de la durée attachée à ce swap.

Au chapitre précédent, nous avons présenté deux types de modèles. Chacun d'entre eux ont leur utilité. Le premier produit des prévisions basées sur le meilleur estimé (\tilde{S}) et le deuxième permet d'obtenir les prévisions ajustées pour le risque ($\tilde{\tilde{S}}$). Rappelons que la prime de risque est attribuable à la réduction de la tendance donné par λ . Comme on s'y attend, plus λ est élevée, plus le prix de l'actif financier est élevé. Il est important de mentionner que si la prime de risque est trop élevée alors le nombre d'investisseurs potentiels devrait diminuer à cause du coût trop élevé de l'actif. Il est donc important de trouver le point d'équilibre entre le rendement requis et le risque que l'entreprise est prête à prendre.

6.1.1 Obligation de longévité

Avant de tarifier les annuités de notre portefeuille, nous introduisons le prix d'une obligation de longévité (Meyricke, Sherris 2014) [22]. Cette obligation a une date d'échéance de 25 années. Elle sert d'exemple afin de mieux comprendre la suite du travail.

$$V(0) = \sum_{i=1}^{25} B(0, i) e^{\delta i} E_0^P [e^{-\int_0^i m_x(v) dv}] \quad (6.1)$$

Où:

1. δ est la prime de risque moyenne annuelle
2. $B_0(i)$ est le facteur d'actualisation pour la période de 0 à i
3. $E_0^P [e^{-\int_0^i m_x(v) dv}]$ est le meilleur estimé de la survie au moment i

Nous supposons que $\delta = 0.002$. Cette hypothèse découle de la prime de risque moyenne annuelle utilisée pour la tarification de l'obligation de longévité émise par Eurobank en 2003 pour la population du Royaume-Uni. Les coupons de cette obligation, souscrite par BNP Paribas, varient en fonction du taux de mortalité cumulatif observé d'une cohorte. L'hypothèse

utilisée pour la prime de risque est la même que celle utilisée par Meyricke-Sherris (2014) [22]. Il faut tout de même utiliser cette prime de risque avec quelques précautions puisque l'obligation n'a pas trouvée preneur sur le marché financier. Elle fût retirée de celui-ci en 2005. Plusieurs raisons sont évoquées pour expliquer cet échec. Chen et Cummins (2010) [10] soulèvent entre autres que la date d'échéance était trop rapprochée ce qui enlevait de l'efficience à la couverture. De plus, ce type d'obligation monopolise beaucoup de capital requis considérant la réduction de risque potentielle que le produit offre. Le risque associé aux modèles employés est, lui aussi, très élevé. Ce sont ces quelques raisons et biens d'autres qui ont mené à l'échec de ce produit. Nous utilisons tout de même cette obligation afin de déterminer la valeur de λ . Similairement, nous établissons la valeur d'une obligation en utilisant le modèle ajusté pour le risque de la façon suivante :

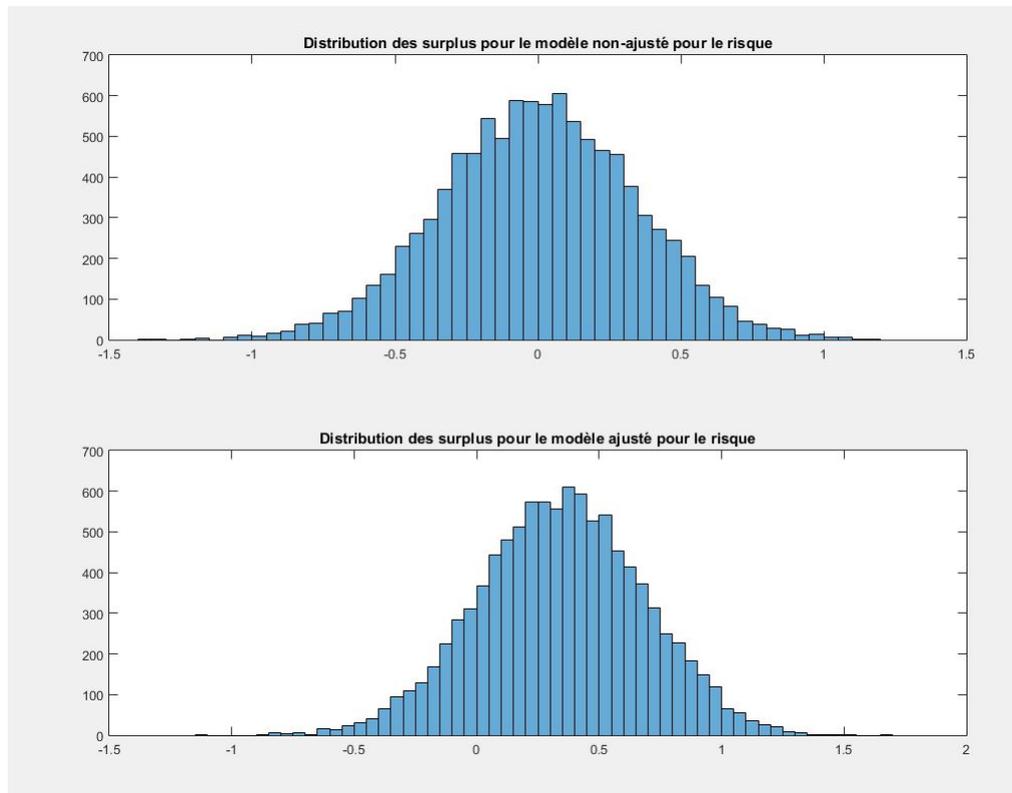
$$V_0 = \sum_{i=1}^{25} B(0, i) E_0^{RA} [e^{-\int_0^i m_x(v) dv}] \quad (6.2)$$

Afin d'évaluer la valeur de λ , nous calculons la valeur de l'obligation à l'aide de la formule 6.1. Cette valeur représente le prix que demanderait PNB\EIB pour cette obligation liée avec les taux de survie de la population des femmes canadiennes âgées de 65 ans. Nous utilisons ce prix pour déterminer la valeur de λ qui est telle que les équations 6.1 et 6.2 soient égales. Comme il n'existe pas de formule analytique, nous effectuons plusieurs simulations en faisant varier λ et utilisons celui qui satisfait la contrainte imposée.

6.2 Calibration pour le risque

Dans cette section, nous calibrons le modèle ajusté pour le risque à l'aide de la méthodologie proposée à la section précédente. Nous avons déterminé que la valeur de l'obligation, selon l'équation 6.1 est de 14.7672\$. Nous calculons ensuite la valeur de λ de telle sorte que la valeur de l'équation 6.2 est égale à 14.7672\$. En utilisant la méthode de l'approximation successive, nous trouvons qu'un λ de 0.1425 nous permet d'obtenir une valeur d'obligation de 14.7662 ce que nous considérons acceptable. La figure 6.1 montre que l'ajustement pour le risque ne fait que déplacer la moyenne de la distribution vers la droite sans modifier les autres mesures, c'est-à-dire l'écart-type, le skewness ou la kurtose.

Figure 6.1: Distribution des surplus sans/avec l'ajustement pour le risque



6.3 Le portefeuille et ses couvertures

Dans cette section, nous construisons le portefeuille de rentes viagères ainsi que les produits de couverture. Nous rappelons que, selon notre contexte, les rentes sont vendues à des femmes canadiennes âgées de 65 ans au moment de la vente (donc $T = 0$).

Actifs

Certains fonds de retraite offrent des produits que l'on appelle annuité ou rente viagère. Celles-ci sont construites de la façon suivante : les rentiers (acheteurs du produit) donnent un montant d'argent (actif du vendeur) en échange duquel ils recevront un montant annuel fixe (passif du vendeur) jusqu'à leur décès. Le risque associé au produit dépend donc de la durée pour laquelle les paiements seront effectués. Nous supposons dans notre cas que le fonds de retraite offre des contrats où l'annuité est de 1\$ payable à la fin de l'année si la femme a survécu. Comme notre modèle suppose que la totalité des rentières décèdent avant l'âge de 110 ans, nous avons cet âge

comme maximum possible du modèle. L'annuité pour une cohorte âgée de x années a donc pour valeur à $t = 0$ (Fung & al 2015)[12].

$$a_x = \sum_{i=1}^{110-x} B_0(i) \tilde{S}_x(i) \quad (6.3)$$

Où $B_0(i)$ est le facteur d'actualisation au temps i et $\tilde{S}_x(i)$ est donné par $E_0^{RA}[e^{-\int_0^i m_x(v)dv}]$ et correspond donc à l'espérance du taux de survie ajusté pour le risque de la cohorte âgée de x années au temps i . Le total de la valeur des annuités au temps 0 contenues dans le portefeuille est donc donnée par:

$$A = Na_x \quad (6.4)$$

Où N est le nombre de rentières dans le portefeuille. Les taux de survie utilisés sont ceux ajustés pour le risque puisque ce sont des produits offerts par les entreprises qui s'exposent au risque de longévité et qui, par conséquent, veulent être rémunérées pour cela.

Passifs et surplus

Les actifs du portefeuille déterminés, évaluons les passifs futurs. Puisque ceux-ci sont aléatoires, nous devons les estimer. Comme la condition imposée au contrat d'annuité est que les paiements cessent l'année précédant le décès de la rentière, la valeur actuelle des primes dues à la k -ième rentière est alors donnée par :

$$l_k = \sum_{i=1}^{\lfloor \tau_k \rfloor} B_0(i) \quad (6.5)$$

Où $\lfloor \tau_k \rfloor$ est la variable stochastique déterminée par l'année qui précède le décès de la rentière. Celle-ci est obtenue au moyen d'une simulation de Monte Carlo (voir la procédure ci-dessous). Nous reprenons la technique de simulation de Fung et al.(2015) [12]. Le total des engagements dans le portefeuille devient alors :

$$L = \sum_{i=1}^N l_i \quad (6.6)$$

Une fois les passifs estimés, nous pouvons étudier la rentabilité du portefeuille. Les surplus D constituent la distribution des profits et pertes du portefeuille. Ceux-ci sont donnés par $D = A - L$. Comme nous utilisons la méthodologie de Fung & al. (2015) [12], nous observons la distribution

des gains/pertes moyennes, que nous dénotons par \bar{D} . Nous avons donc $\bar{D} = D/N$ (Fung 2015) [12]. De la même façon que Bacinello & al (2010) [3], nous simulons un portefeuille de décès pour chacune des scénarios de taux de mortalité calculés précédemment. Puisqu'il semble naturel de considérer que le processus des intensités de mortalité est aléatoire ainsi que le moment de décès des rentiers, nous nous retrouvons alors dans le cas d'un processus de Poisson inhomogène. La procédure utilisée pour déterminer les scénarios de décès ainsi que la distribution des surplus est la suivante :

- a) Simuler les taux de mortalité $m_x(t)$ pour chacune des périodes. Dans notre cas, comme nous supposons le décès de toutes les rentières à 110 ans et que nous étudions un portefeuille composé uniquement de femmes âgées de 65 ans, alors nous simulons 45 années de taux de mortalité;
- b) Calculer les intensités à chacun des instants t de mortalité qui sont données par $I(t) = \sum_{i=1}^t m_x(i)$ pour les valeurs de $t \in \{1, 2, 3, \dots, T\}$ et ce pour chacune des simulations de taux de mortalité ;
- c) Générer des nombres aléatoires, ϵ suivant une loi exponentielle de paramètre 1. Ils peuvent être obtenus en prenant $-\ln(1 - U)$. Où U est une variable aléatoire uniforme entre 0 et 1. Comme l'hypothèse du portefeuille est de contenir 4000 rentiers, nous allons générer 4000 variables ϵ ;
- d) Pour chacun des nombres aléatoires, ϵ , déterminer la plus grande valeur de t telle que $I(t) < \epsilon$. Cette valeur de t sera le τ_k désiré; c'est-à-dire le moment du décès du k-ième rentier ;
- e) Calculer les valeurs de L_x de l'équation (6.5) ainsi que les valeurs de D et \bar{D} ;
- f) Répéter les étapes b) à e) pour chacune des simulations de taux de mortalité.

Une hypothèse simplificatrice de cette méthodologie est que les taux de mortalité du portefeuille représentent bien ceux de la population. Cette hypothèse est imposée parce que nous n'avons pas de données provenant d'un portefeuille réel de rentes où nous pouvons calibrer le modèle et ensuite utiliser les taux de mortalité prévus pour la partie b) de la simulation et ainsi tester les produits de couverture.

6.3.1 Les S-forwards/S-Swaps

Les forwards de survie (S-forwards) sont des instruments de base et les plus intuitifs lorsqu'il s'agit de transférer le risque d'un portefeuille exposé au risque de longévité à une tierce partie (LLMA 2010)[20]. Ils permettent de réduire les pertes potentielles. En contrepartie, ils réduisent aussi les gains potentiels. Le payoff, P , d'un forward de survie est :

$$P = N(e^{-\int_0^T m_x(v)dv} - K) \quad (6.7)$$

Où K est le taux forward fixé au départ et N est le nombre de personnes dans la cohorte. Puisqu'un contrat forward n'a pas de coût à l'initiation, la relation 6.8 doit absolument être respectée (Fung al. 2015)[12]. Pour qu'il en soit ainsi, K doit être donné par $\mathbb{E}_0^{RA}(e^{-\int_0^T m_x(v)dv}) = \tilde{S}_x(T)$. Nous nous référons à un forward x ans afin de signifier l'âge initial de la cohorte attachée au taux de survie.

$$B_0(T)\mathbb{E}_0^{RA}(e^{-\int_0^T m_x(v)dv} - K) = 0 \quad (6.8)$$

L'expression $e^{-\int_0^T m_x(v)dv}$ représente le taux de survie réalisé de la cohorte âgée de $(x+T)$ années à l'année T . Avec ce type de payoff, une compagnie d'assurance qui est exposée au risque de longévité, c'est-à-dire qu'elle doit payer des rentes annuelles, souhaite être long ce forward. Puisque celui-ci a un payoff positif lorsque le taux de survie est plus grand qu'anticipé. Ce qui compense pour les primes annuelles payées en surplus à celles espérées. Si à l'inverse, le taux de mortalité est plus grand qu'anticipé, la compagnie d'assurance doit payer la contre-partie. Cependant, les passifs en lien avec les rentes qu'elle possède sont réduits.

Ces instruments peuvent être transformés en swap de survie (S-swap). Ces swaps sont un portefeuille de forwards avec différentes échéances mais sur la même cohorte déterminée à l'initiation. Dans le cadre de l'étude, nous utilisons un swap de survie comme instrument de couverture au lieu d'un seul S-forward puisque nous considérons qu'il s'agit du produit le plus approprié pour couvrir le risque des portefeuilles de rentes. Notre choix s'explique par l'appariement entre les paiements des rentes et les flux monétaires provenant du S-swap. Puisqu'il s'agit d'un portefeuille de S-forwards, il est naturel que le taux swap varie dans le temps. De plus, de la même manière que pour le S-forward, le contrat doit avoir une valeur de 0 à

l'initiation. Conséquemment, l'équation 6.8 doit être respectée pour chacune des années (Fung al. 2015)[12] ou, plus directement, chacun des S-forwards constituant le S-swap.

Afin d'obtenir le taux de survie de la population, représenté dans le payoff du swap de survie par $e^{-\int_0^T m_x(v)dv}$, il faut compter le nombre de variables, τ_k , simulées (à l'étape d) de la simulation des décès) qui sont plus grandes que T et diviser ce nombre par N. Il est important de noter que cette procédure fonctionne seulement grâce à l'hypothèse du rapprochement des taux de mortalité du portefeuille et de ceux de la population.

Les surplus d'un portefeuille couvert à l'aide d'un swap sont donc donnés par :

$$D_{swap} = A - L + F_{swap}$$

Un autre aspect important, des produits dérivés, est le calcul de la valeur au marché lors d'une date entre l'initiation et l'échéance du contrat. Puisque les entreprises doivent effectuer le calcul de la valeur de leur portefeuille fréquemment et que les swaps de survie n'y échappent pas, il est important de mentionner le calcul à effectuer lorsque cette évaluation doit être faite (Fung al. 2015) [12].

$$F(t) = B_t(T)(\bar{S}_x(t)\tilde{S}_{x+t}(T) - K)$$

Où $\bar{S}_x(t)$ est le taux de survie cumulatif observé pour la période débutant à 0 et se terminant à t. De plus, il est important d'être en mesure d'évaluer la valeur au marché de ces actifs afin de permettre l'établissement d'un marché liquide. Un nombre plus élevé de transactions permettrait de diminuer la prime de risque liée à la liquidité des contrats swap et de ce fait les rendraient plus attrayants comme instruments de couverture. Les raisons pour lesquelles les gestionnaires voudraient transiger ces contrats sont nombreuses. Un exemple où les gestionnaires pourraient désirer cesser la couverture de leur portefeuille, est la situation où les taux de mortalité associés à leur portefeuille ont augmenté significativement pendant quelques années.

Afin de calculer cette valeur au marché, il faut établir de nouvelles prévisions en fonction de l'âge de la cohorte à cette période. Ce qui indique que x deviendra alors x+t où t est le temps écoulé depuis l'émission du forward.

6.3.2 Les caps de longévité

Les caplets sont des options d'achat (call) européennes sur taux. Les caplets de longévité permettent, si exercés, au détenteur de recevoir un paiement équivalent à la différence entre le dernier taux de mortalité observé et le taux cible de l'option. Les payoffs d'un caplet sont donc donnés par :

$$\max((e^{-\int_0^T m_x(v)dv} - K), 0)$$

Puisque les payoffs de ce produit ne peuvent être négatifs, il est alors normal de devoir déboursier une prime pour obtenir ce produit. La valeur à l'initiation d'un caplet est :

$$\mathbb{E}^{RA}[(e^{-\int_0^T m_x(s)ds} - K)^+]$$

Cette espérance peut être calculée à l'aide d'une simulation de Monte Carlo. La valeur de K est fixée au départ. La valeur de $e^{-\int_0^T m_x(s)ds}$ est déterminée à l'aide des étapes a) à d) de la procédure expliquée à la section 6.3. Cependant, les intensités utilisées dans cette simulation sont celles calculées à partir des taux prévus par le modèle ajusté pour le risque. De cette manière, l'espérance calculée contient une prime pour le risque. Dans le cadre de notre scénario de portefeuille contenant des rentes, nous utilisons un cap qui consiste en un portefeuille de caplets ayant différentes échéances. Connaissant cette valeur, celle d'un cap est donc :

$$\sum_{t=1}^{\hat{T}} V_{\text{caplet}_x}(t, K(t)) \quad (6.9)$$

Où $K(t)$ sont les différents prix de levée des caplets constituant le cap et \hat{T} est la date d'échéance du cap qui, elle, correspond à la date d'échéance la plus lointaine parmi celles des différents caplets. Nous supposons que le cap est acheté au moment de la vente des rentes et que les payoffs sont payés à la fin de chacune des années s'il y a exercice de l'option.

Comme un cap apporte un rendement à un portefeuille de rentes lorsque le taux de survie est plus élevé que prévu, nous croyons que l'insertion de ce produit dans un portefeuille de rentes devrait réduire l'ampleur des pertes potentielles. Dans la situation où les taux de mortalité seraient plus élevés que prévus, le portefeuille serait en position de gains moyens. Puisque les payoffs du cap ne peuvent être négatifs, les gains du portefeuille ne seraient

qu'affectés par la prime payée pour la couverture. Ce qui laisse place à des gains plus importants si le taux de mortalité de la cohorte devait augmenter significativement comparativement au swap. Conséquemment, nous nous attendons à ce que l'asymétrie des rendements du portefeuille soit positif. Les surplus du portefeuille couvert à l'aide d'un cap sont donnés par :

$$D_{Cap} = A - L + F_{Cap} - C_{cap} \quad (6.10)$$

Où F_{Cap} correspond au total des payoffs du cap et C_{Cap} correspond au coût du cap.

6.4 Résultats et discussion

Dans cette section, nous simulons les surplus d'un portefeuille de rentes et employons 2 stratégies différentes de couverture. Cette simulation est réalisée à l'aide de la méthodologie décrite à la section 6.3. Nous comparons les statistiques élémentaires des distributions des surplus entre ceux non-couverts et les surplus couverts à l'aide de différents instruments. Les instruments de couvertures étudiés sont le swap et le cap. Plusieurs échéances sont testées afin d'étudier l'impact d'une stratégie de couverture s'étalant sur une plus longue période. Les résultats sont présentés aux tableaux 6.1 et 6.2. On remarque, grâce aux mesures de risque (VaR et ES), que les produits de couverture effectuent correctement leur travail. Les graphiques 6.2, 6.3 et 6.4 permettent de le constater visuellement. On observe aussi que plus la période de couverture est longue, plus ils réduisent ces mesures. Afin de bien expliquer le phénomène, nous explicitons le portefeuille utilisant le swap comme outil de couverture. Les surplus du portefeuille sont donnés par:

$$\begin{aligned} D_{swap} &= \mathbb{E} \left(\underbrace{A}_{\text{actif}} - \underbrace{L}_{\text{Engagements}} + \underbrace{F_{swap}}_{\text{couverture}} \right) \\ D_{swap} &= \mathbb{E} \left[N \sum_{i=1}^{110-x} B_0(i) \tilde{S}_x(i) - \sum_{k=1}^N \sum_{i=1}^{\lfloor \tau_k \rfloor} B_0(i) + N \sum_{i=1}^{\hat{T}} B_0(i) (e^{\int_0^i m_x(s) ds} - K(i)) \right] \\ &= N \sum_{i=1}^{110-x} B_0(i) \tilde{S}_x(i) - \mathbb{E} \left[\sum_{k=1}^N \sum_{i=1}^{\lfloor \tau_k \rfloor} B_0(i) \right] + \mathbb{E} \left[N \sum_{i=1}^{\hat{T}} B_0(i) (e^{\int_0^i m_x(s) ds}) \right] \\ &\quad - N \sum_{i=1}^{\hat{T}} B_0(i) K(i) \end{aligned}$$

Comme les portions fixes du portefeuille de forwards sont données par $\tilde{S}_x(i)$ nous pouvons remplacer $K(i)$. Nous obtenons alors :

$$D_{swap} = N \sum_{i=1}^{110-x} B_0(i) \tilde{S}_x(i) - \mathbb{E} \left[\sum_{k=1}^N \sum_{i=1}^{\lfloor \tau_k \rfloor} B_0(i) \right] + \mathbb{E} \left[N \sum_{i=1}^{\hat{T}} B_0(i) (e^{\int_0^i m_x(s) ds}) \right] - N \sum_{i=1}^{\hat{T}} B_0(i) \tilde{S}(i)$$

En posant $\hat{T} = 110 - x$, nous retrouvons :

$$D_{swap} = \mathbb{E} \left[- \sum_{k=1}^N \sum_{i=1}^{\lfloor \tau_k \rfloor} B_0(i) \right] + \mathbb{E} \left[N \sum_{i=1}^{\hat{T}} B_0(i) e^{\int_0^i m_x(s) ds} \right] \quad (6.11)$$

Il reste donc à montrer que l'espérance de l'équation 6.11 est de 0. Nous pouvons exprimer L comme le nombre de paiements effectués au moment i . Ce qui se traduit par

$$\mathbb{E}[L_i] = \mathbb{E} \left[(N - n_i) B_0(i) \right], \forall i = \{1 \dots T\}$$

Avec n_i une variable aléatoire qui est donnée par le nombre de décès observés au moment i . Du côté du forward, nous pouvons approximer $e^{-\int_0^i m_x(s) ds}$ par $\frac{N-n_i}{N}$. Nous utilisons la même notation que celle des engagements puisque par hypothèse les décès dans le portefeuille ont la même distribution que ceux observés dans la population. On se retrouve donc avec :

$$\mathbb{E} \left[- (N - n_i) B_0(i) + N B_0(i) \left(\frac{N - n_i}{N} \right) \right] = 0$$

Selon Fung & al. (2015), il est moins justifié de couvrir le portefeuille pour une période de plus de 30 ans puisque la majorité des rentières seront décédées. Conséquemment, à ce moment, le risque de longévité contenu dans ce portefeuille pour cette cohorte sera beaucoup plus faible. Selon notre calibration du modèle, la barrière d'inefficacité de la couverture se situe plutôt autour de 35 ans. Cette légère augmentation provient de l'amélioration temporelle des taux de mortalité entre les 2 études. Selon notre calibration, nous estimons qu'une femme canadienne sur 5, âgée aujourd'hui de 65 ans, atteindra l'âge de 100 ans. À ce moment, il ne restera que 800 personnes vivantes, en moyenne, sur les 4000 du départ à qui il faudra payer des rentes. Avec un nombre restreint de rentier, le montant à déboursier pour les produits de couvertures est trop élevé pour l'avantage qu'ils apportent.

Au niveau de la couverture à l'aide d'un cap, nous observons que les performances au niveau de la réduction des risques sont moins bonnes comparativement à celle effectuée à l'aide d'un swap. Cependant, la moyenne des surplus est plus grande. Comme prévu, le skewness du portefeuille couvert à l'aide d'un cap est positif. Ces résultats sont en concordances avec ceux obtenus par Fung & al (2015). En somme, ces produits offrent une certaine flexibilité quant au niveau de risque qu'un gestionnaire de portefeuille pourrait désirer.

Portefeuille	Moyenne	Écart-type	Skewness	Kurtosis	VaR99	ES99
Sans couverture	0.3471	0.3399	-0.0028	2.9383	-0.4453	-0.5421
Avec swap	0.2758	0.2475	-0.0771	2.9261	-0.3129	-0.3876
Avec cap	0.3043	0.2951	0.2655	3.0627	-0.3822	-0.4821

Tableau 6.1: Statistiques des surplus selon le type de portefeuille $T=20$

Portefeuille	Moyenne	Écart-type	Skeness	Kurtosis	VaR99	ES99
Sans couverture	0.3471	0.3399	-0.0028	2.9383	-0.4453	-0.5421
Avec Swap	0.1513	0.1216	-0.2999	3.0448	-0.1515	-0.2073
Avec cap	0.2226	0.2386	0.9236	3.8779	-0.3324	-0.4133

Tableau 6.2: Statistiques des surplus selon le type de portefeuille $T=30$

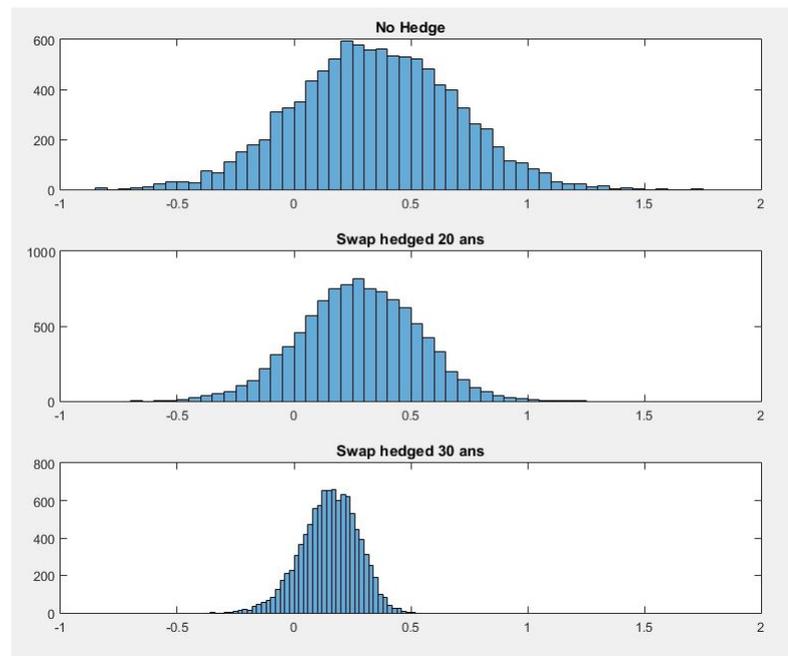


Figure 6.2: Distribution des surplus selon la couverture

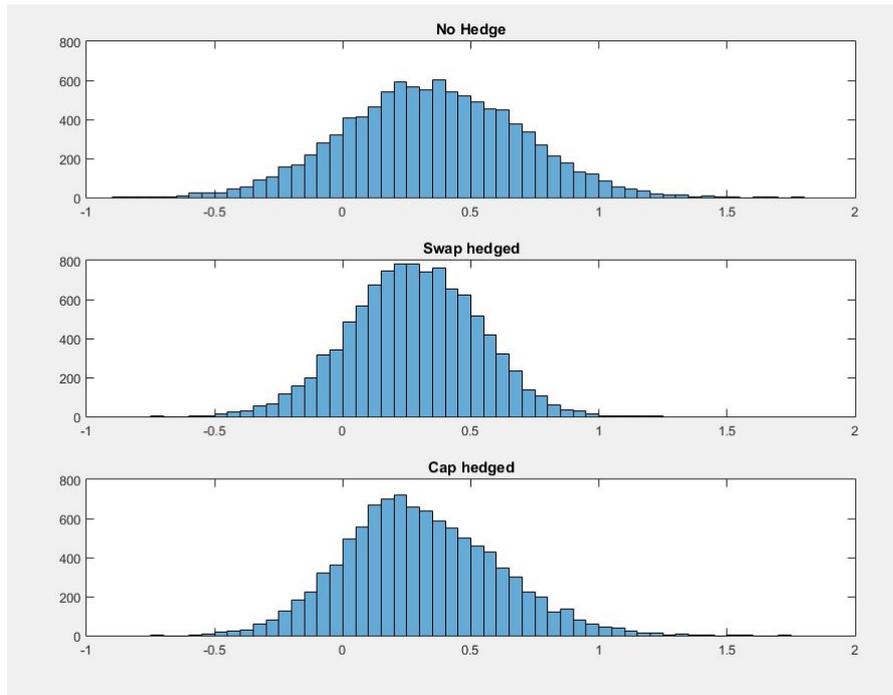


Figure 6.3: Surplus selon une protection de 20 ans

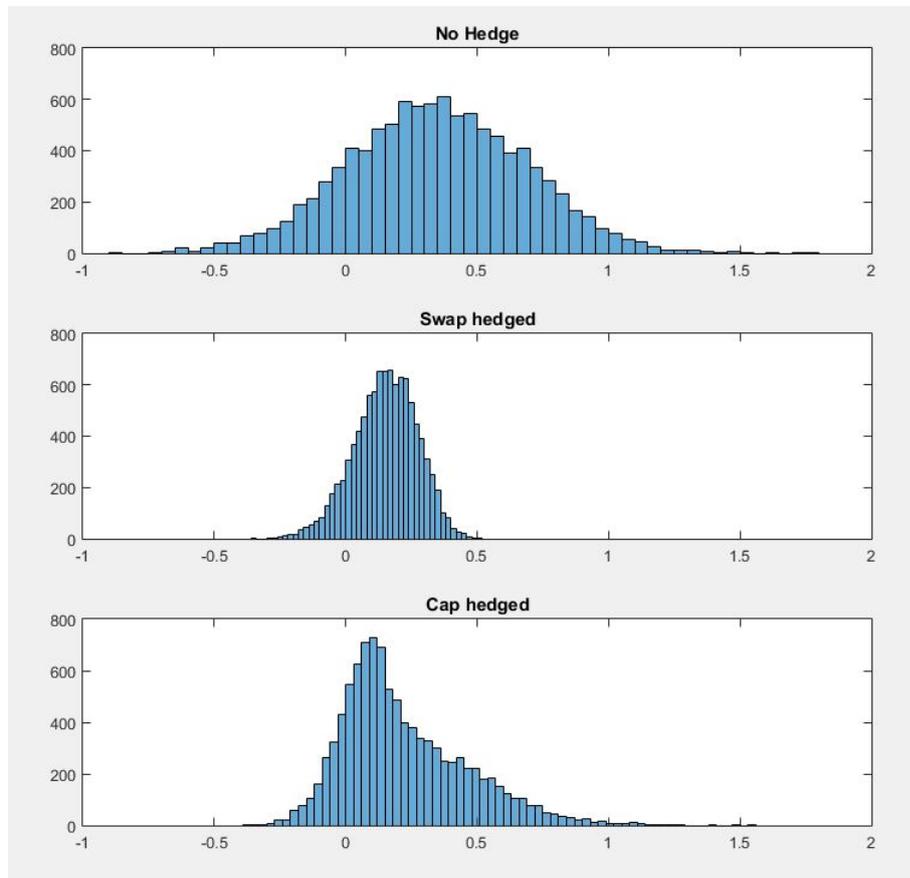


Figure 6.4: Surplus selon une protection de 30 ans

7. Portefeuille multiples cohortes

À la suite de la démonstration de la possibilité de couvrir un portefeuille de rentes contenant une unique cohorte d'âge, nous abordons la couverture d'un portefeuille de rentières appartenant à différentes cohortes à l'aide d'un swap sur une seule cohorte. Nous faisons cela dans le but de rapprocher le portefeuille de la réalité. Nous nous attendons à ce qu'il soit possible de mettre en place une telle stratégie puisque les améliorations des taux de mortalité ne sont pas propres à une seule cohorte. Elles devraient donc être du même ordre pour un groupe d'âge relativement près. Dans le cas où cette couverture ne serait pas efficace, nous tenterons de trouver le nombre minimal de swaps sur différentes cohortes qui sont nécessaires afin d'obtenir des couvertures efficaces.

7.1 Mise en situation

Nous construisons un portefeuille de 20 000 rentières. Soit 4000 femmes appartenant à chacune des cohortes suivantes: 65, 66, 67, 68 et 69 ans. Les valeurs de chacune des annuités contenues dans ce portefeuille sont déterminées à l'aide de l'équation 6.3. Comme nous avons 5 cohortes différentes, nous obtenons donc 5 prix différents. Nous supposons que la totalité des rentières décèdent avant 110 ans. La méthodologie employée est la même que celle pour le portefeuille constitué d'une seule cohorte. Nous croyons qu'un effet de diversification devrait influencer les surplus du portefeuille. Cette hypothèse est basée sur le fait qu'il existe une matrice de covariance non-diagonale entre les cohortes. Ce qui nous indique qu'il existe une certaine corrélation entre les résidus pour différents âges. Afin de tester cette hypothèse, nous vérifions l'efficacité de différentes couvertures utilisant des portefeuilles de swaps. Les surplus du portefeuille sont donnés par la moyenne pondérée des surplus de chacune des cohortes. Les poids de cette moyenne sont donnés par le quotient entre le nombre de femmes

appartenant à la cohorte et le nombre total de femme dans la totalité du portefeuille. Nous notons \bar{D} comme la moyenne des surplus espérés qui, dans notre cas, est donnée par la moyenne arithmétique puisque nous supposons que chacune des cohortes sont constituées du même nombre de femmes. Le nombre de swaps dans un portefeuille est donné par j et x_j représente le j -ième âge de l'ensemble de la cohorte dont les payoffs du swap dépendent. Nous considérons le payoff moyen du portefeuille de swap \bar{S} . Selon ces conditions, \bar{D} est donné par:

$$\begin{aligned} \bar{D} &= \bar{A} - \bar{L} + \bar{S} \\ &= \frac{N}{5} \sum_{x=65}^{69} a_x - \frac{1}{5} \sum_{k=1}^N \sum_{x=65}^{69} l_{x,k} + \frac{N}{j} \left(\sum_{x=x_1}^{x_j} \sum_{i=1}^{\hat{T}} B_0(i) (e^{-\int_0^i m_x(s) ds} - \tilde{S}_x(i)) \right) \end{aligned}$$

7.2 Résultats et discussion

Nous présentons d'abord les statistiques descriptives des surplus pour le portefeuille simple cohorte comparativement à celles du portefeuille multi-cohortes. Celles-ci sont présentées au tableau 7.1. Nous remarquons qu'il y a une diminution de l'écart-type qui est créé par l'ajout de cohortes dans le portefeuille. Ce qui entraîne une réduction des mesures de risque. Au niveau de la moyenne la baisse de celle-ci s'explique par l'inclusion d'autres cohortes dans le portefeuille.

Statistiques	Simple cohorte	Multi-cohortes
Moyenne	0.3471	0.3112
Écart-type	0.3399	0.2066
Skewness	-0.0028	-0.0663
Kurtosis	2.9383	2.9833
VaR99	-0.4453	-0.1700
ES99	-0.5421	-0.2517

Tableau 7.1: Statistiques descriptives des portefeuilles simples et multi-cohortes

Nous présentons au tableau 7.2 une comparaison entre les statistiques du portefeuille multi-cohortes non-couvert et différents portefeuilles couverts utilisant un portefeuille de swaps comme instrument de couverture. Nous présentons les couvertures similaires, ainsi que les plus efficaces. Ceci dans

le but d'alimenter la discussion qui suit. Dans ce tableau, nous utilisons la notation S_x^T afin d'expliciter les cohortes contenues dans le portefeuille de swaps. Par exemple, $S_{65,67}^{20}$ signifie que nous avons couvert le portefeuille à l'aide d'un portefeuille constitué de deux swaps de 20 ans. L'un d'entre eux ayant comme partie fixe les taux de survie des femmes âgées de 65 ans ajustés pour le risque. Le deuxième est celui dont la partie fixe est donnée par les taux de survie ajustés des femmes âgées de 67 ans. La première observation dont nous faisons mention est que la couverture à l'aide d'un seul swap n'est pas efficiente selon notre modèle et ce peu importe l'âge de la cohorte sous-jacente. Ceci vient infirmer la première hypothèse que nous avons émise concernant la couverture du portefeuille à l'aide d'un seul swap. C'est pour cette raison que nous nous concentrons davantage sur les portefeuilles de swaps. Nous observons que dans le cas d'une stratégie de couverture sur une période de 20 ans, un portefeuille constitué de 3 swaps semble être suffisant pour offrir des résultats appréciables. Il est important de souligner que plus la période de couverture est grande, plus le nombre de swaps nécessaires à la stratégie de protection est grand. Ceci montre que plus le temps avance, moins les variations de taux de mortalité sont du même ordre pour des cohortes d'un âge rapproché. Ce qui semble être une faiblesse du modèle retenu pour l'étude. Il est vérifié que dans chacune des périodes, le portefeuille constitué de 5 swaps offre toujours la meilleure couverture. Ce qui est évident puisque dans ce cas, les swaps sont liés à leur cohorte et ce dans leurs proportions.

Couverture	Moyenne	Écart-type	Skewness	Kurtosis	VaR99	ES99
Sans couverture	0.3112	0.2066	-0.0663	2.9833	-0.1700	-0.2517
S_{65}^{10}	0.2794	0.1877	-0.0532	3.0162	-0.1627	-0.2371
$S_{65,67}^{10}$	0.2777	0.1791	-0.0861	2.9780	-0.1430	-0.2127
$S_{65,67,69}^{10}$	0.2759	0.1748	-0.0902	2.9975	-0.1395	-0.2031
S_{65}^{20}	0.2406	0.1770	-0.0303	3.0687	-0.1770	-0.2439
$S_{65,67}^{20}$	0.2379	0.1572	-0.0919	3.0380	-0.1332	-0.1956
$S_{65,67,69}^{20}$	0.2344	0.1480	-0.1056	3.0305	-0.1145	-0.1741
$S_{67,68,69}^{20}$	0.2314	0.1454	-0.0859	3.0158	-0.1110	-0.1660
S_{65}^{30}	0.1155	0.2033	-0.0218	3.0251	-0.3638	-0.4290
$S_{65,67}^{30}$	0.1140	0.1476	-0.0335	2.9466	-0.2374	-0.2786
$S_{65,67,69}^{30}$	0.1117	0.1175	-0.0538	3.0104	-0.1704	-0.2097
$S_{66,67,68,69}^{30}$	0.1101	0.0776	-0.0666	3.0438	-0.0714	-0.1025
$S_{65,66,67,68,69}^{30}$	0.1112	0.0617	-0.2745	3.1195	-0.0412	-0.0686

Tableau 7.2: Statistiques des surplus selon le portefeuille de swaps utilisé

Au niveau du cap, nous utilisons C_x^T comme notation. Les variables T et x ont les mêmes significations que pour le swap. Nous présentons les résultats en lien avec les mêmes cohortes que celle du swap afin de faciliter les comparaisons.

Couverture	Moyenne	Écart-type	Skewness	Kurtosis	VaR99	ES99
Sans couverture	0.3233	0.2167	-0.0165	3.0156	-0.187	-0.2606
C_{65}^{10}	0.3175	0.2106	0.00296	3.0157	-0.1774	-0.2476
$C_{65,67}^{10}$	0.3172	0.2097	0.0079	3.0219	-0.1732	-0.2441
$C_{65,67,69}^{10}$	0.3164	0.2089	0.0111	3.022	-0.1714	-0.243
C_{65}^{20}	0.2761	0.1922	0.1006	3.1457	-0.1596	-0.2213
$C_{65,67}^{20}$	0.2745	0.1856	0.165	3.1623	-0.1396	-0.1975
$C_{65,67,69}^{20}$	0.2724	0.1816	0.1858	3.1914	-0.132	-0.1857
$C_{67,68,69}^{20}$	0.2699	0.1798	0.1877	3.1589	-0.1295	-0.1792
C_{65}^{30}	0.2017	0.1827	0.0373	3.034	-0.2202	-0.2838
$C_{65,67}^{30}$	0.202	0.1605	0.1924	3.1167	-0.1532	-0.2043
$C_{65,67,69}^{30}$	0.1978	0.1493	0.3341	3.2177	-0.1129	-0.1583
$C_{66,67,68,69}^{30}$	0.1956	0.143	0.4806	3.2567	-0.0832	-0.1125
$C_{65,66,67,68,69}^{30}$	0.1968	0.1422	0.5592	3.32	-0.0694	-0.0953

Tableau 7.3: Statistiques des surplus selon le portefeuille de caps utilisé

Au niveau du cap, les conclusions sont similaires à celles du swaps.

C'est-à-dire qu'afin d'obtenir une couverture de portefeuille efficiente, la stratégie doit être déployée sur une période minimale d'environ 20 ans. De plus, on remarque qu'un portefeuille constitué de 3 caps semble être suffisant afin de couvrir adéquatement le portefeuille de rentes. Il semble que la même relation entre la date d'échéance ainsi que le nombre de caps à inclure dans la stratégie de couverture tient. Ce qui montre que les relations d'interdépendances entre les âges semblent être le point faible du modèle. Toutefois, en dépit de cette faiblesse, il est clair que les produits effectuent leur travail adéquatement. Ce qui démontre la pertinence d'en développer un marché.

Conclusion

Dans ce mémoire, nous étudions la modélisation des taux de mortalité ainsi que l'effet potentiel de certains produits dérivés de longévité s'ils étaient disponibles sur le marché. Ces produits fournissent une couverture pour les améliorations des taux de mortalité qui sont au delà des prévisions actuarielles. Les sous-estimations d'amélioration pourraient potentiellement avoir des conséquences importantes en lien avec la solvabilité future des fonds de retraites.

Les hypothèses de recherche émises sont confirmées en partie. Nous observons que le modèle de Will-Sherris modélise très bien les données canadiennes. Les prévisions effectuées à partir de la calibration semblent être précises pour une période de temps d'environ 20 ans. En ce qui concerne la couverture d'un portefeuille à l'aide de produits de longévité, nos résultats suggèrent qu'il est possible de couvrir un portefeuille constitué d'une seule cohorte à l'aide d'un swap dont le payoff est dépendant du taux de survie observé de cette cohorte. Cependant, au niveau d'un portefeuille constitué de plusieurs cohortes, nos résultats suggèrent qu'il est plus adéquat d'utiliser un portefeuille de swaps contenant minimalement 3 cohortes différentes. Au niveau du cap, les résultats sont similaires. C'est-à-dire qu'afin d'obtenir une couverture efficace, la stratégie doit, elle aussi, être déployé sur une période d'environ 20 ans. La grande différence entre les 2 stratégies réside au niveau de la distribution des surplus. Nous avons observé que celle des surplus du portefeuille couvert à l'aide d'un cap est asymétrique vers la droite. Alors que pour le portefeuille couvert à l'aide d'un swap, la distribution semble rester centrée. Cependant, nous pouvons obtenir des couvertures appréciables avec un nombre restreint de swaps/caps ce qui est encourageant. Des recherches plus poussées entre autres sur des modèles tenant mieux compte de l'interdépendance des taux de mortalité entre les cohortes devraient être développés. L'objectif étant de vérifier s'il est pos-

sible de couvrir un portefeuille de rentes à l'aide d'un seul swap. Dans un autre ordre d'idées, afin de développer un marché liquide, un indice sur la mortalité devrait être développé afin de permettre aux investisseurs de faire l'évaluation de ces produits continuellement et non pas une seule fois annuellement.

Bibliographie

- [1] Alders, M., & De Beer, J. (2005). An expert knowledge approach to stochastic mortality forecasting in the Netherlands. In N. Keilman (Ed.), Swedish Social Insurance Agency, Stockholm. Perspectives on Mortality forecasting (Vol. II. Probabilistic Models, p. 39-64).
- [2] Artzner, P., Delbaen, F., Eber, J. M., Heath, D., 1999, Coherent measures of risk, *Mathematical finance* Vol. 9 (3), July 1999, p.203-228
- [3] Bacinello, A.R., Biffis, E., Millossovich, P., 2010, Regression-based algorithms for life insurance contracts with surrender guarantees, *Quantitative Finance*, November 2010, v. 10, iss. 9, p. 1077-1090
- [4] Barrieu, P., Bensusan, H., El Karoui, N., Hillairet, C., Loisel, S., Ravanelli, C., Salhi, Y., 2012, Understanding, modelling and managing longevity risk: key issues and main challenges, *Scandinavian Actuarial Journal*, 3, p. 203-231
- [5] Bell, W., 1997, comparing and assessing time series methods for forecasting age-specific fertility and mortality rates, *Journal of official Statistics*, 13(3), p. 279-303
- [6] Black, F., Scholes, M., 1973, The Pricing of Options and Corporate Liabilities, *Journal of Political Economy*. Vol. 81 Issue 3, p637.
- [7] Booth, H., Maindonald, J., Smith. L., 2002. Applying Lee-Carter under conditions of variable mortality decline. *Population studies*, 56, p. 325-336
- [8] Booth, H., Tickle, L., 2008, Mortality modelling and forecasting: A review of methods, ADSRI Working Paper No.3
- [9] Cairns, A. J. G., 2004, *Interest Rate Models: An Introduction*, Princeton University Press (2004).

- [10] Chena, H., Cummins, J. D., 2010. Longevity bond premiums: The extreme value approach and risk cubic pricing: *Insurance: Mathematics and Economics* 46 (2010), p. 150-161
- [11] Dahl, M. (2004). Stochastic Mortality in Life Insurance: Market Reserves and Mortality-linked Insurance contracts, *Insurance: Mathematics and Economics* 35 (2005), p. 113-136.
- [12] Fung, M., Ignatieva, K., Sherris, M., 2015. Managing Systematic Mortality Risk in Life Annuities: An Application of Longevity Derivatives UNSW Business School Research Paper No. 20015ACTL04
- [13] Gompertz, B. (1825). On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies, *Philosophical Transactions of Royal Society, (Series A)* 115, p.513-583.
- [14] Hollmann, F. W., Mulder, T. J., & Kallan, J. E. (2000). Methodology and assumptions for the population projections of the United States: 1999 to 2100: Working Paper 38, Population Division, U.S. Bureau of the Census.
- [15] Joanis, M., 2016. Essai index de longévité. Université de Sherbrooke
- [16] Johnson, R.A., and Wichern, D.W., 2007. Applied Multivariate Statistical Analysis. 6th ed. Prentice Hall, New York. P.360
- [17] Lee, R. D., 2000. The Lee-Carter Method for Forecasting Mortality, with Various Extensions and Applications. *North American Actuarial Journal*, (4) 1, p. 80-91
- [18] Lee, R. D., and L. R. Carter, 1992, Modeling and Forecasting U.S. Mortality, *Journal of the American Statistical Association*, 87 P. 659-675.
- [19] Li H., De Waegenareb A., Melenbergb B., 2015. The choice of sample size for mortality forecasting: A Bayesian learning approach. *Insurance: Mathematics and Economics* 63 p. 153-168
- [20] Life & longevity markets association, 2011 A framework for pricing longevity exposure, www.llma.org

- [21] Markowitz, H., 1952, Portfolio selection, The journal of finance, Vol. 7 No. 1 (Mar. 1952), p.77-91
- [22] Meyricke, R., Sherris, M., 2014, Longevity risk, cost of capital and hedging for life insurers under Solvency II. Insurance: Mathematics and Economics 55, p. 147-155
- [23] Mitchell, D., Brockett, P., Mendoza-Arriaga, R., and K. Muthuraman. 2013. Modeling and Forecasting Mortality Rate. Insurance: Mathematics and Economics 52: p. 275-285
- [24] Pelsser, A., 2008. On the applicability of the Wang transform for pricing financial risks. ASTIN bulletin 38, p. 171-181
- [25] Stuart J. Pocock, D. G., Shirley A. A. Regression of Area Mortality Rates on Explanatory Variables: What Weighting is Appropriate?, Journal of the Royal Statistical Society. Series C (Applied Statistics) Vol. 30, No. 3 (1981), p. 286-295
- [26] Tabeau, E., Ekamper, P., Huisman, C., Bosch, A., 2001, *Predicting Mortality from Period, Cohort or Cause-specific Trends: A Study of Four European Countries*, Forecasting mortality in developed countries: insights from a statistical, demographic and epidemiological perspective (159-187). Kluwer Academic Publishers, Dordrecht.
- [27] Vartiainen E, Puska P, Pekkanen J, Tuomilehto J, Jousilahti P (1994) Changes in risk factors explain changes in mortality from ischemic heart disease in Finland. BMJ 309 p. 23-27
- [28] Waldron, H. (2005). Literature review of long-term mortality projections. Social Security Bulletin, 66(1), p. 16-30
- [29] Yang, S., Wang, C., 2012, Pricing and securitization of multi-country longevity risk with mortality dependence. Insurance: Mathematics and Economics 52, p. 157-169
- [30] Wang, S., 2000, A class of distortion operators for pricing financial and insurance risks. The journal of risk and insurance, Vol 67 p. 15-36
- [31] White, K., 2002, Longevity advances in high-income countries, 1955-96, Population and development review, 28(1), p. 59-76

- [32] Wilmoth, J., 1993. Computational methods for fitting and extrapolating the Lee-Carter model of mortality change. Technical Report, Department of Demography, University of California, Berkeley
- [33] Wills, S., Sherris, M., 2011. Integrating financial and demographic longevity risk models: an Australian model for financial applications. UNSWAustralian School of Business Research Paper No. 2008ACTL05.
- [34] Yue, J., Yang, S., Huang H., A Study of the Lee-Carter Model with Age-Shifts, 2008
- [35] Zahra Boutaggout, F., modélisation et management du risque de longévité: Application des instruments financiers de longévité, 2017, Université de Sherbrooke
- [36] <http://www.phac-aspc.gc.ca/cphorsphc-respcacsp/2008/fr-rc/cphorsphc-respcacsp05b-fra.php>
- [37] <http://www.mortality.org/>