



HAL
open science

Outils théoriques et opérationnels adaptés au contexte de l'assurance vie en Afrique subsaharienne francophone - Analyse et mesure des risques liés à la mortalité

Aymric Kamega

► To cite this version:

Aymric Kamega. Outils théoriques et opérationnels adaptés au contexte de l'assurance vie en Afrique subsaharienne francophone - Analyse et mesure des risques liés à la mortalité. Gestion des risques [q-fin.RM]. Université Claude Bernard - Lyon I, 2011. Français. NNT: . tel-00654549

HAL Id: tel-00654549

<https://theses.hal.science/tel-00654549>

Submitted on 22 Dec 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Université Claude Bernard Lyon 1



UNIVERSITÉ DE LYON

N° d'ordre : 321-2011

Année 2011

THÈSE DE L'UNIVERSITÉ DE LYON

Présentée devant

UNIVERSITÉ CLAUDE BERNARD LYON 1

ÉCOLE DOCTORALE DE SCIENCES ÉCONOMIQUES ET DE GESTION (486)

Pour l'obtention du

DIPLÔME DE DOCTORAT
(arrêté du 7 août 2006)

Présentée par

Aymric KAMEGA

Outils théoriques et opérationnels adaptés au contexte de l'assurance vie en Afrique subsaharienne francophone

—

Analyse et mesure des risques liés à la mortalité

Directeur de thèse : Frédéric PLANCHET

Composition du jury :

Monsieur Michel BERA (Conservatoire National des Arts et Métiers)
Monsieur Arthur CHARPENTIER (Université du Québec, Canada)
Monsieur Olivier LOPEZ (Université Paris 6)
Madame Véronique MAUME-DESCHAMPS (Université Lyon 1)
Monsieur Abder OULIDI (Université Catholique de l'Ouest)
Monsieur Frédéric PLANCHET (Université Lyon 1)
Monsieur Marc QUINCAMPOIX (Université de Bretagne Occidentale)

SOMMAIRE GENERAL

SOMMAIRE GÉNÉRAL.....	3
INTRODUCTION GÉNÉRALE.....	7
Choix du sujet de l'étude	7
Contexte de l'étude	7
Organisation de l'étude	8
Bibliographie	11
PARTIE I ANALYSE DE LA MORTALITÉ DANS LA ZONE : PRÉSENTATION DU CONTEXTE ET DES DÉFIS TECHNIQUES	13
Chapitre 1 Présentation du marché de l'assurance vie en Afrique subsaharienne francophone	15
1. Organisation du marché : présentation des trois dispositifs d'assurance sur la vie (famille et communauté, assureurs privés, institutions publiques).....	16
2. Présentation du contexte socioéconomique pour les assureurs vie du marché CIMA	27
3. Présentation des difficultés techniques et commerciales pour les assureurs vie du marché CIMA.....	37
4. Synthèse et conclusion.....	45
Bibliographie	49
Chapitre 2 Évolution technique du marché CIMA : construction de tables de mortalité d'expérience pour la zone.....	53
1. Présentation des données collectées.....	54
2. Construction des tables d'expérience.....	62
3. Contrôle des tables d'expérience construites et comparaisons	75
4. Synthèse et conclusion (et présentation de défis techniques complémentaires sur la mortalité).....	82
Bibliographie	84
Annexe A : retraitement des données.....	85
Annexe B : mise à jour du périmètre	89
PARTIE II MORTALITÉ ACTUELLE : MODÉLISATION DE L'HÉTÉROGÉNÉITÉ ET MESURE DES RISQUES ASSOCIÉS DANS LA ZONE.....	91
Chapitre 3 Mesure du risque d'estimation en présence d'hétérogénéité... 93	
1. Introduction.....	94

2. Intervalles et bandes de confiance	95
3. Mesure du risque d'estimation sur les taux ajustés	99
4. Mesure du risque d'estimation sur les provisions	108
5. Évolution du risque d'estimation en présence d'hétérogénéité.....	114
6. Synthèse et conclusion.....	122
Notations des taux de décès	124
Bibliographie	125
Chapitre 4 Réduction du risque d'estimation en présence d'hétérogénéité	
127	
1. Introduction.....	128
2. Choix du modèle à variables explicatives observables	130
3. Modèle de Cox : ajustement et simulation des taux de décès	133
4. Modèle de Lin et Ying : ajustement et simulation des taux de décès...	138
5. Comparaison des ajustements des modèles de Cox et de Lin et Ying et <i>backtesting</i> sur l'hétérogénéité.....	141
6. Comparaison des risques d'estimation entre les modèles indépendants pour chaque sous-population et ceux intégrant l'hétérogénéité à partir de facteurs observables	148
7. Synthèse et conclusion.....	153
Bibliographie	156
PARTIE III MORTALITÉ FUTURE : MODÉLISATION PROSPECTIVE DE LA	
MORTALITÉ ET MESURE DES RISQUES ASSOCIÉS DANS LA ZONE	159
Chapitre 5 Modélisation prospective en l'absence de données sur les	
tendances passées et mesure des risques associés.....	161
1. Introduction.....	162
2. Choix du modèle pour projeter la mortalité	164
3. Ajustement du modèle de Bongaarts	176
4. Application : anticipation de la mortalité d'expérience future.....	183
5. Synthèse et conclusion.....	191
Bibliographie	193
Chapitre 6 Modélisation prospective en présence de données sur les	
tendances passées et mesure des risques associés.....	197
1. Introduction.....	198
2. Construction de tables de mortalité prospectives.....	200
3. Encadrement des taux de décès bruts au titre des fluctuations d'échantillonnage.....	204
4. Encadrement des tables de mortalité au titre des fluctuations d'échantillonnage (risque d'estimation).....	207

5. Impact des fluctuations d'échantillonnage dans les évaluations de l'assureur (risque d'estimation)	213
6. Comparaison du risque d'estimation et du risque d'avis d'expert	215
7. Synthèse et conclusion	217
Bibliographie	219
SYNTHÈSE ET CONCLUSION GÉNÉRALE.....	221
Synthèse de l'étude	221
Perspectives pour la recherche future : choix du niveau de segmentation pertinent pour modéliser la mortalité	222
Perspectives pour la recherche future : mesure et gestion des risques associés à la modélisation de la mortalité	224
Perspectives pour la recherche future : autres modélisations et autres risques 225	
Bibliographie	228
SIGLES	229
BIBLIOGRAPHIE GÉNÉRALE.....	231

INTRODUCTION GENERALE

Choix du sujet de l'étude

Le rapport entre les cotisations officielles en assurance vie et le PIB officiel est égal à moins de 0,5 % en Afrique subsaharienne francophone, contre plus de 3 % en Afrique en général et plus de 4 % au niveau mondial (*cf.* Kone [2007a]). Sur ces bases, il existe donc de bonnes raisons de penser que l'assurance vie dans la région dispose de marges de progression importantes.

Toutefois, lorsque l'on évoque les perspectives de développement du marché de l'assurance vie dans la région, les opinions convergent généralement sur le fait que les contextes économique, social, politique et culturel n'y sont pas favorables. Ces jugements sont souvent accompagnés de nombreux arguments : absence de stabilité politique et économique, faible niveau de revenus par habitant, existence de barrières religieuses ou culturelles, insuffisances dans les opportunités de placements financiers, absence de confiance réciproque entre les assureurs, les assurés et les intermédiaires, poids significatif de l'économie informelle, faible niveau de bancarisation, risques sur la durée de vie humaine importants et mal maîtrisés, insuffisances de personnel qualifié au sein des organismes assureurs, manque de données fiables, etc.

Ces caractéristiques sont toutes fondées, et sont reprises en détail dans le premier chapitre de ce travail (sur ce point, *cf.* également Kone [2007a], Yeatman [2007], Ziguélé [2008]). En pratique toutefois, ce premier chapitre ne se contente pas d'énumérer les éléments de contexte défavorables au marché ; il s'attache également à rappeler qu'il ne s'agit que d'éléments de contexte défavorables, et non rédhibitoires.

En d'autres termes, si les spécificités du marché local freinent son développement actuel, elles ne traduisent en rien une incompatibilité avec son développement futur. Bien au contraire, ces spécificités pourraient se révéler être d'excellents leviers de développement. Pour cela, des solutions innovantes et adaptées au contexte sont toutefois nécessaires. Une réflexion en vue de fournir des outils théoriques et opérationnels adaptés au marché d'Afrique subsaharienne francophone est donc indispensable. La présente étude s'inscrit dans ce cadre.

Contexte de l'étude

Sensible à cette nécessité de disposer d'outils adaptés pour développer le marché de l'assurance vie, l'autorité de contrôle prudentiel de la région¹, associée à la fédération des sociétés d'assurance, a organisé en 2007 les états généraux de l'assurance vie. À l'issue de ces travaux, plusieurs résolutions ont été présentées pour favoriser le développement de l'assurance vie dans la région. Parmi ces résolutions, on note par exemple (*cf.* Kone [2007b]) : le développement de formations adaptées pour les (futurs) collaborateurs, le développement des incitations fiscales, l'instauration de régimes de retraite

¹ Nous verrons dans ce travail qu'il existe une autorité de contrôle unique pour la majorité des pays d'Afrique subsaharienne francophone.

complémentaires aux profits de toutes les populations (y compris du secteur informel), le développement des opportunités de placements ou encore la mise à jour des tables de mortalité réglementaires.

En particulier, sur ce dernier point, le principe de tables de mortalité d'expérience a été adopté en remplacement des tables de mortalité de la population générale française entre 1960 et 1964 (tables PM 60-64 et PF 60-64), toujours en vigueur dans la région lors de l'organisation des états généraux. Les procédures relatives à la construction de ces tables d'expérience ont été initiées en 2008, et les travaux effectifs ont commencé en 2009.

Le travail présenté ici s'inscrit en parallèle des travaux relatifs à la construction de ces tables de mortalité d'expérience réglementaires, et s'appuie donc sur les mêmes données, tout en préservant leur caractère confidentiel. En pratique, ce travail s'attache à présenter différents outils pour apprécier les risques liés à la durée de vie humaine de la population assurée dans la région.

Organisation de l'étude

La thèse est organisée en trois parties. La **première partie** précise le contexte dans lequel se place cette étude sur la mortalité d'expérience en Afrique subsaharienne francophone. Elle permet ainsi de présenter une analyse du marché de l'assurance vie, de justifier la nécessité de construire de nouvelles tables de mortalité pour la région et de décrire la méthode de construction de ces nouvelles tables. Cette première partie se conclut alors par une présentation des défis et des problématiques associés à la constitution de ces tables de mortalité de référence pour la région, notamment au regard de l'hétérogénéité potentielle au sein de la région (celle-ci étant composée de plusieurs pays distincts).

Dans ce contexte, la **deuxième partie** s'attache à prendre en compte l'hétérogénéité entre pays dans la mortalité d'expérience au sein de la région. Cet exercice conduit à segmenter la population en sous-populations, et donc à réduire la taille des échantillons dans les modèles. Avec ces échantillons réduits, il apparaît alors des fluctuations d'échantillonnages plus importantes et donc une plus grande incertitude dans les estimations des paramètres des modèles (risque systématique d'estimation). Une attention particulière est alors accordée au choix de modèles permettant de tenir compte de l'hétérogénéité, tout en limitant le niveau de l'incertitude liée aux fluctuations d'échantillonnage.

Les travaux de la première et de la deuxième parties sont réalisés au titre de la mortalité actuelle uniquement. Cependant, parmi les résolutions des états généraux il apparaît la volonté d'instaurer des régimes de retraite complémentaires gérés par les assureurs, et dans ce cadre la modélisation de la mortalité actuelle n'est pas suffisante. La modélisation de la mortalité future est donc nécessaire, et la **troisième partie** de l'étude s'inscrit dans ce cadre. À cet effet toutefois, deux scénarios distincts sont considérés.

Dans le premier scénario, on considère une situation dans laquelle aucune donnée sur les tendances passées de la mortalité n'est disponible pour la population assurée de la région. Dans ce cas, une modélisation de la mortalité

future à partir d'une extrapolation des tendances passées n'est pas possible. Aussi, en pratique dans ce scénario, qui correspond à la situation réelle aujourd'hui dans la région, la mortalité future est modélisée à partir d'avis d'experts sur l'évolution possible de la mortalité d'expérience dans la région. Une analyse du risque systématique lié au choix des avis d'expert y est également présentée.

Dans le second scénario, des données d'expérience sur les tendances passées de la mortalité sont supposées disponibles (si ce n'est pas le cas dans la région aujourd'hui, de telles données devraient devenir disponibles dans les prochaines années, compte tenu de la volonté de l'autorité de contrôle d'étendre l'analyse de la mortalité d'expérience de la région sur une dizaine d'années d'observation, cf. CIMA [2011]). La modélisation de la mortalité future est alors classiquement réalisée à partir des tendances passées et d'avis d'expert. L'analyse menée dans ce scénario consiste alors à fournir des outils pour mesurer les risques systématiques associés à cette modélisation (au titre des fluctuations d'échantillonnage, liées à la réduction de l'échantillon compte tenu de la segmentation des données par année pour identifier la tendance, et au titre du choix d'avis d'expert).

Au final, et de manière générale, ce travail a pour ambition de fournir des outils adaptés à la modélisation des risques liés la durée de vie humaine au sein de la région, et accorde une attention particulière à la mesure des risques systématiques associés à ces outils, compte tenu du caractère innovant de l'exercice et de l'incertitude entachant les données utilisées (pour lesquels le nombre de référence disponible est fortement limité).

*

* *

Les travaux menés dans le cadre de la présente thèse ont fait l'objet des publications suivantes² :

Kamega A. [2010] « Défis de l'assurance vie en Afrique noire francophone », *La Tribune de l'Assurance* 143.

Kamega A., Planchet F. [2011], « Hétérogénéité : mesure du risque d'estimation dans le cas d'une modélisation intégrant des facteurs observables », *Bulletin Français d'Actuariat*, Vol. 11, No. 21 (papier présenté le 21/06/2011 au Colloque AFIR 2011).

Planchet F., Kamega A. [2011A], « Construction d'une table de mortalité prospective pour un régime de rentes : prise en compte du risque d'estimation », *Cahiers de recherche de l'ISFA*, WP2140 (soumis pour publication).

² Cf. www.ressources-actuarielles.net

Planchet F., Kamega A. [2011B], « Mortalité prospective en cas de petits échantillons : modélisation à partir d'informations externes en utilisant l'approche de Bongaarts », *Cahiers de recherche de l'ISFA, WP2141* (soumis pour publication).

Bibliographie

- CIMA [2011] « Note sur la confection des nouvelles tables de mortalité », *CIMA*.
- Kone M.G.K. [2007a] « État des lieux de l'assurance vie dans les pays de la CIMA », *CIMA*.
- Kone M.G.K. [2007b] « État généraux de l'assurance vie - Rapport final », *CIMA/FANAF*.
- Yeatman J. [2007] « L'assurance en Afrique : une émergence difficile », *Risques, Les cahiers de l'assurance* 71.
- Ziguélé M. [2008] « Comment renforcer les compagnies d'assurances africaines de la zone CIMA », *CAPafrique – Analyses sectorielles*.

PARTIE I

ANALYSE DE LA MORTALITÉ DANS LA ZONE : PRÉSENTATION DU CONTEXTE ET DES DÉFIS TECHNIQUES

Cette première partie précise le contexte dans lequel se place cette étude sur la mortalité d'expérience en Afrique subsaharienne francophone. Elle permet ainsi de présenter une analyse du marché de l'assurance vie, de justifier la nécessité de construire de nouvelles tables de mortalité pour la région et de décrire la méthode de construction de ces nouvelles tables. Cette première partie se conclue alors par une présentation des défis techniques associés à la constitution de références de mortalité pour la région.

Chapitre 1

Présentation du marché de l'assurance vie en Afrique subsaharienne francophone

Sommaire

CHAPITRE 1 PRÉSENTATION DU MARCHÉ DE L'ASSURANCE VIE EN AFRIQUE SUBSAHARIENNE FRANCOPHONE	15
1. Organisation du marché : présentation des trois dispositifs d'assurance sur la vie (famille et communauté, assureurs privés, institutions publiques)	16
1.1.Famille et communauté (solidarité traditionnelle, dispositif informel)..	16
1.2.Assureurs privés du marché CIMA (dispositifs formels).....	19
1.3.Institutions publiques du marché CIPRES (dispositifs formels).....	24
2. Présentation du contexte socioéconomique pour les assureurs vie du marché CIMA	27
2.1.Présentation du contexte social	27
2.2.Présentation du contexte économique	32
2.3.Conclusion sur le contexte socioéconomique	36
3. Présentation des difficultés techniques et commerciales pour les assureurs vie du marché CIMA.....	37
3.1.Quelques problèmes historiques du marché.....	37
3.2.Conséquences des problèmes du marché	41
3.3.Exemples d'actions menées et perspectives.....	42
4. Synthèse et conclusion	45
Bibliographie.....	49

Ce premier chapitre présente le contexte et l'état du marché de l'assurance vie dans les pays d'Afrique subsaharienne francophone, assimilés à la zone CIMA/FANAF¹.

¹ Zone couvrant les pays membres de la CIMA (Conférence Interafricaine des Marchés de l'Assurance) et représentés par des sociétés d'assurance ou de réassurance de la FANAF (Fédération des sociétés d'Assurances de droits National Africaines), à savoir : le Bénin, le Burkina Faso, le Cameroun, la Centrafrique, le Congo Brazzaville, la Côte d'Ivoire, le Gabon, le Mali, le Niger, le Sénégal, le Tchad et le Togo.

De manière générale, l'assurance sur la vie dans la région s'articule autour de trois dispositifs : la famille et la communauté, les organismes d'assurance privés (marché CIMA) et les institutions publiques de l'État (marché CIPRES). L'origine et l'état des lieux de chacun de ces trois dispositifs sont présentés dans la première section.

Puis, à partir de la deuxième section, l'étude se concentre sur le dispositif relatif aux organismes d'assurance privés au sein du marché CIMA. Ainsi, la deuxième section présente le contexte socioéconomique du marché CIMA et la troisième section présente les problématiques techniques et commerciales du marché CIMA. Ces deux sections permettent entre autres d'inscrire les problématiques du marché CIMA dans leur contexte, et d'identifier les actions à mener pour y remédier.

1. Organisation du marché : présentation des trois dispositifs d'assurance sur la vie (famille et communauté, assureurs privés, institutions publiques)

Tous les individus sont soumis à des aléas, y compris sur la durée de vie humaine. Ces aléas peuvent avoir des conséquences financières susceptibles de menacer leur patrimoine. Ainsi par exemple, en cas de décès du chef de famille, la famille endeuillée peut ne plus disposer des ressources financières nécessaires pour maintenir son train de vie. À l'inverse, en cas de survie plus longue qu'anticipé d'un individu, la joie de cette longévité peut être nuancée par des problèmes financiers si celui-ci ne dispose pas d'une épargne suffisante et que plus personne n'est en mesure de l'aider.

Il convient alors de permettre aux individus de se protéger contre ces risques. À cet effet, on peut leur fournir des services d'assurance sur la vie, dont le principe s'appuie sur plusieurs méthodes de gestion des risques (incitation à la précaution, mutualisation des risques, partage des risques, transferts des risques, etc.).

Les individus ont classiquement accès à trois types de structure véhiculant l'assurance vie et la gestion des risques (*cf.* Denuit et Charpentier [2004]) : la famille (nucléaire, élargie ou extrapolée à des groupes sociaux), les organismes d'assurance privés et l'État (directement ou via des organismes publics spécialisés). Dans cette section, l'histoire et le rôle de chacune de ces trois solutions en Afrique subsaharienne francophone sont présentés, respectivement dans les sous-sections 1.1, 1.2 et 1.3.

1.1. Famille et communauté (solidarité traditionnelle, dispositif informel)

Selon Ewald [2010], la notion de risque désigne aujourd'hui le mode d'être collectif de l'homme en société. Pour l'auteur, l'assurance est ainsi désormais une obligation morale et sociale, elle n'est plus libre et ne se limite plus au fruit de cette vertu privé qu'est la prévoyance individuelle.

En Europe occidentale, cette obligation morale et sociale de l'assurance s'illustre principalement par la protection sociale assurée par l'État providence.

En Afrique subsaharienne francophone, cette obligation morale et sociale s'illustre davantage au niveau de la solidarité traditionnelle familiale et communautaire, compte tenu de son poids dans la structure sociale. Pour l'illustrer, cette sous-section présente quelques références historiques et un état des lieux du poids et du rôle de la solidarité traditionnelle en Afrique subsaharienne (*cf.* respectivement les paragraphes 1.1.1 et 1.1.2).

1.1.1. Quelques références historiques sur la solidarité traditionnelle

Pour introduire la solidarité ancestrale (de l'époque de l'empire du Ghana à la fin du XIX^e siècle), on distingue les dispositions sociales qui émanent de la sphère publique (et/ou professionnelle) et celles relatives à la sphère privée.

Solidarité ancestrale : structure sociale d'ordre public

Les structures sociales d'ordre public (et/ou professionnel) portent principalement sur les castes.

Aussi, Diop [1979] précise que les avantages et inconvénients liés à l'appartenance à une caste sont parfaitement équilibrés dans le système, puisque les individus de caste supérieure ne peuvent pas exploiter matériellement les individus de caste inférieure. Au contraire, les individus de caste supérieure doivent céder leurs richesses si une personne de caste inférieure s'adresse à eux, et en échange, ils disposent d'une supériorité morale (les individus de caste inférieure devant abdiquer toute leur personnalité devant eux).

Ce système de compensation et de solidarité constitue ainsi un mécanisme de prévoyance collective qui permet de limiter les frustrations et de contribuer à l'apaisement social.

Solidarité ancestrale : structure sociale d'ordre privé

Sur le plan privé, la société africaine est principalement caractérisée par une solidarité intergénérationnelle.

La segmentation de la vie d'un homme décrite par Hampaté Bâ [1972] illustre parfaitement cet aspect de la civilisation africaine. Ainsi, l'auteur rappelle que la tradition considère que la vie d'un homme comporte deux phases : une ascendante, jusqu'à 63 ans, et une descendante, jusqu'à 126 ans. La première est la phase active, elle est composée de trois parties : une première de 0 à 21 ans (la personne passe d'une dépendance totale à sa mère à une influence du milieu extérieur à travers l'école de la vie et des maîtres), une deuxième de 21 ans, âge du rite d'initiation, à 42 ans (la personne mûrit des enseignements reçus et demeure à l'écoute des sages) et une troisième de 42 à 63 ans (l'homme est censé avoir atteint la maturité et est tenu d'enseigner ce qu'il a appris). Lors de la seconde phase, qui commence à 63 ans, la personne est censée avoir achevé sa vie active et n'est plus astreinte à aucune obligation (il peut toutefois continuer à enseigner s'il le souhaite).

On retrouve dans ce développement l'expression de la solidarité générationnelle décrite par Diakitè [2008], à savoir le soin porté à l'éducation des enfants et au respect des anciens. En outre, nombre d'auteurs ont tenté de

développer ces valeurs. Ainsi, concernant l'éducation des enfants, Sow [2007] rappelle que tout adulte a le droit de rectifier ou de corriger tout enfant en faute, qu'il soit le sien ou non, dans l'optique d'assurer de la pérennité des solidarités. Il ajoute que c'est l'une des expressions les plus significatives de la solidarité sociale. Concernant le respect des anciens, Diop [1979] rappelle qu'en Afrique la sagesse, somme d'expériences vécues et de connaissances acquises, était fonction de l'âge (à l'inverse de l'Occident où dès le bas âge l'instruction pouvait conférer des connaissances supérieures à celles de beaucoup de personnes âgées).

Au-delà de cette solidarité intergénérationnelle, la générosité, qui au niveau du peuple prend la forme pratique de l'hospitalité, est aussi considérée comme une valeur morale essentielle (cf. Diop [1979]). Dans son ouvrage sur l'histoire de l'Afrique noire, Ki-Zerbo [1978] reprend ainsi de nombreux témoignages de voyageurs, occidentaux ou arabes, louant la grande hospitalité des populations d'Afrique subsaharienne. Cette hospitalité exclue toutefois tout parasitisme (comme le rappelle un dicton repris par Ki-Zerbo [1978] : « quand l'étranger arrive, nourris-le pendant deux jours ; le troisième jour, donne lui un outil ») et reste ainsi cohérente avec l'équilibre des systèmes de solidarité socioéconomiques présentés ci-dessus.

Diop [1979] appelle toutefois à une prudence dans l'interprétation de cette hospitalité et précise qu'elle n'est pas le reflet d'une « nature morale spécifique du Nègre ». Selon l'auteur, elle est principalement engendrée par des raisons économiques et matérielles ancestrales (à une époque où il n'existait ni banque ni hôtel), et est donc susceptible de s'estomper au fil des années.

1.1.2. La solidarité traditionnelle aujourd'hui

Aujourd'hui, la solidarité traditionnelle en Afrique subsaharienne francophone se manifeste principalement au sein de la sphère privée, et s'illustre notamment par le déclin de la solidarité spontanée (solidarité familiale intergénérationnelle) et par le développement de la solidarité organisée (solidarité communautaire intra-générationnelle, à travers des associations et tontines).

Le déclin de la solidarité « spontanée »

Une première expression de la solidarité africaine est celle que l'on retrouve au sein de la famille. À la différence de la famille nucléaire qui caractérise l'Occident, la famille en Afrique a un contenu étendu. Aussi, il n'est pas rare en Afrique qu'un homme s'occupe non seulement de son épouse et de ses enfants, mais également de ses parents, frères, cousins, neveux, beaux-frères, etc. (cf. Omvane Nkouélé [1988]). Cette solidarité spontanée aboutit à une assistance mutuelle intergénérationnelle et rassure les individus en ce qui concerne leur protection et leur sécurité. Cette solidarité spontanée de la famille ou du clan en Afrique se manifeste également par des aides matérielles ou financières apportées en cas de d'événement heureux ou malheureux (aide à la réhabilitation d'un commerce détruit, prestations en cas de décès, hospitalisation, mariage, naissance, etc.).

Cependant, l'individualisation et l'urbanisation de la société ces dernières années atténuent la solidarité spontanée africaine. La logique de la dette et du contrat intergénérationnel est de plus en plus mis à mal par les nouvelles générations. Malgré les pressions familiales et l'attachement aux traditions, les « enfants » aspirent à une vie organisée à leur façon et libérée de ces contraintes (cf. Ndembou [2007]).

Le développement de la solidarité « organisée »

Les changements sociaux évoqués ci-dessus conduisent au déclin des formes de solidarité spontanées traditionnelles. Toutefois, l'évolution de la société a maintenu, voire renforcé, les formes de solidarité organisées telles que les tontines (ou associations de solidarité).

Les tontines en Afrique peuvent se définir comme des regroupements de personnes, à adhésion individuelle et volontaire, en vue d'organiser un système d'entraide populaire d'épargne ou de crédit. Ces regroupements sont d'ordre familial, régional, socioprofessionnel, amical, linguistique, etc. (cf. Bleu [1984]) et présentent les avantages de fournir des services financiers de manière extrêmement souple, tout en maintenant un respect des traditions et coutumes (à travers leur mode de fonctionnement, leurs fêtes annuelles généralement organisées, leur mode de célébration d'événements heureux ou malheureux, etc.).

En Afrique, on distingue généralement les tontines rotatives (dans ce cas les participants s'engagent à verser une somme prédéterminée à une fréquence donnée, et pour chaque tour de versement, un des participants est désigné pour être le bénéficiaire des fonds des autres participants), et les tontines accumulatives (dans ce cas les cotisations ne sont pas redistribuées à un des membres mais accumulées dans une caisse, avant d'être redistribuées aux membres, lorsqu'ils le décident, au prorata de ce qu'ils ont cotisé). Dans bien des cas toutefois, les deux modes de tontines sont pratiqués simultanément.

Bien souvent, les tontines accumulatives offrent également des services d'assurance à leurs membres. Ces services d'assurance sont organisés selon les principes classiques de l'assurance : mutualisation des risques entre tous les membres qui acquittent une cotisation préalable et droits à prestations en cas de réalisation des risques assurés (cf. Yeatman [2007]). Ils couvrent généralement les risques sur le décès (d'un membre, de son conjoint ou de la famille proche) et sur l'hospitalisation, et sont également utilisés en cas d'événement heureux tel que le mariage d'un membre ou la naissance d'un enfant.

1.2. Assureurs privés du marché CIMA (dispositifs formels)

Cette sous-section présente la naissance et l'organisation du marché de l'assurance vie classique, c'est-à-dire celle des assureurs privés, au sein de la zone CIMA/FANAF.

1.2.1. Naissance de l'assurance classique en Afrique subsaharienne francophone

Ce paragraphe présente différentes étapes de l'introduction de l'assurance vie classique au sein de la zone CIMA.

Premières traces de l'assurance classique en Afrique

Sous sa forme « moderne » et en se limitant aux anciennes colonies des pays d'Afrique subsaharienne, l'assurance est associée à l'arrivée des Anglais au XIXe siècle, lorsque les Lloyd's de Londres et d'autres grandes compagnies anglaises envoyaient des courtiers et des experts maritimes dans les ports pour accompagner le commerce international. Puis, peu à peu ces assureurs anglais ont étendu leur offre de garanties pour couvrir les risques des entreprises étrangères installées sur place, ainsi que ceux de leurs dirigeants expatriés.

Quelques années après, face à l'augmentation de la population européenne et à la croissance du nombre d'entreprises étrangères sur place, parallèles au développement économique et démographique de la région, les principales sociétés françaises se sont à leur tour intéressées à ces marchés. Leurs premiers agents généraux sont nommés en Afrique subsaharienne francophone quelques années avant la seconde Guerre mondiale (*cf.* Yeatman [2004]).

Développement du marché en zone franc CFA par les acteurs français

Avant l'indépendance, le développement de l'assurance en zone franc CFA² était principalement rythmé par trois acteurs français, comme le rappelle Yeatman [2004] dans son étude sur l'assurance française dans cette zone.

Soucieux de ne pas laisser les marchés des territoires d'outre-mer aux mains d'assureurs étrangers, l'État français, premier acteur français de ce développement, nationalise après la Libération en 1946 les trente-quatre sociétés d'assurances les plus importantes. C'est le début d'une extension rapide du réseau d'agents généraux issus de sociétés françaises. Le deuxième acteur français de l'essor de l'assurance en zone franc CFA est le GFA (Groupement Français d'Assurances). Constitué en 1947, le GFA est né du désir de dix sociétés françaises, de taille moyenne et restées privées après les nationalisations, de mettre en commun leurs moyens pour prospecter les territoires d'outre-mer. Sur le terrain, les agents du GFA se développent et prennent une part significative du marché « en recherchant la clientèle de la population africaine souvent négligée par les sociétés plus traditionnelles ». Le troisième acteur français est le groupe de mutuelles professionnelles Monceau et leur réassurance Mutuelle Centrale de Réassurance (MCR). Sous leur impulsion, une série de sociétés mutuelles ont vu le jour en zone franc CFA et ont pu bénéficier de leur assistance technique et de la réassurance de la MCR.

Pour accompagner ces trois sources de développement, des organisations professionnelles françaises ont pris des mesures en vue d'améliorer le suivi et la gestion des risques locaux : tenue de statistiques communes, inspection des

² Pour mémoire, à cette époque, ce sigle représentait les « Colonies Françaises d'Afrique ».

risques, etc. Ces initiatives sont bénéfiques en termes de structure et de développement des marchés de l'assurance africains. Elles éloignent toutefois un peu plus les Africains de la prise en main de leur propre marché ; et les indépendances n'y changeront pas grand chose sur le fond.

En effet, si, après les indépendances, chaque État se dote d'une législation propre sur les assurances, il n'en demeure pas moins qu'elle est quasiment copiée sur la loi française et que la direction des assurances du ministère français des Finances y joue un rôle prépondérant. Malgré quelques mesures phares qui se sont révélées assez dissuasives pour les assureurs étrangers, telles que l'obligation d'agrément local pour assurer les risques situés sur place et les contraintes de représentation locale des actifs représentatifs des engagements techniques, au lendemain des indépendances la présence française demeure très marquée dans l'organisation des marchés de l'assurance en zone franc CFA.

1.2.2. Mutation des structures et organisation régionale : de la CICA à la CIMA

Ce paragraphe présente les différentes mutations des structures chargées d'organiser le marché de l'assurance vie classique dans la région.

Émergence et évolution de la CICA

C'est dans le contexte d'une présence française très marquée qu'est signée la première convention de la Conférence Internationale de Contrôle des Assurances (CICA), le 27 juillet 1962 à Paris. Composée des anciennes colonies françaises d'Afrique de l'Ouest et centrale, de Madagascar et de la France, cette organisation a pour objectif d'harmoniser les législations et réglementations nationales, de coordonner l'exercice du contrôle des entreprises et de coordonner la formation en assurance des cadres africains (*cf.* Nonyu Moutassie [2005]). Le siège de la CICA est à Paris.

Au fil de la formation des cadres africains, les États prennent conscience de l'importance du secteur pour le développement de leur pays. Cette perception s'accroît par ailleurs à l'issue de l'Assemblée Générale de la Conférence des Nations Unies sur le Commerce Et le Développement (CNUCED) en 1972. Parmi les recommandations, la Conférence sommat ainsi les pays en développement de créer des sociétés à capitaux totalement ou majoritairement nationaux et dirigées par des cadres nationaux (*cf.* Nonyu Moutassie [2005]). Si, en réponse, certains pays membres se sont orientés vers la nationalisation des opérateurs du secteur, d'autres ont favorisé la création de sociétés de droit local entièrement privées et opérant dans un marché ouvert et concurrentiel (*cf.* Ziguélé [2008]). Dans les deux cas, l'objectif est toutefois le même : rendre africaines les sociétés opérant sur les marchés africains. Pour parvenir à cet objectif, une CICA plus africaine s'avérait toutefois nécessaire.

Ainsi, le 27 novembre 1973 une deuxième convention de la CICA est signée à Paris par les seuls États africains (à l'exception de la Mauritanie et de Madagascar) et en 1976 le siège de l'organisation est transféré de Paris à Libreville (République Gabonaise). Non signataire de cette convention, la France ne bénéficie plus que du statut d'observateur (*cf.* Nonyu Moutassie [2005]).

À l'aube des années 90, c'est l'heure du bilan de trente années d'existence de la CICA. De précieuses réalisations ont vu le jour sous l'égide de la Conférence, telles que des sociétés nationales et des services nationaux de contrôle des assurances, l'Institut International des Assurances (IIA), la Fédération des sociétés d'Assurances de droit National Africaines (FANAF) ou encore la compagnie commune de réassurance des États membres de la CICA (CICA-RE). Toutefois, à l'issue de ces trente années, le développement des marchés de l'assurance demeure limité en raison, entre autres, d'une mauvaise gestion de la plupart des sociétés nationales, de l'inefficacité des services de contrôle au niveau national, de l'impuissance de la CICA comme organe de contrôle et d'harmonisation au niveau régional, du nombre pléthorique d'intermédiaires, de l'absence de marché financier, de la faiblesse de l'assurance vie, de marchés nationaux trop exigus pour permettre un développement satisfaisant de l'industrie des assurances, ou encore de tarifs inadaptés (*cf.* Nonyu Moutassie [2005]). De surcroît, la CICA ne dispose d'aucun pouvoir de décision et les pouvoirs publics sont indifférents vis-à-vis des études et contrôles effectués par celle-ci : la CICA ne semble pas assez intégrée dans la zone. Aussi, le 20 septembre 1990, une troisième convention est signée à Paris par les mêmes États africains. Il s'agit de la Convention de coordination pour la Promotion et le Développement de l'Industrie des Assurances (CPDIA), caractérisée par la création d'un Conseil des Ministres des assurances des États membres et d'une commission interétatique de contrôle des assurances. Néanmoins, avant la ratification par la majorité des États signataires, apparaît l'idée d'aller plus loin que la CPDIA en créant une organisation intégrée³.

Émergence et fonctionnement de la CIMA

L'idée de la création d'une organisation intégrée conduisit, le 10 juillet 1992, à la signature du Traité instituant une organisation intégrée de l'industrie des assurances dans les États africains, dénommée Conférence Interafricaine des Marchés d'Assurances (CIMA). Le Traité instituant la CIMA est convenu par les gouvernements du Bénin, Burkina Faso, Cameroun, Centrafrique, Congo Brazzaville, Côte d'Ivoire, Gabon, Guinée Équatoriale, Mali, Niger, Sénégal, Tchad et Togo (*cf.* CIMA [1992]) et est entrée en vigueur en 1995 (la Guinée Bissau y adhère en 2002, *cf.* Nonyu Moutassie [2005]).

Le traité instituant la CIMA fait, entre autres, référence aux objectifs de la Conférence, à la nouvelle législation du marché et aux organes et pouvoirs mis en place (*cf.* CIMA [1992]).

Les objectifs assignés à la CIMA sont notamment :

- d'améliorer la couverture des risques en les adaptant aux réalités africaines ;
- d'encourager la rétention des affaires au plan national et régional ;

³ Les Ministres chargés des assurances étaient alors particulièrement soucieux d'éviter une crise comparable à celle rencontrée dans le système bancaire quelques années plus tôt (*cf.* Diallo [2008]).

- de favoriser l'investissement local des provisions générées ;
- de créer des structures communes afin de favoriser le développement des entreprises d'assurances, de favoriser la constitution d'un marché élargi et intégré sur l'ensemble des pays membres et de mettre en place de nouveaux instruments financiers ;
- de poursuivre la politique d'harmonisation et d'unification des dispositions législatives et réglementaires.

En vue de répondre à ces objectifs, la Conférence s'est dotée d'une législation unique des assurances : il s'agit du Code des Assurances des États membres de la CIMA.

Enfin, toujours en vue d'assurer la pleine réalisation des objectifs du Traité, trois organes sont réorganisés ou créés autour de la CIMA :

- le Conseil des Ministres (organe directeur de la CIMA qui assure la réalisation des objectifs du Traité) ;
- la Commission Régionale de Contrôle des Assurances (CRCA, organe régulateur de la CIMA qui assure le contrôle des sociétés, assure la surveillance générale et concourt à l'organisation des marchés nationaux) ;
- le Secrétariat Général (organe qui assure la préparation, l'exécution et le suivi des travaux du Conseil des Ministres et de la CRCA).

Depuis 1995, les nombreuses mesures prises par la CRCA (avertissements, blâmes, mises sous surveillance permanente, suspensions de dirigeants, retraits d'agrèments, etc.) et le Conseil des Ministres (modifications régulières du Code CIMA, dont notamment les relèvements du capital social minimal) ont contribué au développement du chiffre d'affaires des organismes assureurs et à un assainissement du marché (*cf.* Ziguélé [2008]).

1.2.3. L'assurance vie classique dans le marché CIMA aujourd'hui

Malgré les évolutions positives depuis 1995, au regard des mesures prises par la CRCA et le Conseil des Ministres, de nombreuses difficultés subsistent sur le marché CIMA de l'assurance en général. De surcroît, le marché est marqué par la faiblesse de l'assurance vie classique, alors même que ce secteur de l'assurance participe au développement et à la modernisation des économies des pays (*cf.* notamment la section 3 pour une présentation des difficultés de l'assurance vie dans la région).

Ainsi par exemple, l'assurance vie (décès et épargne-retraite) représente moins d'un tiers du chiffre d'affaires de l'assurance en zone CIMA/FANAF (*cf.* FANAF [2011]), alors qu'à titre de comparaison elle représente plus de deux tiers du chiffre d'affaires de l'assurance en France (*cf.* FFSA [2010]).

En outre, selon Kone [2007], il apparaît que le taux de pénétration officiel de l'assurance vie (rapport entre les cotisations officielles en assurance vie et le PIB officiel) est égal à 0,2 % au niveau de la zone CIMA/FANAF, contre plus de 7 % en France, plus de 4 % au niveau mondial et plus de 3 % au niveau de

toute l'Afrique (dont 10,9 % en Afrique du Sud⁴ – un des taux les plus importants au monde –, 5,8 % en Namibie, 2,6 % aux Îles Maurice, 0,8 % au Kenya, 0,7 % au Maroc, etc.).

Pour mémoire toutefois, ces chiffres officiels se doivent d'être nuancés compte tenu des nombreux systèmes d'assurance vie endogènes (réseaux de solidarité familiaux, tontines, etc., *cf.* sous-section 1.1) existant aujourd'hui au sein de la région et non pris en compte dans les chiffres officiels. Ces seuls indicateurs officiels ne sont pas de bons baromètres du développement du secteur de l'assurance vie (classique et informel) dans la zone.

1.3. Institutions publiques du marché CIPRES (dispositifs formels)

Après la famille et la communauté (solidarité traditionnelle) et les organismes d'assurance privés, on présente ici un troisième véhicule de gestion des risques au sein de l'Afrique subsaharienne francophone : l'État, ou plus exactement les institutions publiques de l'État chargées de la solidarité professionnelle (prévoyance sociale). Il s'agit du deuxième niveau de protection sociale dans la région (le premier niveau, relatif aux filets sociaux de sécurité et visant les populations les plus pauvres, étant négligeable dans la région).

Cette sous-section est segmentée en trois parties : une première qui permet de présenter l'origine et le périmètre des garanties couvertes par la prévoyance sociale dans la région (*cf.* paragraphe 1.3.1), une seconde décrivant son fonctionnement avant les années 1990 (*cf.* paragraphe 1.3.2) et une troisième présentant l'émergence et le rôle de la CIPRES depuis le début des années 1990 (*cf.* paragraphe 1.3.3).

1.3.1. Origines et risques couverts par la prévoyance sociale

Comme le rappelle la CIPRES⁵, la prévoyance sociale en Afrique trouve ses sources dans la Déclaration universelle des droits de l'homme du 10 décembre 1948, qui précise dans son article 22 que « toute personne (...) a droit à la sécurité sociale (...) », et dans la Convention n°102 de l'Organisation Internationale du Travail (OIT) du 28 juin 1952, qui fixe la norme minimale de la sécurité sociale.

En particulier, la Convention n°102 de l'OIT définit les prestations qui forment le noyau de la norme minimale, à savoir : les soins médicaux, les indemnités de maladie, les prestations de chômage, les prestations de vieillesse, les prestations en cas d'accident du travail et de maladie professionnelle, les prestations aux familles, les prestations de maternité, les prestations d'invalidité ou les prestations de survivants.

⁴ Ce niveau exceptionnel en Afrique du Sud tient toutefois de l'organisation des prestations sociales dans ce pays, dans laquelle les salariés sont couverts pour leur risque vieillesse par les assureurs privés et non par des organismes publics ne faisant pas partie des marchés d'assurance (*cf.* Yeatman [2007]).

⁵ *Cf.* <http://www.lacipres.org/prevoyance.html>.

Ces diverses prestations sont généralement organisées en régimes et en branches. Parmi les régimes, on distingue généralement le régime des fonctionnaires et celui des salariés relevant du Code du travail. Parmi les branches, on identifie généralement trois catégories (*cf.* CIPRES [2007]) : celle des prestations familiales (y compris prestations de maternité), celle des accidents du travail et des maladies professionnelles et celle de la vieillesse, de l'invalidité et du décès (prestations de survivants).

1.3.2. Gestion technique et administrative de la prévoyance sociale avant la CIPRES

Sur le plan technique, le système de financement des prestations est contributif. Aussi, chaque branche est financée par des cotisations à la charge des salariés et/ou des employeurs du secteur formel. En pratique toutefois, la structure de la population active (près de 70 % de la population active est dans le secteur primaire et près de 75 % des emplois non agricoles sont des emplois informels selon le PNUD et l'OIT, *cf.* notamment Hugon [2009]), les problèmes économiques et les politiques d'ajustements structurels menées à partir des années 1980 (gelant les recrutements sur des décennies) limitent la taille de la population cotisants aux différentes branches (sur ce point, *cf.* par exemple l'article de Juompan-Yakam [2010] sur le cas du financement des prestations vieillesse dans la région). Au final, malgré une pyramide des âges favorable dans la région (*cf.* paragraphe 2.1.3), il ressort une faible assiette des cotisations qui, associée à un niveau de recouvrement des cotisations perfectible, crée des tensions sur l'équilibre technique des régimes.

Sur les plans administratif et financier, il apparaît également des dysfonctionnements importants, notamment en raison d'ambiguïtés dans la définition des responsabilités (due à l'immixtion de l'État), du manque de formalisation de certaines opérations, du déficit général d'information des assurés (notamment sur les conditions d'accès aux droits), de délais de paiement et de liquidation des droits insatisfaisants, d'un niveau excessif des frais de gestion, etc. (*cf.* CIPRES [2007]).

C'est sur la base de ce constat, que les Ministres des Finances de la zone franc ont décidé, lors d'une réunion d'avril 1991 à Ouagadougou (Burkina Faso), de mettre en commun leurs moyens en créant une Organisation supranationale chargée de l'assainissement, de la surveillance et de la régularisation du secteur de la prévoyance sociale dans la zone.

1.3.3. Émergence et rôle de la CIPRES

La Conférence Interafricaine de la Prévoyance Sociale (CIPRES) a été instituée par un Traité signé en septembre 1993 à Abidjan, par quatorze pays d'Afrique subsaharienne (les douze pays de l'actuelle zone CIMA/FANAF, augmentés des Comores et de la Guinée Équatoriale)⁶.

⁶ La République Démocratique du Congo a été admise au sein des États membres de la CIPRES le 27 février 2009.

La CIPRES est un outil de contrôle et de régulation du secteur de la prévoyance sociale en vue d'une meilleure couverture des populations desdits États. Ses objectifs sont notamment (*cf.* CIPRES [1993]) :

- d'instituer un contrôle régional de la gestion des organismes de prévoyance sociale en vue de la rationalisation de leur fonctionnement ;
- de fixer des règles communes de gestion ;
- de réaliser des études et d'élaborer des propositions tendant à l'harmonisation des dispositions législatives et réglementaires applicables aux organismes et aux régimes de prévoyance sociale ;
- de faciliter la mise en œuvre, par des actions spécifiques au niveau régional, d'une politique de formation initiale et permanente des cadres et techniciens des organismes de prévoyance sociale dans les États membres.

Par ailleurs, en vue d'assurer la pleine réalisation des objectifs du Traité, trois organes sont réorganisés ou créés autour de la CIPRES (à l'image de ce que l'on observe pour la CIMA, *cf.* paragraphe 1.2.2) :

- le Conseil des Ministres (organe directeur de la CIPRES, garant de la réalisation des objectifs du Traité),
- la Commission régionale de surveillance de la prévoyance sociale (organe chargée de veiller à la bonne gestion des organismes de prévoyance sociale dans les États membres et de participer à la régulation du secteur de la prévoyance sociale ainsi qu'à la réalisation des objectifs du présent Traité),
- l'Inspection régionale de la prévoyance sociale (organe qui assure la préparation, l'exécution et le suivi des travaux du Conseil des Ministres et de la Commission régionale).

Depuis le démarrage des activités techniques de la CIPRES, de nombreuses réalisations ont été effectuées dans le cadre des objectifs assignés : adoption d'un plan comptable définissant la nomenclature unique, élaboration de fiches techniques, missions d'inspections, organisations de sessions de formation, etc.

Aussi, selon la CIPRES [2007], certains résultats encourageants ont été enregistrés : augmentation des ressources de cotisations, amélioration du service rendu aux bénéficiaires dans le règlement des prestations, amélioration dans les procédures de traitement des opérations, meilleure maîtrise des charges administratives, etc. De nombreux défis se posent toutefois aux autorités concernant la prévoyance sociale dans la zone : extension aux populations actuellement exclues (secteur primaire, emploi informel), développement d'un filet social de sécurité pour tous (notamment les plus vulnérables), développement de la branche maladie, etc.

2. Présentation du contexte socioéconomique pour les assureurs vie du marché CIMA

La section précédente présente les trois structures d'assurance sur la vie existants, que l'on peut ventiler en deux catégories : une concernant les seuls fonctionnaires et salariés du secteur formel (les institutions publics), et une autre concernant toute la population, mais basées sur une initiative individuelle ou collective (la famille et la communauté d'une part et les organismes d'assurance privés d'autre part). Dans les deux catégories, il s'agit de régimes contributifs (à l'exception de la solidarité traditionnelle familiale).

À partir de cette section, on se concentre principalement sur l'assurance vie proposé par les organismes privés. Bien que ce dispositif soit théoriquement ouvert à toute la population de la région, en pratique, seule une petite partie d'entre elle est concernée. L'un des principaux défis auquel doivent alors faire face les organismes privés est d'élargir le périmètre de la population couverte, et à cet effet ils sont tenus d'intégrer les caractéristiques du contexte socioéconomique local dans leurs démarches techniques et commerciales. C'est dans ce contexte que cette section s'attache à présenter l'environnement social (*cf.* sous-section 2.1) et économique (*cf.* sous-section 2.2) de la région.

2.1. Présentation du contexte social

Cette sous-section présente les spécificités sociales des pays d'Afrique subsaharienne francophone (assimilée à la zone CIMA/FANAF), tant sur le plan culturel, religieux et psychologique que sur le plan démographique.

2.1.1. Croyances et religion

Il semble acquis que l'Afrique dispose d'une culture riche, de traditions très marquées et d'un investissement religieux très prononcé. Si ces propriétés suscitent fascination et admiration chez certains, ils sont également facteurs de difficultés et d'obstacles chez d'autres, notamment chez les assureurs qui proposent leurs produits d'assurance vie.

Des croyances a priori défavorables aux assureurs vie...

Sur un plan socioculturel, le contexte de la providence mentionné par Bleu [1984] dans les années 80 semble demeurer aujourd'hui, et c'est ainsi que dans certaines traditions africaines ou religieuses, la conception de la vie favorise l'attentisme et considère que l'Homme n'est pas en mesure de maîtriser les effets d'événements futurs susceptibles de toucher sa sécurité (*cf.* également Kone [2007]). De même, certaines traditions et religions dénoncent le caractère immoral de la réalisation de profits suite à l'avènement d'un malheur (*cf.* Haman Bello [1984] et Kone [2007]).

Concernant les barrières culturelles et religieuses, une attention particulière peut par ailleurs être accordée à l'islam. Pour mémoire, la charia, ensemble de règles de conduites applicables aux musulmans, classe les actions humaines en cinq catégories : *fard* (actions obligatoires), *mandub* (actions recommandées), *mubâh* (actions indifférentes), *makrûh* (actions désapprouvées), *haram* (actions interdites). Parmi les actions *haram*, l'islam interdit dans les transactions commerciales la *riba*, qui représente l'intérêt et est généralement traduit par

l'usure, le *gharar*, c'est-à-dire l'incertitude ou l'ambiguïté et par extension le transfert de risque, et le *maisir*, qui correspond à la prise excessive de risque (cf. Kassim [2008]). La moitié de la population de la zone CIMA/FANAF étant musulmane (cf. Tableau 1 ci-après), il ne fait aucun doute que ces restrictions et considérations y limitent le développement de l'assurance vie classique (bien que l'assistance mutuelle soit recommandée entre les musulmans, cf. Sylla Ndoye [1981]).

...mais qui restent compatibles avec le développement de l'assurance vie

Quelle place donc pour l'assurance vie dans ce contexte ? En pratique, il apparaît que malgré les éléments présentés ci-dessus, l'assurance vie et les croyances des populations en zone CIMA/FANAF ne sont pas incompatibles.

Concernant la providence et le respect des traditions, cette compatibilité est matérialisée par les nombreuses structures d'entraide et d'assistance traditionnelles que l'on retrouve aujourd'hui, même dans les villages et les régions les plus reculés. Il s'agit ici de la fameuse solidarité traditionnelle, présentée dans la section précédente (cf. notamment la sous-section 1.1), dont le fonctionnement est comparable à celui de l'assurance vie classique.

Concernant la religion musulmane en particulier, cette compatibilité s'illustre par exemple par le fort développement de l'assurance *takaful* dans les pays du Golfe et en Malaisie ces dernières années. L'assurance *takaful* est un mécanisme d'assurance basé sur l'entraide dans la voie tracée par le Coran et la charia (cf. Daniel [2007]). Ainsi, au sein d'une assurance vie *takaful* il n'y a pas de primes mais des dons (*tabarru*), il y a mutualisation des sinistres et des donations, le bénéficiaire n'est qu'un récipiendaire de l'indemnité et devra à ce titre la répartir entre les héritiers selon l'ordre déterminé par la loi coranique et, entre autres, les placements admis sont d'une part ceux dont la rémunération résulte d'un partage du sort entre l'investisseur et le bénéficiaire de l'investissement et d'autre part ceux étant en harmonie avec les préceptes religieux (cf. Daniel [2007] et Kassim [2008]).

2.1.2. Comportements sociaux et éducation

Au-delà des aspects culturels et religieux, l'Afrique subsaharienne francophone est également marquée par les comportements sociaux et le niveau d'éducation de ses populations.

Une psychologie a priori défavorable aux assureurs vie...

Au-delà de ses croyances traditionnelles et religieuses, la population est caractérisée par une mentalité et une psychologie qui semblent indisposer à la souscription de contrats d'assurance vie. Trois exemples illustrent ces propos.

Le premier est qu'au sein de certaines populations d'Afrique subsaharienne, la mort est un sujet tabou (cf. par exemple Aku Sama [1996] et Betoé Minko [1986]), ou du moins une réalité traumatisante dans on n'aime pas parler (cf. Direction des Assurances (Bénin) [2007]). En outre, quand bien même ce premier obstacle serait levé, certains assurés potentiels seraient

touchés par la crainte d'être tués par les bénéficiaires du contrat (cf. Aku Sama [1996]).

Par ailleurs, l'absence de concret dans les contrats d'assurance vie ne satisfait pas certaines populations. En effet, nombre d'assurés potentiels s'accommodent mal d'une simple promesse, qui plus est parfois de long terme, qu'est la sécurité en assurance vie : toute sortie d'argent doit se matérialiser par un bien palpable. « En général, l'Africain ne compte qu'avec le concret en matière d'échanges » (cf. Debayou [1988]), et l'assurance vie ne semble pas être perçue comme tel.

Le troisième handicap identifié ici est l'incompréhension des produits par les prospects. En effet, nombre des termes utilisés en assurance vie semblent ésotériques pour l'essentiel des assurés potentiels (cf. Bleu [1984]). En relevant qu'au sein de la zone CIMA/FANAF près de 50 % de la population des plus de 15 ans est estimée analphabète (cf. Tableau 1 ci-après), on prend la mesure de la barrière terminologique qui peut exister dans le secteur.

...mais qui reste compatible avec le développement de l'assurance vie

La mentalité des populations et leur incompréhension des garanties mettent en avant la nécessité d'une pédagogie et d'une explication des garanties et des produits. Ces caractéristiques n'illustrent pas une incompatibilité avec le secteur de l'assurance vie, comme en témoignent l'importance des solidarités traditionnelles (cf. sous-section 1.1), elles mettent seulement en exergue une culture de l'assurance vie privée et professionnalisée encore trop faible au sein des pays de la région.

2.1.3. Structure de la population générale

Les paragraphes précédents présente la population de la zone CIMA/FANAF sous différents angles (culturels, religieux, comportementaux, éducation, etc.), en insistant sur les freins qu'ils peuvent occasionner pour le développement du marché de l'assurance vie. Mais comment est structurée cette population générale ? Qu'en est-t-il de ses caractéristiques démographiques aujourd'hui ? Qu'en sera-t-il dans les prochaines décennies ?

Présentation de la structure de la population

Les ménages, et donc les populations, sont au cœur de toute activité économique, surtout en Afrique où « l'investissement démographique » se substitue communément à l'investissement économique, l'enfant représentant effectivement une utilité économique en tant que force de travail, assurance vieillesse ou encore source de transferts futurs (cf. Hugon [2009]).

Dans ce contexte, le tableau suivant présente des indicateurs démographiques relatifs aux 12 pays de la zone (chiffres issus de Bafd/OCDE [2008], sauf précision contraire).

Tableau 1 - Indicateurs démographiques de la zone CIMA/FANAF

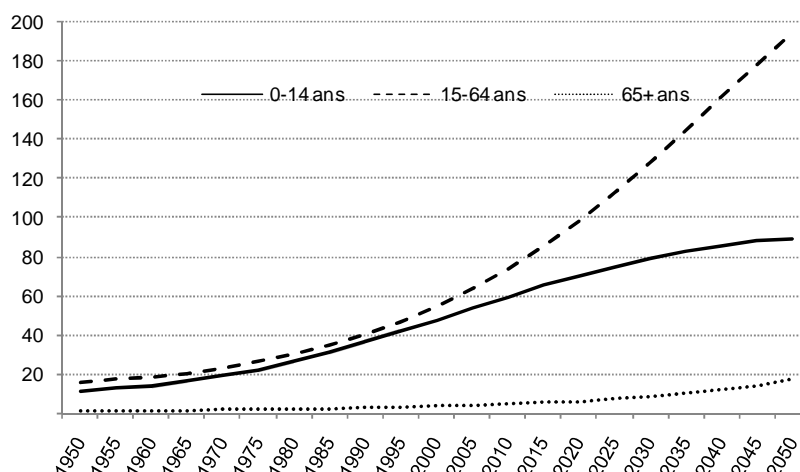
Pays / Zone	Pop. totale (en millions, 2007)	Distribution par âge (% du total, 2007)			Pop. urbaine (% du total) (2007)	Taux d'analphabétisme des plus de 15 ans (2007)		VIH/Sida Prévalence chez l'adulte (2005)	Part des musulmans* (%) (2004)
		0-14 ans	15-64 ans	65+ ans		Hommes	Femmes		
Bénin	9,0	44%	54%	3%	47%	39%	70%	1,8%	15%
Burkina Faso	14,8	46%	51%	3%	18%	59%	80%	2,0%	50%
Côte d'Ivoire	19,3	41%	56%	3%	46%	35%	55%	7,1%	38%
Mali	12,3	48%	49%	4%	41%	58%	79%	1,7%	94%
Niger	14,2	48%	49%	3%	26%	71%	88%	1,1%	95%
Sénégal	12,4	42%	54%	4%	52%	47%	65%	0,9%	94%
Togo	6,6	43%	54%	3%	37%	21%	47%	3,2%	20%
UEMOA	88,6	44%	52%	3%	38%	49%	70%	2,9%	61%
Cameroon	18,5	41%	55%	4%	50%	15%	27%	5,4%	20%
Centrafrique	4,3	42%	54%	4%	43%	32%	54%	10,8%	15%
Congo	3,8	42%	55%	3%	62%	8%	17%	5,3%	0%
Gabon	1,3	35%	61%	5%	93%	10%	18%	7,9%	1%
Tchad	10,8	46%	51%	3%	25%	39%	54%	3,5%	44%
CEMAC	38,8	43%	54%	3%	45%	23%	36%	5,6%	24%
CIMA/FANAF	127,4	44%	53%	3%	40%	41%	60%	3,7%	50%

* Source : Frémy et Frémy [2005]

La zone CIMA/FANAF compte ainsi plus de 125 millions d'individus, dont 44 % ont moins de 15 ans et 53 % ont entre 15 et 64 ans.

Le graphique ci-après présente la « Révision 2006 » des projections démographiques mondiales publiées par la Division de la population des Nations Unies⁷.

⁷ Hypothèses de projections : poursuite des tendances migratoires récentes et prise en compte des politiques migratoires annoncées ; convergence à long terme des taux de fécondité vers 1,85 enfant par femme ; allongement et convergence de l'espérance de vie à long terme ; réduction de moitié en 30 ans du nombre de nouveaux individus atteints par le Sida et de l'intensité de l'infection ; prise en compte d'aucune pandémie majeure. Cf. <http://esa.un.org/unpp/index.asp?panel=4>.

Graphique 1 - Population de la zone CIMA/FANAF (en millions)

Selon ces projections, la population de la zone devrait doubler dans une génération.

Par ailleurs, au-delà de l'explosion démographique générale, ces projections illustrent l'explosion de la population africaine en âge de travailler. Cette population jeune permet notamment de maintenir un taux de dépendance vieillesse inférieur à 10 % jusqu'en 2050, ce qui se traduit par exemple par un besoin en termes de prise en charge sociale des retraités limité (à titre de comparaison, le taux de dépendance vieillesse est évalué à 25 % aujourd'hui en France).

Une structure de population générale favorable au développement de l'assurance vie

Avec la forte croissance démographique et la forte proportion de jeunes, la zone CIMA/FANAF dispose d'un contexte favorable pour le développement de produits d'assurance vie en général (une forte population permet une forte mutualisation des risques) et de produits d'épargne/retraite en particulier (le taux de dépendance vieillesse demeurant inférieur à 10 % jusqu'en 2050).

En outre, ces constats sur la croissance démographique et la proportion de jeunes permettent de disposer d'une grande flexibilité du facteur travail, notamment en termes de requalifications professionnelles ou de mobilité géographique (*cf.* Ouédraogo [1999]), ce qui peut conduire à un dynamisme économique, puis à un dynamisme du marché de l'assurance vie. Toutefois, la croissance de l'activité économique ne sera réalisable qu'à condition d'organiser la participation de ces jeunes au processus de production. Tout dépendra ainsi de la capacité du système économique, puis du marché de l'assurance vie, à s'adapter.

2.2. Présentation du contexte économique

Au-delà des aspects sociaux, l'Afrique subsaharienne francophone est également marquée par un contexte économique spécifique. À ce titre, cette sous-section présente la situation macroéconomique du marché, le poids et le rôle du secteur informel ainsi que le comportement microéconomique des agents de la région.

2.2.1. Macroéconomie et analyse économique

Ce paragraphe présente certaines caractéristiques des économies nationales et sous-régionales de la zone CIMA/FANAF. Après une présentation de la situation monétaire de la zone, il s'agit ici de fournir quelques éléments d'analyse de la situation économique dans la région.

Monnaie et intégration régionale

Les monnaies utilisées dans les douze pays de la zone CIMA/FANAF, les francs CFA, sont liées à l'euro (avant 1999 elles étaient liées au franc français) par un système de parité fixe avec une garantie de convertibilité illimitée du Trésor français (*cf.* Agbohou [2008]). Cette parité est identique pour tous les pays de la zone CIMA/FANAF : 1 euro = 655,957 FCFA. Cependant, la coopération monétaire entre la France et ces douze pays est organisée autour de deux unions monétaires distinctes ayant chacune une banque centrale : l'UEMOA (la banque centrale associée est la BCEAO) pour les pays d'Afrique de l'Ouest et la CEMAC (la banque centrale associée est la BEAC) pour les pays d'Afrique centrale. En outre, il n'y a pas de convertibilité entre les francs CFA émis par la BCEAO (dans ce cas il s'agit des francs de la Communauté Financière d'Afrique) et les francs CFA émis par la BEAC (il s'agit ici des francs de la Coopération Financière d'Afrique centrale).

Dans le cadre d'un développement des économies, les monnaies utilisées au sein de la zone CIMA/FANAF présentent ainsi des avantages, tels que la stabilité et la fiabilité grâce à la parité fixe et à la convertibilité en euros (*cf.* Agbohou [2008] et Hugon [2009]). Cependant, ces monnaies présentent également des inconvénients, tels que les limites en termes d'ajustement économique dues à la parité fixe (*cf.* Ouédraogo [1999]) et les handicaps sur les échanges commerciaux intra-zone dus à l'absence de convertibilité entre les francs CFA de la BCEAO et ceux de la BEAC (*cf.* Agbohou [2008]).

Analyse et développement de l'activité économique

Pour expliquer les mauvaises performances économiques africaines, de nombreux facteurs généraux sont régulièrement avancés (*cf.* Hugon [2009]) : ethnolinguistiques (la zone compte plus de 230 langues et dialectes, *cf.* Frémy et Frémy [2005]), historiques (mauvaise spécialisation, poids de la colonisation), politiques (instabilité, insécurité, coûts de transaction - notamment de corruption - créée par l'autoritarisme), infrastructurels (faiblesse des télécommunications, électricité, transports, services publiques), internationaux (exposition aux chocs extérieurs), etc.

Des facteurs propres à l'économie sont également avancés par Hugon [2009], tels que le gaspillage des ressources (sachant de surcroît que les valorisations des rentes de situation sont également à l'origine de fortes fluctuations), le faible lien entre les rémunérations et la productivité des facteurs (caractérisé par un faible et, surtout, un mauvais investissement) et, entre autres, les distorsions entre les modes de production et de consommation (les Africains ne consommant pas ce qu'ils produisent).

Enfin, certains déséquilibres structurels des économies de la zone sont également matérialisés par des limites budgétaires. Ainsi par exemple, le système fiscal actuel ne présente pas les incitations nécessaires en matière de politique d'industrialisation et d'investissement (*cf.* Ouédraogo [1999]) et ne couvre pas certaines activités qui échappent alors à la fiscalité.

Au sein de la zone CIMA/FANAF, les facteurs justifiant les insuffisantes performances économiques sont donc nombreux et propres au contexte social, culturel, économique et politique.

Dans l'attente de politique de développement adaptée, permettant de redresser la situation économique et sociale dans la région, les populations ont été appelées à créer leur propre système de développement adapté : le secteur informel (*cf.* paragraphe suivant), système qui présente toutefois de nombreuses limites.

2.2.2. Économie informelle

Selon les estimations 2007 de l'OIT, l'emploi informel représente près des trois-quarts de l'emploi non agricole en Afrique subsaharienne. De surcroît, depuis la fin des années 1990, l'OIT considère que le secteur informel génère plus de 90 % des emplois urbains nouveaux (*cf.* Trani [2007]), c'est dire son importance. Aussi, ce paragraphe présente ce phénomène au sein de la zone CIMA/FANAF, en reconstituant l'origine de l'émergence de ce secteur, en donnant une définition à ce type d'activité et en décrivant les avantages et les inconvénients de cette activité au sein de la zone.

Origine de l'économie informelle

Le phénomène qui s'observait à partir de la décennie 1970 dans la plupart des pays de la zone, et qui demeure aujourd'hui, est que « le taux de croissance des industries est assez faible face à la croissance de la population en âge de travailler, ce qui a pour implication évidente une augmentation continue du nombre d'individus en situation de potentiels chômeurs. Le chômage n'étant généralement pas indemnisé (...), le surplus de main-d'œuvre, résultant de l'exode rural et de l'incapacité du secteur moderne à satisfaire la demande, entraîna non pas une augmentation exponentielle du chômage ouvert, mais bien au contraire la création de petites activités permettant aux migrants ruraux et aux citoyens n'ayant pu avoir accès aux emplois du secteur moderne de vivre et de survivre ». C'est en ces termes que Georges [2000] décrit la genèse de l'économie informelle.

Définitions de l'économie informelle

L'origine de l'économie informelle semble admise par tous. Il a toutefois été difficile d'obtenir un consensus autour de sa définition. En effet, depuis que l'OIT a publié son rapport sur la situation de l'emploi au Kenya en 1972, identifiant un type d'activité de travail qui ne correspondait ni aux activités du secteur « traditionnel » ni à celles du secteur « moderne », nombre d'études et de travaux sur ce phénomène ont vu le jour. Il existe quasiment autant de définitions que d'études, ce qui s'explique notamment par la forte hétérogénéité de ce secteur (taille variable des structures, unités et modes de production multiples, forte diversité des statuts des personnes qui y travaillent, etc., cf. Trani [2007]).

Selon Trani [2007], les organisations internationales limitent ce « secteur aux seules activités illégales dans leur fonctionnement, mais licites dans leur contenu ». Si cette approche a le mérite de la simplicité, elle manque de précisions. Aussi, de nombreuses tentatives de définition plus précise se sont développées (approche multicritères, approche fonctionnelle, etc., cf. Georges [2000]).

En vue d'harmoniser les traitements et les études statistiques sur ce secteur, la définition officielle suivante a été adoptée en 1993 (cf. Conférence internationale des statisticiens du travail [1993]) : « ensemble d'unités produisant des biens ou des services en vue principalement de créer des emplois et des revenus pour les personnes concernées. Ces unités, ayant un faible niveau d'organisation, opèrent à petite échelle et de manière spécifique, avec peu ou pas de division entre le travail et le capital en tant que facteurs de production. Les relations d'emploi - lorsqu'elles existent - sont surtout fondées sur l'emploi occasionnel, les liens de parenté ou les relations personnelles et sociales plutôt que sur des accords contractuels comportant des garanties en bonne et due forme ».

Avantages et inconvénients de l'économie informelle

Par définition, l'économie informelle assure un filet de sécurité sociale aux populations et leur permet de satisfaire des besoins fondamentaux. De manière plus générale, l'économie informelle a constitué un second souffle pour l'Afrique et a permis la mise en œuvre de nombreuses avancées. Ainsi, l'OIT estime que 70 % des activités commerciales dans le secteur informel sont assurées par des femmes (cf. Adamah-Tassah [2000]). Grâce à leur dynamisme dans le commerce informel, les femmes ont ainsi réussi à compenser leur faible participation au secteur moderne (due pour l'essentiel à la discrimination et à la faible scolarisation dont elles font l'objet) et à acquérir une autonomie économique et financière leur octroyant un rang social remarquable (c'est par exemple le cas des « Nana-Benz »⁸ du Togo entre les années 1970 et 1990). Une autre illustration de cette avancée est la capacité de l'économie informelle à satisfaire des besoins non couverts par une économie étroite, rigide et

⁸ « Nom donné aux riches et prospères femmes commerçantes jouissant d'un pouvoir économique leur permettant de rouler dans des voitures de prestige dont les limousines Mercedes-Benz » (cf. Adamah-Tassah [2000]).

réglementée. À ce titre, l'informel constitue un véritable régulateur économique et social et témoigne d'un dynamisme porteur d'avenir (*cf.* Hugon [2009]).

Système original et spontané développé initialement pour garantir la survie des populations, l'économie informelle est ainsi un témoignage d'ingéniosité qui permet de satisfaire les besoins du plus grand nombre et de moderniser la situation économique et sociale dans la région.

L'économie informelle présente toutefois de nombreuses limites. Certaines difficultés concernent directement les micro-entrepreneurs. Ainsi, Trani [2007] précise que faute de garanties suffisantes, les prêts bancaires sont généralement exclus de ces activités (les micro-entrepreneurs se tournent alors vers les tontines). Ce faible accès aux ressources financières contraint les acteurs à une faible productivité, à des revenus faibles et incertains, et à privilégier les visions de court terme (toute accumulation est limitée). En outre, les acteurs de ce secteur rencontrent régulièrement une barrière à l'accès des marchés publics et de sous-traitance des grandes entreprises. D'autres difficultés concernent le marché en général. À ce titre, l'un des problèmes que peut engendrer ce système est l'abolition de la barrière entre ce qui est légal et ce qui est illégal. Certains des effets de ce risque sont facilement observables, tels que les dérives de népotisme ou de corruption et les conséquences qui peuvent en découler (comme par exemple la dépendance à l'égard du fonctionnaire qui ferme les yeux sur l'exercice de l'activité illégale contre des avantages personnels). Par ailleurs, un autre grand problème des activités informelles est qu'elles ne garantissent aucune protection sociale aux travailleurs, ce qui accentue la précarité des participants (*cf.* Trani [2007]). En outre, on reprochera à l'économie informelle l'inadéquation entre les compétences techniques et les activités exercées (*cf.* Adamah-Tassah [2000]). Enfin, il apparaît que l'activité informelle élevée dans la région s'accompagne d'une faible bancarisation (égale au rapport entre le nombre de comptes bancaires et la population âgée de plus de 15 ans et estimée à moins de 15 % dans la région), ce qui limite la capacité de financement de l'économie.

2.2.3. Comportements microéconomiques, consommation et épargne des ménages

L'économie informelle répond au besoin du plus grand nombre. Mais quels sont ces besoins ? Quels sont les comportements socioéconomiques, et notamment microéconomiques, des populations de la zone ? Ce paragraphe s'attache à apporter des éléments de réponse à ces questions, en présentant les comportements microéconomiques de ménages, en identifiant leurs ressources et en décrivant leurs modes de consommation et d'épargne.

Comportements microéconomiques des ménages

En Afrique, « les rationalités microéconomiques des agents sont liées à leur appartenance à des réseaux et à des structures sociales » (*cf.* Hugon [2009]), par opposition à la rationalité orthodoxe de l'*Homo œconomicus*.

En d'autres termes, les agents économiques africains ne sont pas à la recherche de la frontière efficiente, mais d'une vertu communautaire et d'une

minimisation des risques. Ils sont rationnels et relationnels et agissent dans une logique d'économie affective et prudente.

Ressources, consommation et épargne des ménages

Au regard des comportements présentés ci-dessus, les consommateurs de la zone CIMA/FANAF semblent plus répondre à des normes qu'à des prix. En parallèle, on observe que les niveaux de consommation observés ne sont pas conformes aux revenus officiels (*cf.* Hugon [2009]). Les relations entre les revenus et la consommation d'une part, et entre les revenus et l'épargne d'autre part sont ainsi atypiques. Dans ce sous-paragraphe, il s'agit de comprendre cette situation.

Les revenus au sein de la zone sont généralement faibles et irréguliers. Néanmoins, la pauvreté du plus grand nombre s'accompagne de la richesse de quelques-uns, et les mécanismes « redistributifs » au sein des réseaux communautaires et informels de la zone permettent généralement aux plus pauvres de survivre et de consommer (à l'image de ce que l'on observait dans la solidarité traditionnelle ancestrale entre castes, *cf.* 1.1.1 pour mémoire).

Concernant les taux d'épargne, ils apparaissent fortement variables selon les ethnies (à titre d'exemple, les Bamilékés au Cameroun sont connus pour leur goût prononcé pour l'épargne). Par ailleurs, en termes d'investissements, « compte tenu de la préférence pour la liquidité et des différences de taux de rentabilité entre les investissements productifs et les investissements financiers ou spéculatifs (immobilier, import-export, placements extérieurs), les agents disposant de capitaux privilégient les activités financières et marchandes aux dépens des activités permettant (...) une accumulation productive » (*cf.* Hugon [2009]). Ce comportement se traduit entre autres par une fuite des capitaux importante (on estime que 40 % de l'épargne africaine est replacée sur les marchés financiers hors d'Afrique) et ne favorise pas le développement des productions nationales.

2.3. Conclusion sur le contexte socioéconomique

Le contexte social et économique au sein des pays d'Afrique subsaharienne francophone n'est pas favorable à la prospérité économique et sociale au sens des théories économiques orthodoxes : faible revenu, bien-être individuel relativisé par le bien-être communautaire, poids important des traditions ancestrales, politiques économiques inadaptées, etc. Ces pays ont toutefois pu s'adapter à ce contexte difficile, notamment grâce à la solidarité de leurs réseaux sociaux et communautaires et, surtout, grâce à une réaction sociale innovatrice : le développement de l'économie informelle (l'emploi informel permettant aux populations de subvenir à leurs besoins fondamentaux, et les biens et services issus de l'activité informelle permettant de répondre à des besoins jusque là non couverts par l'économie officielle). Le poids de l'économie informelle et le rôle des pratiques communautaires dans le fonctionnement de l'économie réelle témoignent ainsi de l'importance de la prise en compte du contexte socioéconomique dans la recherche de solutions de développement.

Dans cette section 2, comprendre le contexte socioéconomique en zone CIMA/FANAF a permis de comprendre certains aspects des problèmes d'économie générale et d'identifier les contours des solutions existantes aujourd'hui et envisageables pour le long terme. En rapprochant, dans la mesure du possible, les problèmes d'économie générale à des problèmes du marché de l'assurance vie, il conviendra dans la section suivante de développer et d'envisager des solutions de long terme pour le marché de l'assurance vie. Une attention particulière devra notamment être accordée à la prise en compte des spécificités liées au contexte.

3. Présentation des difficultés techniques et commerciales pour les assureurs vie du marché CIMA

Cette troisième section présente quelques problèmes du marché de l'assurance vie (*cf.* 3.1), les conséquences qu'ils engendrent au niveau du développement du marché (*cf.* 3.2) et les actions menées et/ou à mener pour y remédier (*cf.* 3.3).

3.1. Quelques problèmes historiques du marché

Le marché de l'assurance vie en zone CIMA/FANAF est historiquement marqué par des problèmes liés au contexte (*cf.* 3.1.1) ou à l'activité de l'assureur (outils financiers et techniques inadaptés et manque de personnel qualifié, *cf.* 3.1.2).

3.1.1. Difficultés liées au contexte

Les difficultés liées au contexte porte essentiellement sur l'escroquerie à l'assurance et les caractéristiques socioéconomique de la population.

Escroquerie à l'assurance

Pour présenter l'escroquerie au sein de la zone CIMA/FANAF, il convient d'en distinguer les causes et les manifestations, à l'image des travaux réalisés par Alla [2000].

La pauvreté, les fluctuations économiques et l'inadéquation entre l'offre et le contexte sont les principales causes de l'escroquerie à l'assurance au sein de la zone CIMA/FANAF. Dans la pratique, il apparaît que l'escroquerie est inhérente aux organismes assureurs, à leurs partenaires ou directement aux assurés (on notera que dans bien des cas, ces trois parties contribuent simultanément à l'escroquerie). Au niveau des organismes assureurs, l'escroquerie résulte des insuffisances dans les compétences du personnel (les failles dans la production des polices et dans l'authentification des sinistres profitent aux assurés indécents et bien informés) et de la course effrénée vers le chiffre d'affaires (les nombreux cas de sous-tarifcation et d'absence de sélection des risques permettent d'augmenter rapidement le chiffre d'affaires de l'assureur, mais offrent également un terrain très fertile à certains assurés opportunistes). Au niveau des partenaires (agents généraux, courtiers, experts médicaux, etc.), les escroqueries ont essentiellement pour origine l'absence de respect d'une déontologie ou d'une éthique. Enfin au niveau de l'assuré, l'escroquerie a pour origine le manque de culture de l'assurance classique : le

contrat d'assurance doit être individuellement rentable pour les assurés, il doit rapporter (même si cela engendre des comportements d'escrocs).

Concernant les manifestations de l'escroquerie, il convient d'identifier celles qui ont lieu avant le sinistre et celles qui ont lieu lors de la réalisation du risque. D'une part, avant le sinistre l'escroquerie se manifeste par de fausses déclarations (avec pour l'assuré l'objectif d'obtenir un tarif plus faible ou une indemnité plus élevée) ou par l'absence de communication sur les éventuelles évolutions du risque en cours de vie de contrat. D'autre part, lors de la réalisation du risque, l'escroquerie est matérialisée par la simulation de sinistres, par la survenance de sinistres réels mais intentionnels ou par la surévaluation des prestations au titre d'un sinistre.

Caractéristiques socioéconomiques de la population

Les caractéristiques socioéconomiques des individus de la région constituent également un frein historique au développement de l'assurance vie. L'essentiel de ces caractéristiques a été présenté dans la section 2.

Ainsi, sur le plan social on rappelle ici qu'au sein de la zone CIMA/FANAF, dans certaines traditions africaines ou religieuses : la conception de la vie favorise l'attentisme, la réalisation de profits suite à l'évènement d'un malheur est considérée immorale et enfin l'intérêt, l'incertitude et la prise de risque sont interdits dans les transactions commerciales. En outre, sur un plan psychologique il apparaît bien souvent que la mort est une réalité traumatisante dont on n'aime pas parler, que l'absence de concret ne satisfait pas les populations et qu'il existe une forte incompréhension des produits d'assurance par les prospects.

De même, sur un plan économique, il apparaît que nombre d'Africains de la zone CIMA/FANAF disposent de revenus faibles et incertains, et privilégient ainsi les visions de court terme. Par ailleurs, les rationalités microéconomiques des agents africains semblent liées à leur appartenance à des structures sociales : ils agissent dans une logique d'économie affective et prudente et leur mode de consommation semblent plus répondre à des normes qu'à des prix. Enfin, il apparaît une faible bancarisation dans la région.

Comme évoqué dans la section 2 toutefois, ces caractéristiques ne traduisent pas une incompatibilité entre les caractéristiques socioéconomiques de la région et le développement de l'assurance vie. Elles illustrent seulement la nécessité pour les assureurs de proposer des solutions innovantes et adaptées, permettant de tenir compte de ces spécificités. À titre d'exemple, les difficultés liées au faible poids de la population possédant un compte bancaire (estimé à moins de 15 %, selon la Banque mondiale) pourraient être résolues avec le développement des services financiers sur téléphonie mobile (le poids de la population possédant au moins un téléphone mobile étant estimé à plus de 35 % dans la région, selon les Nations Unies).

3.1.2. Difficultés spécifiques au rôle de l'assureur

Les assureurs sont donc appelés à proposer des solutions adaptées pour développer le secteur de l'assurance vie au sein de la région. En pratique

toutefois, il apparaît que les assureurs disposent généralement de moyens limités pour atteindre cette ambition, que ce soit au niveau des dispositions opérationnelles (personnels, données) ou au niveau des outils financiers et techniques utilisés.

Limites opérationnelles (imperfection des données, manque de personnel qualifié)

Si l'offre de produits traduit un manque d'originalité et n'illustre pas l'ingéniosité qui caractérise les populations africaines, c'est principalement en raison d'un manque de données de qualité et de personnel qualifié.

En effet, l'une des difficultés importantes pour les assureurs vie de la région se situe au niveau des insuffisances dans la qualité des données et des systèmes d'information.

Par ailleurs, il existe également des difficultés importantes au niveau du personnel et du processus de décision interne.

Sur ce point, les difficultés s'illustrent d'une part au niveau de manque de personnel qualifié. Certes la CIMA dispose de l'IIA au sein de laquelle de nombreux cadres et (futurs) dirigeants de qualité sont formés en assurance. Néanmoins, la zone CIMA manque manifestement d'actuaire (on estime à quelques dizaines le nombre d'actuaire exerçant en assurance vie au sein de la zone CIMA), alors même que ces techniciens sont indispensables pour toute création de solutions d'assurance innovantes. De manière générale, et au-delà des services techniques, les services commerciaux et marketings, également nécessaires à la création de solutions adaptées, manquent régulièrement de personnels qualifiés en leur sein.

D'autre part, il apparaît bien souvent que les personnes qualifiées au sein du marché ont le sentiment de ne pas bénéficier de l'attention de leur supérieur. À ce sujet, nous retenons deux exemples. Le premier est que de nombreux jeunes actuaire exerçant au sein de la zone ressentent un manque de considération et, dans une région très marquée par le respect voué aux anciens, voient régulièrement leurs propositions innovantes perçues comme une forme d'arrogance par leurs aînés. Le second est que « trop de petites sociétés d'assurances appartiennent en Afrique à des hommes d'affaires qui n'ont qu'une vision strictement utilitaire de leur société (...) » (cf. Yeatman [2007]), et « les Directeurs Généraux, salariés de ces hommes d'affaires, deviennent bon gré mal gré des exécutants de mauvaises politiques de gestion » (cf. Ziguélé [2008]), indépendamment de leur propre compétence⁹.

Opportunités limitées dans la gestion financière des placements

À supposer que les assureurs disposent de personnels qualifiés pour le développement de produits et donc du marché, force est de reconnaître que

⁹ De surcroît, dans ces situations, les services de contrôle nationaux avaient, avant l'instauration de la CRCA, rarement un poids politique suffisant pour s'imposer face à ces propriétaires de sociétés d'assurance, ces derniers ayant souvent un accès direct au ministre des Finances, patron de ces contrôleurs nationaux, ou au chef de l'État (cf. Yeatman [2007]).

l'activité de ces derniers serait fortement limitée par la situation des marchés financiers locaux d'autre part et par les tables de mortalité réglementaires d'une part.

Concernant le premier point, les opportunités de placements sur les marchés financiers sont limitées aujourd'hui, ce qui ne permet ni d'optimiser la gestion actif-passif, ni d'offrir aux bénéficiaires des taux de rendement substantiels. Plusieurs éléments expliquent cette situation.

D'une part, comme le rappelle Kone [2007], il convient « de s'interroger sur la viabilité, l'efficacité et la rentabilité des [trois] marchés financiers de la zone CIMA » : la BRVM, la BVMAC et la DSX. La BRVM, la bourse des valeurs mobilière de l'UEMOA, a débuté ses activités en 1998. À fin 2007 la BRVM comptait une quarantaine de sociétés inscrites à la cote action, mais à cette date seuls deux titres concentraient plus des deux tiers des échanges sur l'exercice, en valeur et en volume (cf. Banque de France [2008]). On comprend dans ce contexte que la BRVM est confronté à des problèmes de dynamisme, de liquidité et de profondeur. La situation est par ailleurs plus alarmante au sein de la zone CEMAC. La sous-région d'Afrique centrale compte effectivement deux bourses distinctes, plus jeunes et plus limitées en termes d'opportunités : la DSX et la BVMAC. La DSX se définit comme la bourse des valeurs mobilières sur le territoire camerounais et ne propose que deux actions à fin 2008 (la première action a été cotée en juin 2006). La BVMAC, dont le siège est à Libreville, est par ailleurs la bourse des valeurs mobilière de la CEMAC. Elle a démarré ses activités de cotation en août 2008 avec un emprunt obligataire de l'État gabonais.

D'autre part, le Code CIMA impose de fortes restrictions en matière de placements des actifs en représentation des engagements en dehors de la zone CIMA (cf. art. 335 à 335-2 du Code des Assurances de la CIMA).

L'absence de délocalisation des capitaux et l'état des marchés financiers locaux conduit alors à une surreprésentation des liquidités bancaires dans les placements des assureurs vie de la zone : 37% des placements sont en liquidités bancaires, 32 % en obligations, 10 % en actions, 9 % en immobilier et 12% en prêts et autres placements (cf. FANAF [2011]). Cette situation limite ainsi les promesses de taux de rendement attractifs et entrave la mise en place de couvertures adaptées des engagements. De surcroît, comme le rappelle Kone [2007], ces contraintes ne permettent pas à l'assurance vie « de participer de façon optimale au financement des économies africaines », alors même qu'il s'agit là d'un des rôles majeurs que doit jouer ce secteur pour le développement.

Tables de mortalité réglementaires perfectibles

L'autre limite porte sur les tables réglementaires pour la tarification et le provisionnement. Il s'agit aujourd'hui des tables TD CIMA pour les assurances en cas de décès et les tables TV CIMA pour les assurances en cas de vie (art. 338 et 334-4 du Code des Assurances CIMA). Ces tables se révèlent être respectivement les tables PM60-64 et PF60-64, c'est-à-dire les tables réglementaires en France du début des années 1970 au milieu des années 1980.

Ces tables sont construites sur la base des données démographiques françaises observées entre 1960 et 1964, respectivement pour les hommes et pour les femmes. En d'autres termes, aujourd'hui pour modéliser la mortalité des assurés de la zone CIMA dans le cadre d'une tarification ou d'un provisionnement, on s'appuie sur la mortalité de la population générale française du début des années 1960.

3.2. Conséquences des problèmes du marché

Ces problèmes historiques au sein du marché CIMA ont des répercussions importantes : déficit de confiance, existence d'une offre non adaptée et existence d'une importante population non couverte. Tous ces points sont traités dans cette sous-section.

3.2.1. Problème de confiance

Les relations entre les citoyens et les assureurs ne sont guère satisfaisantes, et cela à deux titres (*cf.* Yeatman [2007]) : d'une part, les assureurs redoutent les sinistres fictifs et les autres prétextes imaginés par les assurés pour recevoir des prestations, et d'autre part, nombre d'assurés « ne sont pas convaincus de la bonne foi des assureurs en ce qui concerne leur promesse de payer les sinistres en échange des cotisations qu'ils exigent ». Dans les faits, en raison de difficultés de trésorerie, on observe de nombreux retards dans le paiement des sinistres et on constate que de nombreux rachats et prestations échues ne sont pas payés aux bénéficiaires depuis plusieurs mois, voire plusieurs années (*cf.* Kone [2007]). Aussi, l'assuré ne comprend pas que l'on refuse de lui payer dans les délais impartis une épargne qu'il a constituée au prix de plusieurs sacrifices, et dans ce contexte la crédibilité des opérations d'assurance vie est gravement affectée.

Par ailleurs, concernant les relations entre l'État et les citoyens, il semble acquis que la méfiance prévaut sur la confiance et que l'informalité et la corruption (à grande et petite échelle) prévalent sur les règles. Cette situation reflète d'ailleurs l'état d'esprit et les mentalités du plus grand nombre au sein de la zone, et on en retrouve donc également des symptômes dans les relations entre l'État et les assureurs, notamment en termes d'harcèlement fiscal. En effet, nombre de contrôles fiscaux au sein de la zone s'apparentent à des intimidations et servent de prétexte pour monnayer un avantage financier indu. Ainsi, selon le CIAN [2008], parmi les 12 pays de la zone, seuls le Togo et le Gabon ne figurent pas dans la liste des 21 pays d'Afrique où la fréquence des contrôles fiscaux est anormalement élevée. Pis, cinq pays de la zone CIMA/FANAF (Cameroun, Centrafrique, Congo, Côte d'Ivoire, Mali) figurent parmi les sept pays d'Afrique où les « contrôles fiscaux [sont] complètement démentiels ».

3.2.2. Offre non adaptée

Les problèmes historiques justifient également les différences entre l'offre d'assurance vie et le contexte du marché CIMA/FANAF.

En pratique, il apparaît que l'offre actuelle est principalement le fruit de l'héritage colonial (*cf.* 1.2.1 pour mémoire). Elle est historiquement adaptée

aux populations occidentales expatriées, mais ne répond pas aux attentes de l'essentiel de la population au sein de la région (compte tenu de leurs caractéristiques socioéconomiques : croyances, traditions, religion, revenus, comportements microéconomiques, etc).

De surcroît, ces produits d'assurance vie doivent faire face à la concurrence des systèmes de solidarité traditionnels (notamment les tontines), qui ont le mérite de s'adapter aux revenus et à l'environnement socioculturel des populations.

3.2.3. Population non couverte

Aujourd'hui l'assurance vie classique s'adresse donc à une faible fraction d'agents économiques, et au sein de laquelle il existe une forte concentration de personnels salariés du secteur formel (entreprises ou administrations). C'est dans ce contexte que Kone [2007] rappelle que « les commerçants, les agriculteurs, les artisans, les transporteurs, les autres acteurs du secteur informel et les professions libérales sont dans la plupart des cas sans couverture sociale » (hormis celles des systèmes et réseaux de solidarité). En rappelant qu'en Afrique subsaharienne près de 70 % de la population active est dans le secteur primaire (cf. Hugon [2009]) et que l'emploi informel représente près de 75 % de l'emploi non agricole, on comprend qu'une grande proportion de la population active de la zone CIMA/FANAF n'est pas couverte pas les produits d'assurance vie classique.

Par ailleurs, certaines pratiques de l'assurance vie classique sont interdites par la charia, or pour mémoire la moitié de la population de la zone musulmane (cf. 2.1.3). S'il ne fait aucun doute que certains musulmans usent des contrats d'assurance vie classiques, il ne fait aucun doute non plus que de nombreux musulmans ne sont pas couverts par les contrats d'assurance vie proposés sur le marché.

3.3. Exemples d'actions menées et perspectives

En réponse aux problèmes évoqués ci-dessus, cette sous-section s'attache à présenter quelques actions menées, notamment par la CIMA (cf. 3.3.1), et quelques actions à mener, notamment par les États (cf. 3.3.2) et par les assureurs (cf. 3.3.3).

3.3.1. Actions menées par la CIMA

Parmi les actions menées par la CIMA il convient de noter la mise en place de tables de mortalité réglementaires prenant en compte l'expérience de la population assurée, ainsi qu'un assouplissement des règles relatives aux placements financiers (processus en cours).

Par ailleurs, concernant l'escroquerie à l'assurance de la part des assureurs, force est de reconnaître que depuis l'instauration d'une autorité de contrôle supranationale lors de la création de la CIMA, à savoir la CRCA, l'activité de contrôle a été épurée, notamment grâce à une plus grande intégrité.

Aussi, Ziguélé [2008] précise que depuis l'entrée en vigueur du Traité et du Code CIMA en 1995, « un important travail d'assainissement du secteur

s'effectue par la CRCA, malgré les difficultés inhérentes à un environnement naguère peu coutumier d'une discipline forte ». Ce travail a été possible car la CRCA « a pu recruter des personnels de haut niveau de compétence que leur statut de fonctionnaires internationaux met davantage à l'abri des pressions de toutes sortes dont sont victimes les fonctionnaires locaux » (*cf.* Yeatman [2007]).

Ce travail d'assainissement est possible grâce à un contrôle sur place et sur pièces des normes de solvabilité, qui consiste entre autres à vérifier la constitution de provisions techniques suffisantes, la couverture des engagements réglementés et l'existence d'une marge de solvabilité satisfaisante (*cf.* Ngbwa [2007]). À ce titre, Ouédraogo [2008] précise que la Commission a déjà pris les décisions et sanctions suivantes entre la 1^{ère} (en 1995) et la 49^{ème} (en 2007) Session Ordinaire de la CRCA¹⁰ : 19 retraits de la totalité des agréments, 6 mesures de restriction de la libre disposition des actifs par les dirigeants, 10 mesures de mise sous administration provisoire, 12 suspensions de dirigeants, 26 mesures de mise sous surveillance permanente, 11 blâmes et 11 avertissements aux dirigeants. En outre, la Commission a examiné 254 plans de financement ou de redressement présentés par des sociétés, dont plusieurs ont été jugés non satisfaisants.

3.3.2. Actions à mener par les États

Les États ont également leur rôle à jouer pour le développement du marché. Ainsi, l'État est en première ligne au sujet du manque de confiance qui existe au sein du marché. S'il ne serait pas juste de considérer que cet état d'esprit est uniquement dû au comportement de l'État, il semble naturel que ce soit l'État qui donne l'exemple et qui montre la voie de la sincérité, armes puissantes pour redonner du crédit aux autorités et pour restaurer un climat de confiance au sein du marché.

Par ailleurs, l'État dispose également d'importants leviers pour contribuer au développement du marché par des mesures d'incitations fiscales, que ce soit au niveau des cotisations (phase de constitution), des prestations (phase de restitution), voire directement sur l'activité des organismes assureurs. Si la crainte des États est l'existence d'un manque à gagner, Kone [2007] rappelle qu'une fiscalité moins restrictive permettrait aux assureurs de collecter plus d'épargne qui serait investie soit dans divers secteurs de l'économie soumis à divers impôts, soit dans des obligations d'État qui leur procureraient des ressources à long terme nécessaires aux projets de développement. Enfin, on peut noter à titre de comparaison que si le secteur de l'assurance vie est si développé en France, c'est précisément du fait de son statut fiscal favorable.

3.3.3. Actions à mener par les assureurs

Enfin, les assureurs pourraient également contribuer au développement du marché de l'assurance vie en zone CIMA/FANAF en proposant des solutions d'assurance adaptées aux populations et au contexte. À cet effet, on peut par exemple citer le développement de produits associés à la solidarité traditionnelle, notamment aux tontines accumulatives.

¹⁰ Chiffres relatif aux sociétés vie et non vie.

L'intérêt des tontines pour le développement de l'assurance vie classique s'illustre à quatre niveaux.

Premièrement, les tontines présentent dans bien des cas des fonctionnements proches de ceux observés en assurance vie classique : elles permettent de mutualiser les risques entre tous les membres qui acquittent une cotisation préalable et offrent des droits à prestations en cas de réalisation des risques assurés (décès ou hospitalisation par exemple). La compatibilité entre les tontines et l'assurance vie classique est donc immédiate (à titre d'exemple en France, le Groupe Mutualiste Le Conservateur propose des tontines et cette activité fonctionne bien)

Par ailleurs, les tontines concernent toutes les strates socioéconomiques de la population africaine : ruraux, urbains, riches, pauvres, etc. Ainsi par exemple, il existe des tontines pour les plus pauvres et les plus vulnérables (dans ce cas la tontine est une solution importante pour gérer les risques financiers menaçant les membres) et pour les plus riches et les plus puissants (dans ce cas les risques financiers menaçants les membres sont limités, et la tontine est alors principalement un lieu de rencontre permettant de constituer un réseau d'influence sociale). En s'associant aux tontines, qui couvrent donc une population très large, les assureurs auraient ainsi la capacité d'autoriser une grande mutualisation (entre les populations les plus risquées et les populations les moins risquées, entre les populations les plus fortunées et les populations les moins fortunées, etc.).

Troisièmement, les tontines s'inscrivent naturellement dans le respect des traditions et des spécificités socioéconomiques des membres, à travers leur mode de fonctionnement, leurs fêtes annuelles, leur mode de célébration d'événements heureux ou malheureux, etc. Les barrières psychologiques ou celles liées aux croyances et aux religions disparaissent donc, ce qui est favorable au développement de l'assurance vie.

Enfin, les tontines présentent l'avantage de proposer un cadre dans lequel les membres retrouvent la confiance qu'ils n'accordent plus aux institutions (assureurs classiques, États, organismes publics de prévoyance sociale). La confiance étant à la base de toute activité d'assurance, on comprend que ce cadre présente un intérêt significatif pour les assureurs.

Pour toutes ces raisons, les assureurs ont donc tout intérêt à considérer ces tontines comme une opportunité de développement, et non comme un concurrent déloyal (car informel). Les assureurs pourraient ainsi proposer aux tontines des contrats collectifs ne modifiant pas leurs règles de fonctionnement internes (si chères à leurs membres), en se chargeant principalement de l'encaissement et de la gestion financière d'une partie des fonds accumulés, de la mutualisation des risques entre tontines et de la responsabilité du versement de tout ou partie de la prestation. De tels contrats collectifs présenteraient l'avantage de permettre aux assureurs d'élargir significativement leurs populations couvertes, tout en limitant les démarches commerciales individuelles. L'intérêt de tels contrats pour les tontines souscriptrices seraient alors de bénéficier d'une réduction du coût de l'assurance, grâce à une plus

grande mutualisation des risques (compte tenu notamment de la capacité de l'assureur à mutualiser les risques de différentes tontines).

À l'image de la microassurance, qui pour mémoire désigne « l'adaptation de services d'assurance à des clients, essentiellement à faibles revenus, n'ayant pas accès à des services d'assurance classique » (cf. Nabeth [2007]), cette solution permet de fournir aux populations les plus vulnérables des solutions d'assurance adaptées au contexte socioéconomique. À la différence de la microassurance toutefois, l'association entre les tontines et les assureurs permet de limiter la démarche commerciale et le travail de pédagogie, et permet d'assurer un plus large périmètre mutualisation (cf. Nabeth [2007] pour une comparaison de l'assurance et de la microassurance en Afrique, Nabeth [2006] pour une présentation complète des défis de la microassurance et DID [2011] pour un état des lieux de la microassurance en zone CIMA).

4. Synthèse et conclusion¹¹

Dans les pays d'Afrique subsaharienne francophone, assimilés aux pays de la zone CIMA/FANAF, le marché de l'assurance vie classique, c'est-à-dire celle des assureurs privés, est à la traîne : il y représente moins d'un tiers du chiffre d'affaires de l'assurance classique, contre plus de deux tiers en France. Aussi, le taux de pénétration moyen de l'assurance vie classique au sein de la zone CIMA/FANAF est inférieur à 0,5 %, contre plus de 3 % pour toute l'Afrique (dont plus de 10 % en Afrique du Sud – deuxième meilleur taux mondial derrière Taïwan –) et plus de 7 % en France.

Les difficultés et limites de l'assurance vie classique

Les problèmes associés au développement de l'assurance vie classique en zone CIMA/FANAF sont globaux, et non uniquement techniques. Ils peuvent toutefois être répartis en trois classes interdépendantes.

Une première classe relève des problèmes liés à l'environnement social. Il s'agit ici des préjugés d'ordres socioculturels et religieux d'une part (à titre d'exemple, de nombreuses dispositions de l'assurance vie classique sont interdites par la charia alors que la moitié de la population de la zone est musulmane) et de la saine concurrence des systèmes d'assurance vie informels d'autre part (à titre d'exemple on peut souligner l'importance des tontines dans la zone).

Une deuxième classe de problèmes relève du contexte économique. Parmi les freins d'ordre économique, on retient essentiellement les revenus faibles et fluctuants des populations (en Afrique subsaharienne, près de 70 % de la population active est dans le secteur primaire et près de 75 % de l'emploi non agricole correspond à de l'emploi informel), leurs comportements économiques souvent dictés par des normes sociales et ethniques (la rationalité affective et prudente de l'agent économique africain diffère de la rationalité orthodoxe de l'*Homo oeconomicus*) et l'absence de climat de confiance (qui émane des autorités et des pouvoirs politiques).

¹¹ Cette conclusion s'appuie principalement sur l'article Kamega [2010], relatif aux défis de l'assurance vie en Afrique noire francophone.

La troisième classe de problèmes concerne les caractéristiques du marché de l'assurance vie. À ce niveau, les difficultés sont directement liées à l'offre, et plus précisément à l'inadéquation entre l'offre (généralement d'inspiration occidentale) et le contexte socioéconomique (propre aux pays d'Afrique subsaharienne francophone). Dans cette classe de problèmes, on inclut également la carence de personnel qualifié, le manque de données fiables et pertinentes et les limites en termes d'outils techniques (les tables de mortalité réglementaires PF 60-64 et PM 60-64 sont inadaptées) et financiers (l'insuffisance des marchés financiers locaux et les contraintes réglementaires sur la délocalisation des placements conduisent à une surreprésentation des liquidités bancaires – plus de 35% – dans les placements)¹².

Les principales conséquences de ces problèmes sont l'escroquerie à l'assurance (inhérente aux assureurs, à leurs partenaires ou aux assurés), l'insuffisante couverture de certains risques (à titre d'exemple le risque de mortalité est souvent mal évalué) et l'absence de couverture pour la plus grande partie de la population (on estime à moins de 10 % la part de la population de la zone couverte par l'assurance vie classique, cette forme d'assurance se limitant souvent au personnel du secteur formel, d'entreprises privées ou d'administrations).

Dans cette analyse du marché, il convient toutefois de distinguer la sous-région des pays d'Afrique de l'Ouest (sous-région UEMOA, hors Guinée Bissau) et celle des pays d'Afrique centrale (sous-région CEMAC, hors Guinée Équatoriale), certains des freins potentiels identifiés pouvant être plus marqués dans une sous-région que dans une autre. En effet, on observe par exemple au sein des pays d'Afrique de l'Ouest des dispositions fiscales plus incitatives, des taux de mortalité plus faibles, des taux de prévalence au VIH/Sida plus faibles, une perception de la corruption plus faible, etc.

Un potentiel de croissance qui prend forme

Que ce soit en Afrique de l'Ouest ou en Afrique centrale, force est de reconnaître que l'assurance vie classique au sein des pays de la zone CIMA/FANAF représente un réel potentiel de croissance, et cela à quatre titres. Premièrement, le taux de pénétration du marché y est très faible. Deuxièmement, les systèmes de solidarité traditionnels (c'est-à-dire l'assurance vie familiale ou communautaire), très importants dans la zone, présentent en l'état de nombreuses limites (faible périmètre de mutualisation, absence de tarification des risques, etc.). Troisièmement, la zone compte aujourd'hui plus de 125 millions d'habitants et devrait en compter près de 300 millions d'ici 2050, dont une forte proportion de jeunes (le taux de dépendance vieillesse y est égal à 6 % aujourd'hui – contre 25 % aujourd'hui en France – et devrait rester sous les 10 % en 2050). Enfin, l'évolution des mentalités dans la zone

¹² Sur ces deux derniers points toutefois, des évolutions sont en cours, conformément aux résolutions des États généraux de l'assurance vie : de nouvelles tables de mortalité réglementaires sont en construction et une réglementation plus souple sur les placements se développe.

(conséquences du changement d'ère politique, de l'expansion des systèmes éducatifs, etc.) devrait être favorable au développement du secteur.

Ce dernier argument est sans nul doute le plus important et le plus appréciable. Ainsi par exemple, certains comportements funestes de nombreux dirigeants en Afrique subsaharienne francophone tendent à s'estomper, du moins dans le secteur de l'assurance vie : le népotisme et les pots-de-vin n'y font plus autorité. L'assainissement du marché et la recherche de solutions de long terme sont donc désormais une priorité, notamment sous l'impulsion de la CIMA et de la FANAF (comme en témoignent les États généraux de l'assurance vie tenus à Douala durant l'été 2007 et les États généraux de la micro-assurance tenus à Abidjan en octobre 2009¹³). Ces évolutions, impulsées par l'autorité de contrôle et l'association professionnelle, se traduisent par des évolutions structurelles au sein des organismes d'assurance vie et par une vive marginalisation des acteurs manquant de professionnalisme. Aussi, le marché est en voie de concentration et compte d'ores-et-déjà trois grands groupes régionaux, Colina, NSIA et Sunu, largement solvables et disposant de ressources financières conséquentes (à titre d'exemple, en fin 2008 l'investisseur américain Emerging Capital Partners a investi 35 millions d'euros dans le groupe NSIA). Avec le groupe international Allianz, ces groupes régionaux sont les principaux acteurs d'un marché de l'assurance vie promis à un bel avenir.

Le potentiel de croissance de l'assurance vie dans la région s'illustre déjà dans les chiffres puisque depuis 2004, le chiffre d'affaires du marché augmente en moyenne de plus de 15 % par an.

De nouveaux défis pour les acteurs du marché

Ce potentiel de croissance du marché est également à rapprocher de nouveaux défis, tant académiques que professionnels. Il s'agit ici, comme indiqué ci-dessus, d'intensifier l'adaptation des solutions aux besoins des populations. En d'autres termes, il incombe au secteur de revoir son positionnement technique et commercial afin de répondre aux exigences et contraintes de l'environnement local d'une part, et afin de couvrir des populations plus larges et plus représentatives d'autre part.

En termes d'évolutions techniques adaptées, les assureurs doivent se repositionner sur le traitement de l'assurabilité et sur la mesure et la gestion des risques. Concernant l'assurabilité, il convient ainsi d'apprécier les barrières socioéconomiques, la présence de risques systématiques et systémiques importants, l'existence de fraudes à l'assurance, les conséquences de l'imperfection des données, etc. Concernant la mesure et la gestion des risques, il s'agit par exemple d'élaborer un suivi pertinent et adapté de la mortalité, des rachats, de la politique de placement et de l'adossement actif/passif.

En termes d'évolutions commerciales adaptées, les assureurs pourraient développer des offres en collaboration avec les systèmes d'assurance vie traditionnels, des offres de produits d'assurance vie *takaful* (c'est-à-dire conformément à la charia), la micro-assurance vie (activité très répandue en

¹³ Cf. les dossiers techniques de ces États généraux sur <http://www.fanaf.com/>.

Afrique australe et en Afrique de l'Est, mais peu développée aujourd'hui en zone CIMA/FANAF) ou encore l'utilisation de la téléphonie mobile (qui connaît une expansion remarquable en Afrique) pour la distribution et la gestion des produits.

Qu'ils s'agissent des évolutions techniques ou commerciales, force est de constater que l'ensemble de ces domaines regorge de spécificités propres à la zone CIMA/FANAF et justifie ainsi la recherche de solutions endogènes, indispensables au développement pérenne de l'assurance vie classique dans la région.

Dans la suite de ce travail, on se concentre principalement sur la recherche de solutions endogènes dans le cadre des évolutions techniques, et non commerciales. En particulier, il s'agira d'étudier les spécificités liées à la mesure et à la gestion des risques en assurance vie dans la région. Pour mémoire, les risques en assurance vie s'illustrent principalement dans le résultat de mortalité, dans le résultat de gestion et dans le résultat financier. En pratique, dans la suite de ce travail on s'attachera à étudier les risques affectant le résultat de mortalité : une attention particulière sera ainsi accordée à la recherche de solutions adaptées pour mesurer et gérer les risques liés à la durée de vie humaine dans la région.

Bibliographie

- Adamah-Tassah D. A. [2000] *Assurance et secteur informel : quelles stratégies de pénétration ?*, Mémoire d'assurance, IIA.
- Agbohohou N. [2008] « Le franc CFA et l'euro contre l'Afrique », *Solidarité mondiale*, 2^{ème} édition.
- Aku Sama A. [1996] *Le marché togolais de l'assurance vie*, Mémoire d'assurance, IIA.
- Alla K. A. [2000] *L'escroquerie à l'assurance et ses implications, quelles solutions ?*, Mémoire d'assurance, IIA.
- BAfD/OCDE (Banque Africaine de Développement/Organisation de Coopération et de Développement Économique) [2008] « Perspectives économiques en Afrique 2007/2008 », OCDE.
- Banque de France [2008] « Rapport Zone franc - 2007 », *Banque de France - La politique et les agrégats monétaires dans l'UEMOA*.
- Betoe Minko M. [1986] *La promotion de l'assurance vie en Afrique : l'exemple du Cameroun*, Mémoire d'assurance, IIA.
- Bleu V. [1984] *Influence des paramètres socio-économiques sur le développement de l'assurance vie – Ex : Cameroun et Côte d'Ivoire*, Mémoire d'assurance, IIA Yaoundé.
- CIAN (Conseil français des investisseurs en Afrique) [2008] *Résister aux sollicitations indues dans le domaine fiscal*, Programme de la réunion de la Commission Anti-corruption du 20 novembre 2008, CIAN.
- CIMA (Conférence Interafricaine des Marchés de l'Assurance) [1992], « Traité instituant la CIMA », CIMA.
- CIPRES (Conférence Interafricaine de la PRÉVoyance Sociale) [1993], « Traité instituant la CIPRES », CIPRES.
- CIPRES (Conférence Interafricaine de la PRÉVoyance Sociale) [2007], « Les systèmes de protection sociale en Afrique : Zone CIPRES », CIPRES.
- Conférence internationale des statisticiens du travail [1993] « Résolution concernant les statistiques de l'emploi dans le secteur informel adoptée par la quinzisième Conférence internationale des statisticiens du travail », OIT.
- Daniel J.-P. [2007] « L'assurance islamique gagne du terrain », *La Tribune de l'assurance* 111.
- Debayou A. S. [1988] *Quelle assurance vie pour l'Afrique ?*, Mémoire d'assurance, IIA.
- Denuit M., Charpentier A. [2004] « Mathématiques de l'assurance non-vie – Tome I : principes fondamentaux de théorie du risque », *Economica*.

- Diakité D. [2008] « Universalité des valeurs et idéal d'humanité en Afrique : témoignages d'explorateurs », *extrait de* Konaré A. B. (sous la direction de) [2008] « Petit précis de remise à niveau sur l'histoire africaine à l'usage du président Sarkozy », *La Découverte*.
- Diallo D. S. [2008] « LA CIMA ou l'histoire d'un modèle d'intégration dans les assurances ».
- DID (Desjardins – Développement international) [2011], « Étude sur la microassurance dans la zone CIMA », *Desjardins – Développement international*.
- Diop C. A. [1979] « Nations nègres et culture », *Présence Africaine*, 4^{ème} édition.
- Direction des Assurances (République du Bénin) [2007], « Étude sur les ruptures anticipées de contrats d'assurance sur la vie », *République du Bénin, Ministère du développement, de l'économie et des finances, Direction générale de l'économie*.
- Ewald F. [2010] « La naissance du risque social », *Risques, Les cahiers de l'assurance* 81-82.
- FANAF (Fédération des sociétés d'Assurances de droits National AFricaines) [2011], « Le marché de l'assurance en Afrique – Données 2005 à 2009 », *FANAF*.
- FFSA (Fédération Française des Sociétés d'Assurances) [2011], « Rapport annuel 2009 », *FFSA*.
- Frémy D., Frémy M. [2005] « Quid 2006 », *Robert Laffont*.
- Georges I. [2000] *Assurance et secteur informel : quelles stratégies de pénétration*, Mémoire d'assurance, IIA.
- Haman Bello I. [1984] *Essai d'analyse de la faiblesse de l'assurance vie en Afrique*, Mémoire d'assurance, IIA.
- Hampaté Bâ A. [1972] « aspects de la civilisation africaine », *Présence Africaine*.
- Hugon P. [2009] « L'économie de l'Afrique », *La Découverte*, 6^{ème} édition.
- Juompan-Yakam C. [2010] « Retraites – Un continent en quête de cotisants », *Jeune Afrique* 2590.
- Kamega A. [2010] « Défis de l'assurance vie en Afrique noire francophone », *La Tribune de l'Assurance* 143.
- Kassim Z. [2008] « Takaful - Concepts et pratique », *Le Réassureur Africain*, 22^{ème} édition.
- Ki-Zerbo J. [1978] « Histoire de l'Afrique noire - D'hier à demain », *Hatier*.
- Kone M.G.K. [2007] « État des lieux de l'assurance vie dans les pays de la CIMA », *CIMA*.

- Nabeth M. [2006] « Micro-assurance – Défis, mise en place et commercialisation », *L'Argus de l'assurance*.
- Nabeth M. [2007] « Assurance et microassurance en Afrique », *Risques, Les cahiers de l'assurance 71*.
- Ndembou S. [2007] « La solidarité africaine relève de la générosité », *extrait de Courade G. (ouvrage dirigé par) [2007] « L'Afrique des idées reçues », Belin.*
- Ngbwa J.-C. [2007] « Apport de la CIMA à la solvabilité des entreprises d'assurances », *Risques, Les cahiers de l'assurance 71*.
- Nonyu Moutassie É. [2005] « Présentation de la CIMA », *CIMA*.
- Omvane Nkouélé F.-X. [1988] *L'assurance sur la vie et le financement de l'économie dans les pays de la CICA*, Mémoire d'assurance, IIA Yaoundé.
- Ouédraogo A. [2008] « Faiblesses actuelles des compagnies d'assurances et moyens susceptibles d'être mobilisés pour leur consolidation pérenne », *CIMA*.
- Ouédraogo O. [1999] « Plaidoyer macroéconomique pour l'Afrique », *Karthala*.
- Sow A.-M. L. [2007] « Ethnies et sociétés islamique en Afrique, un paradoxe ? – Le cas du Fuuta Dyalöö guinéen du XVI^e au XX^e siècles », *L'Harmattan*.
- Sylla Ndoye I. [1981] *Le développement de l'assurance vie dans un contexte islamique. Expérience de la société nationale d'assurance mutuelle vie*, Mémoire d'assurance, IIA.
- Trani J.-F. [2007] « L'économie informelle est la voie pour un développement à l'africaine », *extrait de Courade G. (ouvrage dirigé par) [2007] « L'Afrique des idées reçues », Belin.*
- Yeatman J. [2004] « L'assurance française en Afrique noire francophone », *Risques, Les cahiers de l'assurance 57*.
- Yeatman J. [2007] « L'assurance en Afrique : une émergence difficile », *Risques, Les cahiers de l'assurance 71*.
- Ziguélé M. [2008] « Comment renforcer les compagnies d'assurances africaines de la zone CIMA », *CAPafrique – Analyses sectorielles*.

Chapitre 2

Évolution technique du marché CIMA : construction de tables de mortalité d'expérience pour la zone¹

Sommaire

CHAPITRE 2 ÉVOLUTION TECHNIQUE DU MARCHÉ CIMA : CONSTRUCTION DE TABLES DE MORTALITÉ D'EXPÉRIENCE POUR LA ZONE	53
1. Présentation des données collectées.....	54
1.1. Collecte des données.....	54
1.2. Choix de la période d'observation	57
1.3. Constitution du périmètre final des données.....	57
1.4. Statistiques descriptives agrégées par segment.....	58
2. Construction des tables d'expérience.....	62
2.1. Construction des tables d'expérience.....	62
2.2. Extrapolation des taux ajustés.....	68
2.3. Fermeture des tables	71
2.4. Présentation des tables UEMOA et CEMAC.....	72
3. Contrôle des tables d'expérience construites et comparaisons	75
3.1. Cohérence des tables construites avec les observations.....	76
3.2. Cohérence des tables construites avec les tables OMS	78
3.3. Comparaison des tables construites avec les tables TD/TV CIMA	80
4. Synthèse et conclusion (et présentation de défis techniques complémentaires sur la mortalité).....	82
Bibliographie.....	84
Annexe A : retraitement des données	85
A.1. Liste des traitements	85
A.2. Détail des retraitements par assureur	87
Annexe B : mise à jour du périmètre	89

¹ Ce chapitre s'appuie principalement sur le rapport Planchet et al. [2010], relatif aux travaux de construction des tables réglementaires d'expérience de la zone CIMA.

Comme évoqué dans le chapitre précédent, les tables de mortalité en usage dans la zone CIMA pour les opérations d'assurance sur la vie sont basées sur les statistiques de mortalité de la population générale française observée entre 1960 et 1964 (tables PM/PF60-64). Cependant, les conséquences de multiples fléaux dont le VIH/Sida, la recrudescence des maladies infectieuses et la pauvreté conduisent ces tables à potentiellement créer un biais dans l'estimation de la mortalité réelle des populations concernées.

Dans ce contexte, la CIMA et la FANAF ont décidé d'élaborer des tables de mortalité prudentielles basées sur les statistiques de mortalité des populations d'assurés des pays de la zone CIMA. L'instauration de ces nouvelles tables de mortalité s'inscrit parmi les actions à mener pour la redynamisation de l'assurance vie dans la région (suite aux États généraux de l'assurance vie de 2007).

La réforme ainsi envisagée devrait permettre d'accompagner le développement de ce secteur et de renforcer sa solidité sur la base d'instruments de tarification fiables, en adéquation avec les risques couverts. Elle fournira notamment aux compagnies d'assurance vie les moyens d'effectuer, à l'aide d'outils techniques appropriés, une tarification plus précise du risque de décès et un provisionnement suffisant des contrats en portefeuille.

Dans ce cadre, ce chapitre présente chacune des étapes relatives à la méthode retenue pour l'élaboration de ces nouvelles tables, à savoir :

- la collecte et la validation des données disponibles ;
- la construction des tables par sexe pour chacune des sous-régions CEMAC et UEMOA (sachant toutefois qu'en pratique la CIMA privilégie au final l'utilisation d'une table unique par sexe pour la région²) ;
- l'analyse comparée des tables de mortalité en vigueur (tables TD/TV CIMA) avec les nouvelles tables ainsi établies.

Les travaux présentés dans ce chapitre sont principalement descriptifs : des informations complémentaires sur les principaux outils techniques utilisés ici sont fournies dans les chapitres suivants.

1. Présentation des données collectées

Cette section présente la collecte et l'exploitation des données utilisées pour la construction des tables de mortalité d'expérience pour la zone CIMA.

1.1. Collecte des données

La collecte des données est organisée en plusieurs étapes : demande des données nécessaires auprès des organismes, exploitation des données brutes transmises par les organismes et échanges avec les organismes sur les conclusions de l'exploitation des données brutes transmises en vue de la transmission d'une mise à jour.

² Obtenue par pondération des tables par sous-région.

De manière générale, une attention particulière est accordée à l'analyse de la fiabilité des données, celles-ci étant au cœur du processus de construction.

1.1.1. Demandes de données

Les demandes de données ont été formulées auprès de plus d'une quarantaine d'assureurs vie (sous l'autorité de la CIMA) et auprès des Caisses de Sécurité sociale (sous l'autorité de la CIPRES) de la région d'Afrique subsaharienne francophone (ces dernières ont été sollicitées pour équilibrer la répartition géographique des données au sein de la zone, étant donné que le nombre d'assureurs vie en Afrique central est limité).

Pour les assureurs vie, les données portent exclusivement sur des contrats offrant des garanties en cas de décès ou en cas de vie (à l'exclusion des contrats de rentes viagères). Les contrats de capitalisation ou d'épargne pure ne rentrent donc pas dans le périmètre retenu. En outre, les demandes de données sont relatives à la période d'observation s'étalant du 01/01/2002 au 31/12/2007.

Pour les Caisses de Sécurité sociale, les données portent exclusivement sur les futurs bénéficiaires de pensions vieillesse dont le décès engendre le règlement d'une prestation (pension ou allocation) pour les survivants. Les populations d'assurés déjà bénéficiaires de pensions vieillesse ou uniquement concernées par les prestations familiales, d'invalidité, d'accident du travail, etc. ne rentrent donc pas dans le périmètre retenu.

1.1.2. Première exploitation des données

La demande de données a permis de disposer de données de 37 assureurs vie et 3 Caisses de Sécurité sociale.

À la réception des données brutes, les premiers travaux d'exploitation sont menés à deux niveaux :

- détection des anomalies et des imperfections pour améliorer la fiabilité et la cohérence des données ;
- présentation des statistiques globales (statistiques sur les expositions au risque, l'âge moyen, le nombre de décès et le taux brut de décès moyen par année et par sexe) pour disposer d'une première tendance du volume de la population assurée et du niveau de la mortalité.

Ensuite, ces travaux ont fait l'objet d'une restitution auprès des assureurs vie uniquement. Pour les Caisses, les données n'ont jamais été reçues dans les délais impartis et après les premiers travaux d'exploitations, il s'avère que la fiabilité des informations sur les décès est très limitée : ces données ont donc été exclues du champ de l'étude.

Toutes les restitutions menées auprès des assureurs vie se sont conclues par une demande de mises à jour des données (notamment au regard des anomalies et imperfections identifiées) et ont ainsi permis de limiter le nombre de données à exclure du périmètre des données. À l'issue de la réception des mises à jour des données, une base de données harmonisée est construite. Sur les 37 assureurs vie initiaux, 5 ont été exclus du périmètre des statistiques descriptives en raison de mises à jour insuffisantes.

1.1.3. Retraitement des données

Les données issues des 32 assureurs conservés à ce stade ont ensuite fait l'objet de retraitement lors de la réception des mises à jour. Ce retraitement est présenté en détail à l'annexe A.

On retient ici le tableau ci-dessous, qui reprend le nombre de lignes initial, le nombre de lignes supprimées et le nombre de lignes après retraitements par sous-région.

Tableau 2 – Synthèse quantitative des retraitements de données

Synthèse quantitative	CEMAC	UEMOA	Total (CIMA)
Nombre de lignes avant retraitements	387 247	928 699	1 315 946
Nombre de lignes supprimées	62 921	50 552	113 473
Nombre de lignes après retraitements	324 326	878 147	1 202 473

On note d'une part que le poids des lignes supprimées pour la sous-région CEMAC est nettement supérieur à celui de la sous-région UEMOA (16 % du nombre de ligne initial pour la CEMAC, contre 5 % pour l'UEMOA) et d'autre part que le volume de données après retraitement pour la sous-région UEMOA est nettement supérieur à celui de la sous-région CEMAC (il lui est supérieur de 171 %, soit près du triple).

1.1.4. Pertinence et fiabilité des données

Au-delà de certains critères qualitatifs (dus par exemple à la conformité des données aux demandes formulées), l'appréciation de la pertinence et de la fiabilité des données est réalisée à partir de l'analyse des imperfections dans les données d'une part, et à partir des résultats des statistiques globales d'autre part.

Ainsi, la quantité et la nature des imperfections relevées dans les données individuelles permettent de juger la qualité du système d'information et donc la qualité des données brutes. Dans le cadre de l'exercice mené, une attention particulière est accordée à la qualité des données individuelles, celles-ci étant au cœur de la construction des tables de mortalité. Aussi, entre le 30 juillet 2009 (date de la réception des premières données) et le 23 octobre 2009 (date de clôture définitive pour la réception des données), l'essentiel des travaux a consisté à analyser les données individuelles et à entretenir des échanges (y compris visites sur place) avec les assureurs sollicités afin de fiabiliser les bases de données de l'étude.

Les statistiques globales établies sur les données permettent de compléter cette appréciation de la qualité des données. Sur ce point, il s'agit en particulier de s'assurer de la cohérence du niveau des taux de mortalité annuels sur la période d'observation de l'étude. Dans le cadre de la fiabilisation des données pour la construction des tables proprement dites, cette analyse permet notamment de justifier la suppression de certains jeux de données complets d'assureurs vie (ce point est présenté ci-après en 1.3).

1.2. Choix de la période d'observation

La période d'observation est l'intervalle de temps retenue pour l'étude de mortalité. Afin de limiter l'impact des fluctuations conjoncturelles de la mortalité et d'augmenter le volume d'observations, il est souhaitable que cette période comporte plusieurs années.

La date de fin d'observation est fixée pour que la période d'observation inclue les années d'expérience les plus récentes possibles. Il convient toutefois de tenir compte des déclarations tardives de décès afin de ne pas sous-estimer le risque de mortalité. Étant donné que certains assureurs ont indiqué lors des visites de travail qu'au regard des déclarations tardives le nombre de décès renseigné au titre de 2007 n'était pas exhaustif, il est convenu de ne pas tenir compte de cette année dans l'étude. La période d'observation retenue s'arrête ainsi le 31/12/2006.

Par ailleurs, dans le cadre d'une étude de mortalité instantanée, retenir une période de trois à cinq ans est un choix satisfaisant. Dans la présente étude, l'année 2002 a été exclue car de nombreuses sociétés n'existaient pas ou étaient très récentes à ce moment et il est généralement constaté qu'en période de fort développement de l'activité, les taux de mortalité évoluent sensiblement et biaisent ainsi les résultats globaux. En d'autres termes, la période d'observation retenue commence ainsi le 01/01/2003.

Aussi, dans le cadre de la présente étude, la période d'observation retenue s'étale du 01/01/2003 au 31/12/2006 et compte ainsi quatre années d'observation. Les données ont alors été retraitées de manière à ne retenir que les événements survenus pendant ce laps de temps. On retient ainsi que les données comptent des tronçatures à gauche (relatives aux entrées après le 01/01/2003) et des censures à droite (relatives aux sorties avant le 31/12/2006 pour un motif autre que le décès).

Aussi, dans l'ensemble de ce travail, il est précisé que plutôt que de travailler sur des effectifs (ou enregistrements), des « expositions au risque » sur la période d'observation (effectifs pondérés par leur durée de présence) sont utilisées, cette notion étant dans le présent contexte mieux adaptée que le simple décompte des personnes.

1.3. Constitution du périmètre final des données

Pour garantir une fiabilité statistique aux données utilisées pour la construction des tables, une mise à jour du périmètre des données est réalisée sur la base de trois critères (le détail des assureurs concernés par ces suppressions est présenté à l'annexe B).

Le premier critère porte sur le volume d'activité. En pratique, les jeux de données dont l'exposition au risque est inférieure à 1 000 sont supprimés du périmètre de la construction des tables (pour mémoire, une exposition inférieure à 1 000 sur la période d'observation correspond à un effectif assuré inférieur à 250 en moyenne par année).

Le deuxième critère porte sur le nombre de décès. Ici, il s'agit de supprimer du périmètre, après prise en compte du premier critère, les jeux de données pour lesquels aucun décès n'est observé.

Le troisième critère porte sur le niveau des taux de décès bruts et sur leur hétérogénéité au sein d'un pays. Ainsi, lorsque les taux de décès bruts sont anormalement bas, non justifiables et significativement erratiques au sein d'un pays (en l'occurrence lorsqu'au sein d'un même pays les taux entre assureurs varient de plus de 1 000 %) les jeux de données sont supprimés.

Tous les chiffres présentés ci-après tiennent compte de ces suppressions complémentaires dans le périmètre des données (pour information, le détail de ces suppressions est présenté au sein de l'annexe B).

1.4. Statistiques descriptives agrégées par segment

En vue d'apprécier la cohérence et l'homogénéité des données transmises après retraitements et mise à jour du périmètre, on présente ici des statistiques descriptives de base (exposition au risque, âge moyen, taux de décès moyen, etc.) par segment (la segmentation portant sur le sexe et la sous-région) et sur l'ensemble de la période d'observation.

Dans le cadre de ces statistiques, le tableau suivant présente les expositions au risque pour chacun des quatre segments de l'étude.

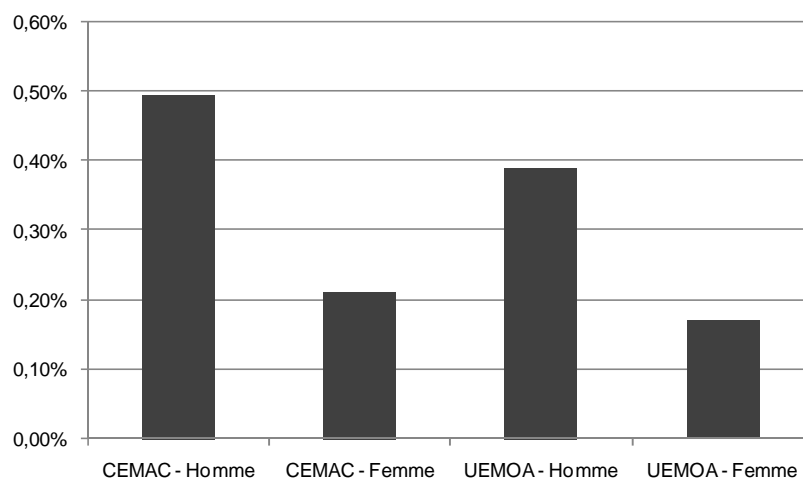
Tableau 3 – Répartition des expositions au risque par segment

Sexe / Sous-région	CEMAC	UEMOA	Total (CIMA)
Homme	467 557	783 105	1 250 662
Femme	139 899	161 219	301 118
Total (tous sexes)	607 456	944 324	1 551 780

On note que le volume final de données pour la sous-région UEMOA est supérieur à celui de la sous-région CEMAC (l'exposition de la sous-région UEMOA représente 61 % de l'exposition totale).

1.4.1. Statistiques descriptives agrégées par segment

Les taux de décès moyens par segment sont présentés sur le graphique suivant.

Figure 1 - Taux de décès moyen par segment

À première vue, ces résultats semblent cohérents car :

- les taux de décès de la population masculine représentent environ le double des taux de la population féminine, ce qui est cohérent avec l'écart généralement observé entre la mortalité des hommes et celle des femmes (*cf.* par exemple les tables réglementaires françaises TH/TF00-02) ;
- les taux de mortalité de la sous-région CEMAC sont supérieurs à ceux de la sous-région UEMOA, ce qui est cohérent avec les données de l'OMS et de l'ONU relative à la population générale.

Sur ce dernier point toutefois, on note que l'écart observé ici avec les données d'expérience est plus important que celui observé avec les données de la population générale (données OMS et ONU). En effet, il ressort des données d'expérience tous sexes confondus que les taux de mortalité dans la CEMAC sont 25 % supérieurs à ceux de l'UEMOA (pour un âge moyen de 42 ans), alors qu'au niveau de la population générale l'écart tous sexes confondus est de 16 % pour les 30-54 ans et de 12 % pour les 20-64 ans.

Cette différence entre la population assurée et la population générale pourrait toutefois s'expliquer en partie par des imperfections dans les données, au niveau de la population assurée comme au niveau de la population générale, et invite ainsi à la prudence. En particulier, on note que dans le cas de la population générale, les taux de mortalité UEMOA et CEMAC sont issus des tables de mortalité nationales de l'OMS³, par tranches d'âge de 5 ans et par sexe (au titre de l'année 2006), pondérées par les populations nationales correspondantes des Nations Unies⁴ (au titre de l'année 2005). Aussi, on rappelle ici que les tables de l'OMS sont développées dans le cadre spécifique

³*Cf.* http://apps.who.int/whosis/database/life/life_tables/life_tables.cfm

⁴Ou plus précisément de la Division de la population du Département des affaires économiques et sociales des Nations Unies (<http://esa.un.org/unpp/index.asp?panel=2>).

de l'évaluation d'espérance de vie en bonne santé (HALE) par l'OMS, et présente à ce titre certaines limites non discutées ici (*cf.* lien de la note de base de page ci-avant sur les données de l'OMS).

Le tableau suivant présente des statistiques complémentaires permettant de préciser et d'interpréter les observations de la Figure 1 (les intervalles de confiance sont obtenus en effectuant l'approximation usuelle d'une loi binomiale par une loi normale).

Tableau 4 – Statistiques par segment

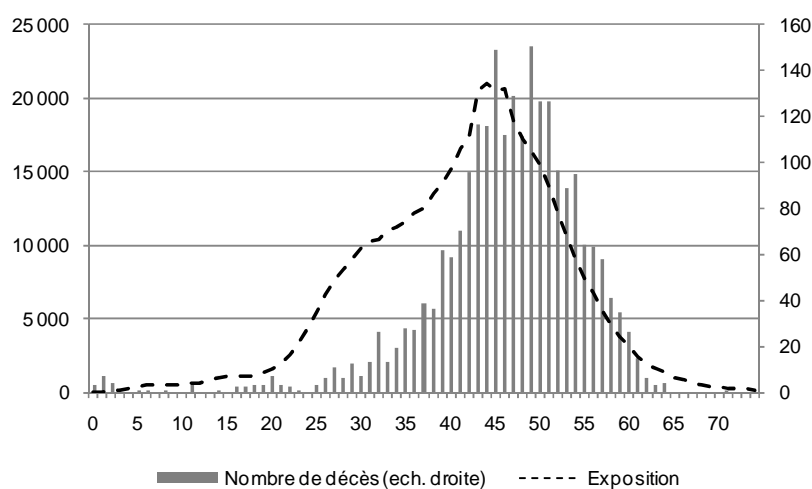
Segment	Exposition 2003-2006	Âge moyen	Taux de décès moyen	Taux de décès moyen (borne inf. à 95%)	Taux de décès moyen (borne sup. à 95%)	Âge moyen au décès
CEMAC - Homme	467 557	42 ans	0,4947%	0,4745%	0,5149%	46 ans
CEMAC - Femme	139 899	39 ans	0,2109%	0,1868%	0,2349%	41 ans
UEMOA - Homme	783 105	42 ans	0,3902%	0,3764%	0,4041%	47 ans
UEMOA - Femme	161 219	42 ans	0,1693%	0,1492%	0,1894%	45 ans
Total (CIMA)	1 551 780	42 ans	0,3826%	0,3729%	0,3923%	46 ans

On note ici que les taux de mortalité de la population féminine de la CEMAC sont supérieurs à ceux de la population féminine de la UEMOA, malgré un âge moyen et un âge moyen au décès plus faibles dans la CEMAC.

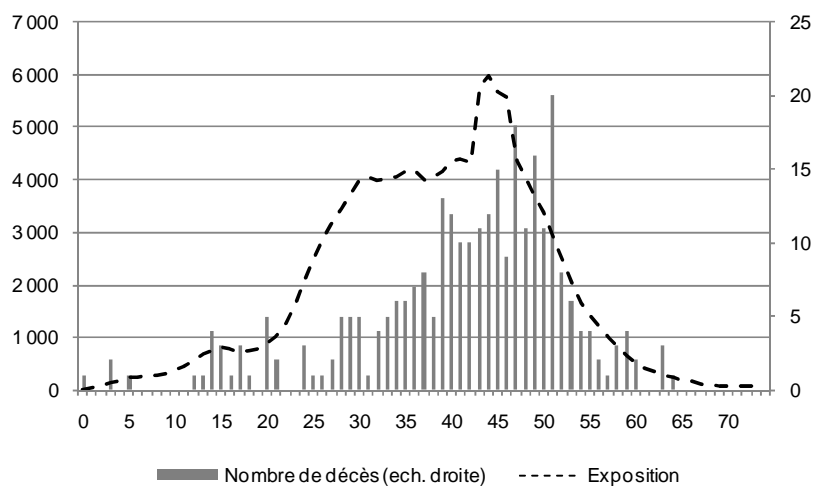
1.4.2. Illustration de l'exposition au risque et du nombre de décès par âge

Les illustrations suivantes présentent pour chaque segment les expositions au risque et les nombres de décès par âge.

Figure 2 - Exposition et décès (CEMAC, homme)

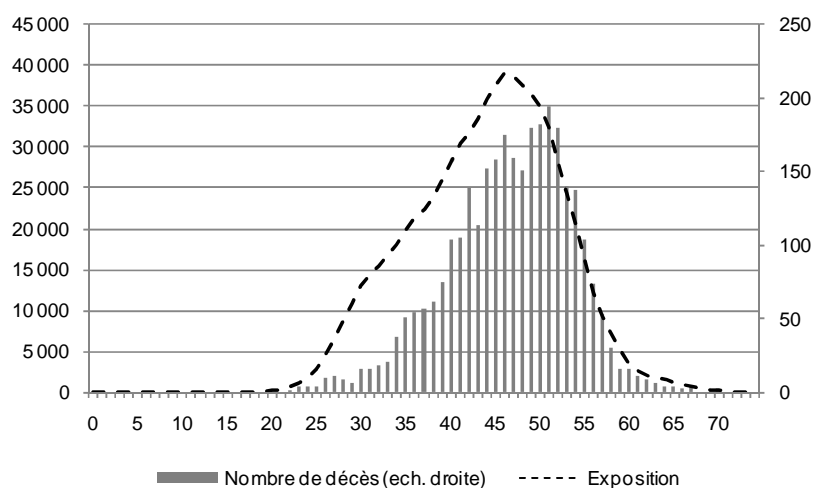


Le graphique ci-dessus met en évidence le faible nombre d'observation avant 20 ans et après 65 ans pour les hommes de la CEMAC.

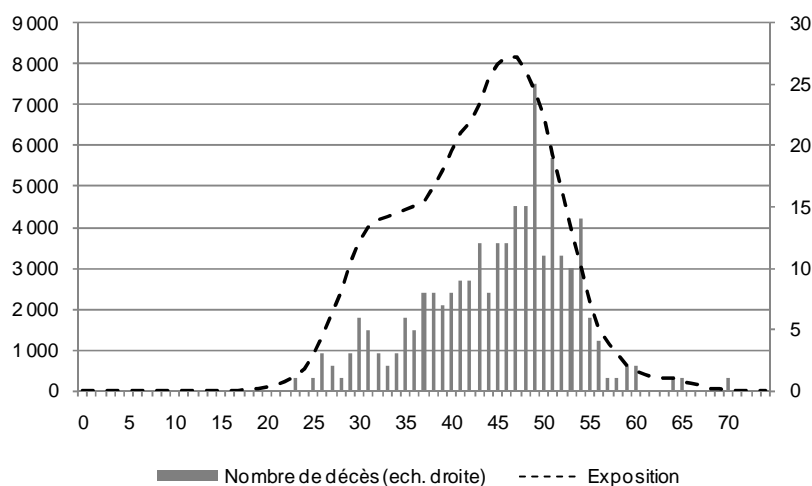
Figure 3 - Exposition et décès (CEMAC, femme)

La quantité d'information pour la population féminine de la sous-région CEMAC est plus limitée que celle de la population masculine. Aussi, on observe des « sauts » dans le nombre de décès et dans les expositions au risque (les fluctuations étant d'autant plus importantes que le volume d'information est faible).

Les deux figures suivantes présentent les mêmes illustrations pour la sous-région UEMOA.

Figure 4 - Exposition et décès (UEMOA, homme)

Pour la sous-région UEMOA, le graphique ci-dessus met en évidence le faible nombre d'observations avant 20 ans et après 65 ans pour les hommes.

Figure 5 - Exposition et décès (UEMOA, femme)

À l'image de la sous-région CEMAC, la quantité d'information pour la population féminine de la sous-région UEMOA est plus limitée que celle de la population masculine. Aussi, on observe certains « sauts » dans les nombres de décès.

2. Construction des tables d'expérience

Cette section présente les étapes de construction des tables par sexe pour la CEMAC et l'UEMOA. Ces tables d'expérience sont construites à partir des données de la population assurée. Compte tenu de l'étroitesse des plages d'âges exploitables (35-55 ans pour la CEMAC et 30-55 ans pour l'UEMOA, *cf.* respectivement 2.1.1 et 2.1.3), les constructions sont réalisées *via* des positionnements par rapport à des tables de référence externes. Cette approche présente l'avantage de proposer des taux de mortalité aux âges faibles et aux âges élevés, malgré l'absence de données pour ces âges. Toutefois, en pratique l'extrapolation des taux de décès aux âges faibles et aux âges élevés fait l'objet d'un retraitement particulier afin de tenir compte de l'écart de mortalité mesuré par l'OMS pour ces âges (en l'absence de données d'expérience pour ces âges, la cohérence avec les statistiques de l'OMS est privilégiée, *cf.* 2.2 et 2.3).

2.1. Construction des tables d'expérience

Cette sous-section présente la construction des taux bruts et leur ajustement sur les plages d'âges exploitables dans les données d'expérience.

2.1.1. Construction et ajustement des taux bruts de la table CEMAC H

Ce paragraphe présente la construction et l'ajustement des taux bruts d'expérience pour les hommes de la sous-région CEMAC.

Construction des taux bruts

Les taux bruts de mortalité par âge sont établis à partir de la méthode de Hoem. Selon cette approche, on a (cf. détails dans Planchet et Thérond [2006]) :

$$\hat{q}_x = \frac{d_x}{\sum_i ([\alpha_i; \beta_i] \cap]x; x+1])}$$

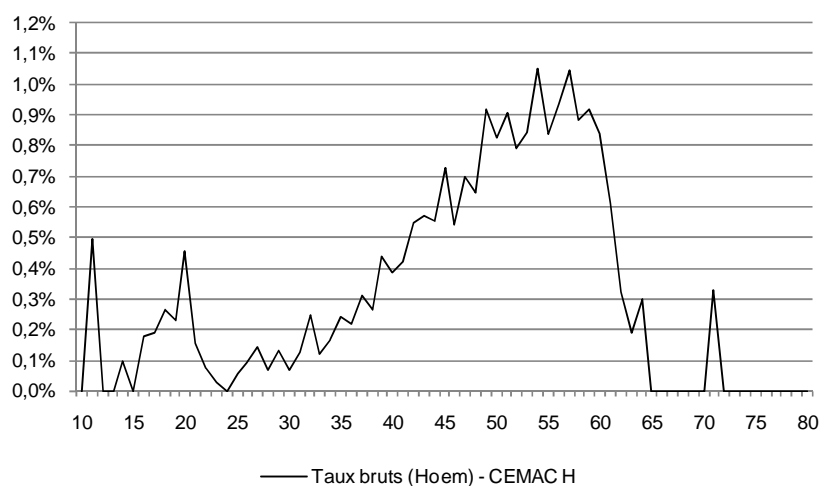
où :

- \hat{q}_x est l'estimateur de la probabilité de décéder à l'âge x ;
- d_x est la réalisation de la variable aléatoire représentant le nombre de décès observés sur $]x; x+1]$;
- $[\alpha_i; \beta_i]$ représente l'intervalle pour lequel l'assuré i est sous observation.

En d'autres termes, cette approche consiste à calculer les taux bruts par âge à partir du rapport entre le nombre de décès et l'exposition au risque de la population par âge. Proche de l'estimateur classique de Kaplan-Meier, cet estimateur permet d'assurer la cohérence avec les expositions au risque qui seront ensuite utilisées pour le calcul des nombres de décès prédits.

Le graphique ci-dessous présente les taux bruts par âge pour les hommes de la sous-région CEMAC.

Figure 6 - Taux de décès bruts (CEMAC, homme)



Ajustement des taux bruts

L'examen des taux de décès bruts ci-dessus montre que la plage d'âge pertinente pour un ajustement est 35-55 ans. En effet, avant 35 ans les taux sont erratiques et au delà de 55 ans les taux diminuent ce qui indique un

vraisemblable problème de données. Compte tenu de la plage d'âge de faible amplitude utilisable, l'approche retenue est celle d'un positionnement par rapport à une référence externe, en l'espèce la table française TH00-02 relative aux hommes de la population générale (la justification du choix de la TH00-02 aux dépens de la TD CIMA est présentée en 2.1.5).

L'estimation des taux de mortalité q_x est contrainte par le fait que l'on doit avoir $q_x \in [0;1]$. Le logit du taux de mortalité permet alors de se ramener à une valeur libre dans $]-\infty; +\infty[$ (en outre, l'utilisation des logits est également justifiée par le fait qu'empiriquement on constate que $\ln\left(\frac{q_x}{1-q_x}\right)$ est approximativement affine au-delà de 25 ans). Aussi, l'ajustement est réalisé à partir d'une régression des logits d'expérience sur les logits de référence (modèle de Brass, présenté en détail dans les chapitres suivants), soit :

$$y_x = \ln\left(\frac{q_x^{ex}}{1-q_x^{ex}}\right) = a \ln\left(\frac{q_x^{ref}}{1-q_x^{ref}}\right) + b + \varepsilon_x.$$

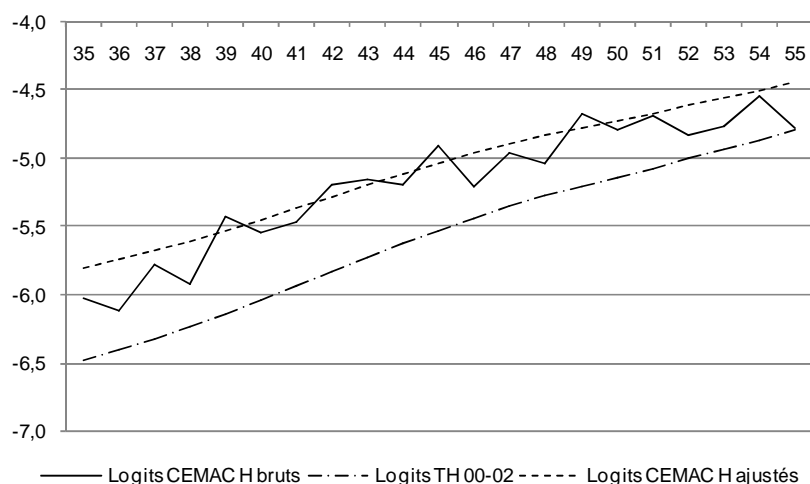
La régression par la méthode des moindres carrés ordinaires conduit aux estimations suivantes $\hat{a} = 0,8236$ et $\hat{b} = -0,5878$. En termes de tests statistiques, il ressort de la statistique de Student que le coefficient de la variable est significativement différent de 0 au seuil de 5 % (probabilité de rejeter à tort l'hypothèse nulle égale à $5,8 \cdot 10^{-11}$), ce qui n'est en revanche pas le cas de la constante (probabilité de rejet à tort l'hypothèse nulle égale à 11,2 %). Dans l'ensemble toutefois, la statistique de Fisher, qui permet de tester la nullité de l'ensemble des coefficients du modèle, justifie le choix du modèle retenu (au seuil de 5 % nous rejetons l'hypothèse nulle car la *p-valeur* associée est $5,8 \cdot 10^{-11}$). De surcroît, le coefficient de détermination R^2 ajusté est égal à 89,5 %, autrement dit le modèle obtenu explique 89,5 % de la variance des logits des taux de décès observés.

Par ailleurs, en pratique, la fonction logistique introduit un biais d'estimation. En effet, entre 0 et $\frac{1}{2}$ la fonction logistique est concave. On déduit alors de l'inégalité de Jensen que dans une zone où les taux de décès sont petits, et si l'on estime les taux de décès par \hat{q}_x supposé sans biais, alors (*cf.* également Planchet et Thérond [2006]) : $E\left[\text{logit}\left(\hat{q}_x\right)\right] \leq \text{logit}\left(q_x\right)$. En d'autres termes, les logits empiriques ainsi obtenus sont biaisés négativement. Pour corriger ce biais, les logits ajustés (négatifs) sont abattus en valeur absolue de 2 % (ce qui conduit de fait à les augmenter). On a ainsi :

$$\hat{y}_x^* = \hat{y}_x \times (1 - 2\%).$$

Le graphique suivant illustre une comparaison des logits d'expérience, des logits de référence et des logits ajustés et abattus.

Figure 7 - Logits (CEMAC, homme)



Ensuite, à l'issue de la modélisation et de l'abattement des logits, les taux de décès sur la plage d'âge exploitable sont calculés à partir de leur logit avec la formule suivante :

$$\hat{q}_x^{CEMAC H} = \frac{e^{\hat{y}_x^*}}{1 + e^{\hat{y}_x^*}}$$

2.1.2. Ajustement des taux bruts de la table CEMAC F

Le volume de données de la population féminine est limité et l'approche retenue pour déterminer les taux ajustés des femmes de la CEMAC est différente de celle retenue pour les hommes.

En effet, les taux ajustés de la table CEMAC F sur la plage d'âge exploitable sont obtenus par un abattement des taux CEMAC H. Le niveau de cet abattement est déduit du modèle à risque proportionnel de Cox (*cf.* par exemple Hill et al. [1990] ou encore Klein et Moeschberger [2005] dans le cadre d'une utilisation en présence de données tronquées et censurées). Cette méthode permet de positionner la mortalité d'un groupe spécifique (en l'occurrence celle des femmes assurées de la CEMAC) par rapport à une mortalité de référence (en l'occurrence celle des hommes assurées de la CEMAC). Elle permet ainsi de prendre en compte les données d'expérience de la population féminine, malgré les limites liées au faible volume des données. On a alors (*cf.* hypothèses sous-jacents dans le chapitre sur l'hétérogénéité) :

$$q_x^{CEMAC F} = 1 - \left(1 - q_x^{CEMAC H}\right)^{\exp(\theta)},$$

où θ désigne le paramètre à estimer.

En pratique, il ressort que l'estimation du paramètre est significative au seuil de 5 % (que le test de significativité soit réalisé avec la statistique du

rapport des vraisemblances ou avec la statistique de Wald, cf. chapitre sur l'hétérogénéité et Klein et Moeschberger [2005] pour une présentation de ces tests) et donne $\hat{\theta} = -0,657$. En d'autres termes, le coefficient d'abattement est égal à $\exp(\hat{\theta}) = 0,52$.

Une comparaison des taux de décès bruts et des taux de décès ajustés pour les femmes de la sous-région CEMAC est présentée sur la Figure 19.

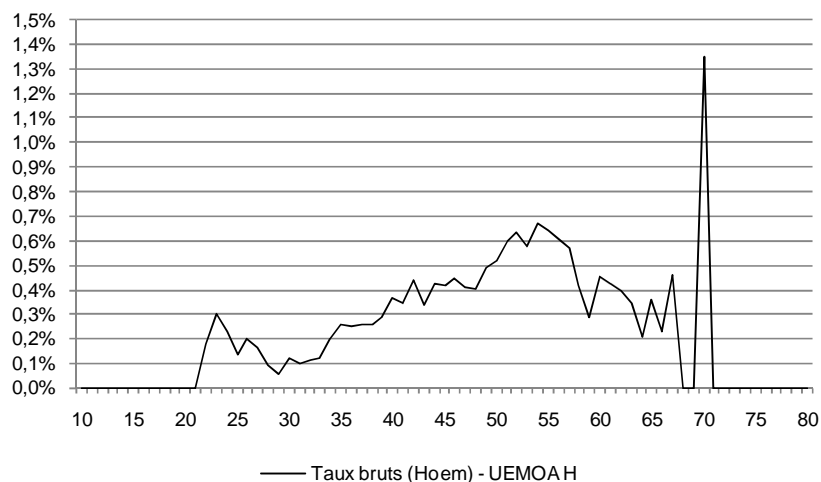
2.1.3. Construction et ajustement des taux bruts de la table UEMOA H

Ce paragraphe présente la construction et l'ajustement des taux bruts d'expérience établie pour les hommes de la sous-région UEMOA. L'approche est comparable à celle retenue pour la population masculine de la sous-région CEMAC et le détail de la méthode retenue n'est donc pas repris ici (on pourra toutefois se référer au paragraphe 2.1.1 pour plus de précisions sur ce point).

Construction des taux bruts

Le graphique ci-dessous présente les taux bruts par âge pour les hommes de la sous-région UEMOA (établis à partir de la méthode de Hoem).

Figure 8 - Taux de décès bruts (UEMOA, homme)



Ajustement des taux bruts

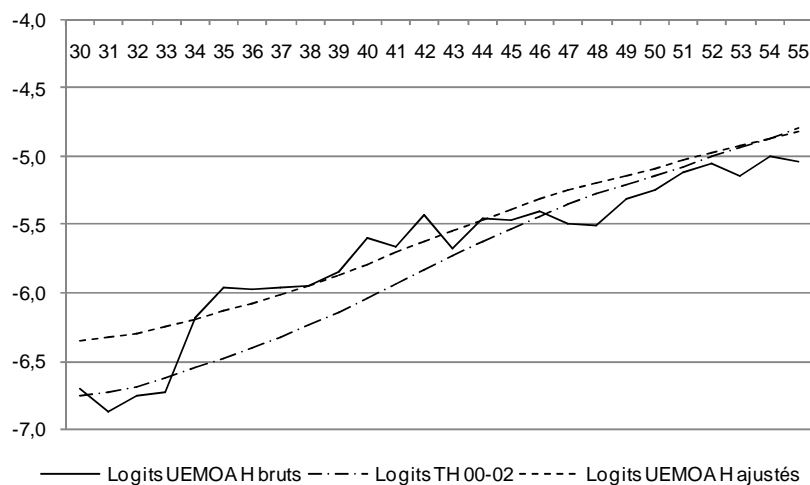
L'examen des taux de décès bruts montre que la plage d'âge pertinente pour un ajustement est 30-55 ans. En effet, avant 30 ans les taux présentent une évolution injustifiable et au delà de 55 ans les taux diminuent ce qui indique un vraisemblable problème de données. Compte tenu de la plage d'âge de faible amplitude utilisable et à l'image des travaux relatifs à la sous-région CEMAC, l'approche retenue est celle d'un positionnement par rapport à une référence

externe, en l'espèce la table française TH00-02. L'ajustement est réalisé à partir du modèle de Brass (régression des logits d'expérience sur les logits de référence, cf. paragraphe 2.1.1).

La régression par la méthode des moindres carrés ordinaires conduit aux estimations suivantes $\hat{a} = 0,7989$ et $\hat{b} = -1,0821$. En termes de tests statistiques, il ressort de la statistique de Student que le coefficient de la variable est significativement différent de 0 au seuil de 5 % (probabilité de rejeter à tort l'hypothèse nulle égale à $4,0 \cdot 10^{-12}$). De même, la statistique de Student illustre que la constante est significativement différente de 0 au seuil de 5 % (probabilité de rejeter à tort l'hypothèse nulle égale à 0,7 %). Dans l'ensemble, la statistique de Fisher justifie le choix du modèle retenu (au seuil de 5 % nous rejetons l'hypothèse nulle car la *p-valeur* associée est $4,0 \cdot 10^{-12}$). De surcroît, le coefficient de détermination R^2 ajusté est égal à 86,5 %.

À l'image des travaux pour la CEMAC, les logits ajustés (négatifs) sont ensuite abattus en valeur absolue de 2 %. Le graphique suivant illustre une comparaison des logits d'expérience, des logits de référence et des logits ajustés et abattus.

Figure 9 - Logits (UEMOA, homme)



Les taux de décès pour la plage d'âge exploitable sont ensuite déduits de ces valeurs à partir de la formule présentée à la fin du paragraphe 2.1.1.

2.1.4. Ajustement des taux bruts de la table UEMOA F

Cette partie présente l'ajustement des taux de la table d'expérience relative à la population féminine de la sous-région UEMOA. L'approche est comparable à celle retenue pour la population féminine de la sous-région CEMAC et le détail de la méthode retenue n'est donc pas repris ici (on pourra toutefois se référer au paragraphe 2.1.2 pour plus de précisions sur ce point).

Ainsi, à l'image des travaux sur la sous-région CEMAC (cf. 2.1.2), les taux de la table UEMOA F sur la plage d'âge exploitable sont obtenus par un

abattement des taux de la table d'expérience UEMOA H. Ici, le coefficient d'abattement est égal à $\exp(\hat{\theta}) = 0,53$ (estimation du paramètre significative au seuil de 5 %, que le test de significativité soit réalisé avec la statistique du rapport des vraisemblances ou avec la statistique de Wald).

Une comparaison des taux de décès bruts et des taux de décès ajustés pour les femmes de la sous-région UEMOA est présentée sur la Figure 17.

2.1.5. Choix de la référence TH00-02

Dans la construction des tables de mortalité, la table de référence retenue est la table TH00-02. Cette référence a été retenue compte tenu de la bonne qualité des ajustements qu'elle permet d'obtenir. D'autres choix auraient pu être effectués. À titre d'information, nous obtenons le comparatif suivant avec l'utilisation de la table TD CIMA comme référence.

Tableau 5 – Statistiques par segment

R2 ajusté	TH00-02	TD CIMA
CEMAC H	89,5%	86,2%
UEMOA H	86,5%	87,1%

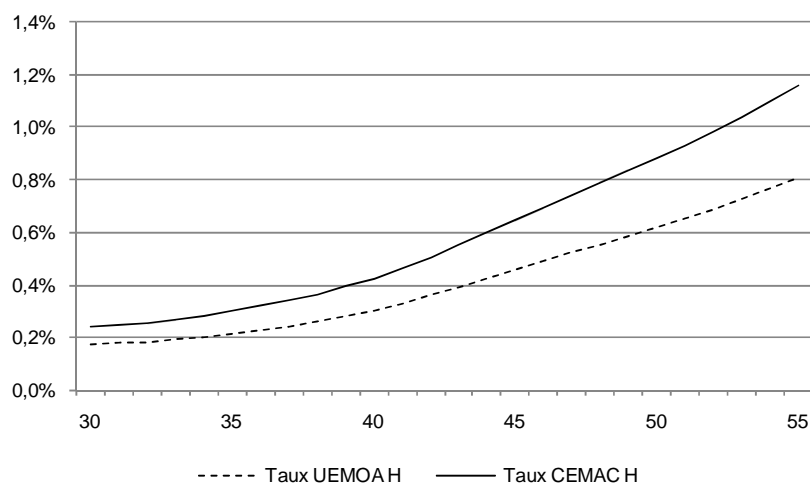
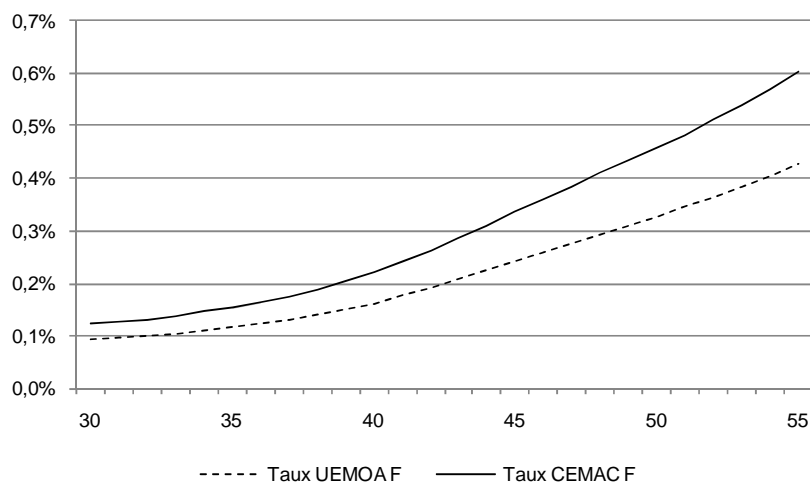
Pour la zone CEMAC, la qualité de l'ajustement est légèrement meilleure avec le choix de la TH00-02 et pour l'UEMOA les deux choix sont comparables. Cela conduit à préférer la référence TH00-02 pour bénéficier de cette bonne qualité d'ajustement.

2.2. Extrapolation des taux ajustés

À l'issue des ajustements des taux sur les plages d'âges exploitables, il convient d'extrapoler ces taux à gauche (pour les âges les plus faibles) et à droite (pour les âges les plus élevés), afin de disposer de tables complètes pour les sous-régions UEMOA et CEMAC.

2.2.1. Illustration des taux ajustés UEMOA et CEMAC

Les graphiques suivants présentent les taux de mortalité ajustés sur les plages d'âges exploitables, c'est-à-dire les plages d'âges retenues pour les ajustements statistiques.

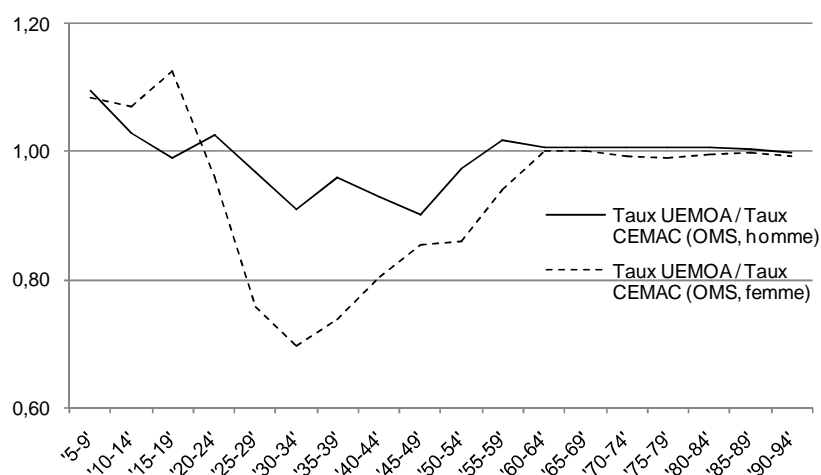
Figure 10 - Taux de décès CEMAC et UEMOA pour les 30-55 ans (H)**Figure 11 - Taux de décès CEMAC et UEMOA pour les 30-55 ans (F)**

Il ressort de ces graphiques que sur tous les âges retenus pour les ajustements, les taux de mortalité CEMAC sont supérieurs à ceux de l'UEMOA, pour les hommes et pour les femmes. Aussi, une extrapolation « en l'état » (c'est-à-dire à partir des modèles retenus pour l'ajustement) des taux de mortalité à gauche (pour les âges faibles) et à droite (pour les âges élevés) conduirait à retenir des taux CEMAC supérieurs aux taux UEMOA pour les âges faibles et pour les âges élevés. Les statistiques de l'OMS présentées ci-dessous en 2.2.2 montrent que ces résultats ne seraient alors pas cohérents avec les données de la population générale.

2.2.2. Comparaison des taux UEMOA et CEMAC issus de l'OMS

Le graphique suivant présente les rapports entre les taux de mortalité UEMOA et CEMAC par tranches d'âge, établis à partir des données de l'OMS (pour mémoire, on rappelle ici que les données OMS présentent également certaines limites, cf. 1.4.1).

Figure 12 - Rapports entre les taux OMS UEMOA et CEMAC (H et F)



On retient de l'illustration ci-dessus que :

- pour les âges faibles (assimilés ici aux moins de 20 ans), les taux de mortalité UEMOA sont supérieurs ou égaux à ceux de la CEMAC, pour les hommes et pour les femmes ;
- pour les âges élevés (assimilés ici aux plus de 65 ans), les taux de mortalité CEMAC sont égaux à ceux de l'UEMOA, pour les hommes et pour les femmes.

Ces observations diffèrent de celles attendues des extrapolations « en l'état » des taux de mortalité ajustés dans les sous-régions CEMAC et UEMOA (cf. paragraphe 2.2.1 ci-dessus).

Par ailleurs, pour mémoire aucune donnée d'expérience exploitable pour les âges faibles et les âges élevés n'est disponible. Aussi, pour ces âges la seule référence disponible en termes de cohérence entre les taux UEMOA et les taux CEMAC est celle issue des statistiques de l'OMS. En pratique, les extrapolations aux âges faibles et aux âges élevés sont ainsi réalisées sur les taux UEMOA et CEMAC en tenant compte des écarts OMS entre les deux sous-régions : il s'agit ainsi d'extrapolations retraitées.

2.2.3. Extrapolation des tables UEMOA et CEMAC

Ce paragraphe présente le retraitement des extrapolations retenu pour les tables UEMOA et CEMAC, pour les âges faibles et les âges élevés.

Extrapolation des tables UEMOA et CEMAC aux âges faibles

Pour les âges faibles, l'extrapolation des taux ajustés sans retraitement conduirait à des taux CEMAC supérieurs à ceux de l'UEMOA, alors que les statistiques OMS illustrent que les taux UEMOA sont supérieurs ou égaux à ceux de la CEMAC pour la population générale.

En l'absence d'informations exploitables sur les taux de décès d'expérience à ces âges, les taux des deux tables sont alors considérés égaux entre 0 et 20 ans. En pratique, l'extrapolation à gauche des taux de la CEMAC est minorée afin d'obtenir des taux de mortalité égaux entre les deux sous-régions entre 0 et 20 ans. Cette approche est privilégiée à une majoration des taux UEMOA entre 0 et 20 ans afin de ne pas observer de diminution significative des taux UEMOA entre 20 et 29 ans.

Ainsi, pour les hommes (respectivement pour les femmes), entre 0 et 20 ans les taux CEMAC H (respectivement CEMAC F) sont supposés égaux aux taux UEMOA H (respectivement UEMOA F) établis à partir de l'extrapolation issue du modèle ajusté. Entre 20 et 29 ans, les taux CEMAC sont ensuite majorés à un rythme régulier afin d'atteindre les taux issus des ajustements à compter de 30 ans.

Extrapolation des tables UEMOA et CEMAC aux âges élevés

Pour les âges élevés, l'extrapolation des taux ajustés sans retraitement conduirait à des taux CEMAC supérieurs à ceux de l'UEMOA, alors que les statistiques OMS illustrent que les taux UEMOA sont égaux à ceux de la CEMAC pour la population générale.

En l'absence d'informations exploitables sur les taux de décès d'expérience à ces âges, les taux des deux tables seront alors considérés égaux au-delà de 65 ans. En pratique, l'extrapolation à droite des taux de l'UEMOA est majorée afin d'obtenir des taux de mortalité égaux entre les deux sous-régions au-delà de 65 ans. Cette approche est privilégiée à une minoration des taux CEMAC car dans le cadre de garanties décès (qui représentent près 90 % de l'exposition au risque totale), elle est facilement justifiable.

Ainsi, pour les hommes (respectivement pour les femmes), entre 66 et l'âge de fermeture de la table (*cf.* sous-section suivante) les taux UEMOA H (respectivement UEMOA F) sont supposés égaux aux taux CEMAC H (respectivement CEMAC F) établis à partir de l'extrapolation issue du modèle ajusté. Au préalable, entre 56 et 65 ans, les taux UEMOA sont majorés à un rythme régulier afin de devenir égaux aux taux CEMAC à compter de 66 ans.

2.3. Fermeture des tables

Dans une perspective de cohérence avec les indicateurs pour la population générale (issus des données OMS) des zones concernées, les taux de décès au-delà de 90 ans sont modifiés et l'âge limite des tables est ramené à 105 ans. Compte tenu de l'absence totale d'observations pour ces âges, les taux fournis sont largement indicatifs et n'ont pas l'ambition de décrire fidèlement le niveau de mortalité d'éventuels assurés. En tout état de cause, plusieurs études montrent qu'aux âges élevés, le choix de la méthode de fermeture a un impact limité sur

le niveau de la provision (cf. par exemple Quashie et Denuit [2005] et Planchet et Thérond [2006]).

En pratique, les tables sont fermées à 105 avec un modèle de Gompertz (taux de décès exponentiel) entre 91 et 104 ans pour les taux de la table CEMAC H uniquement (étant donné qu'aux grands âges, les taux des trois autres tables sont directement déduits de ceux de la table CEMAC H).

Pour mémoire, le modèle de Gompertz modélise le taux de décès instantané comme suit :

$$\mu_x = b \times c^x,$$

et les paramètres retenus ici sont : $\hat{b} = 4,0.10^{-9}$ et $\hat{c} = 1,212$. Un test de cohérence de cette fermeture avec celle des tables OMS est présenté en 3.2.2.

2.4. Présentation des tables UEMOA et CEMAC

Les graphiques ci-dessous présentent les taux de mortalité des quatre tables d'expérience finales.

2.4.1. Présentation des tables UEMOA et CEMAC

La Figure 13 et la Figure 14 présentent les tables de mortalité construites pour les hommes et pour les femmes.

Figure 13 - Présentation des tables UEMOA H et CEMAC H

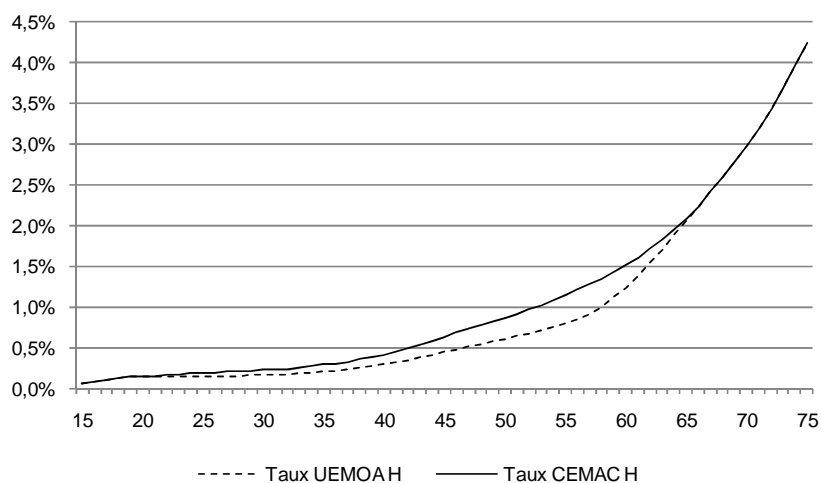
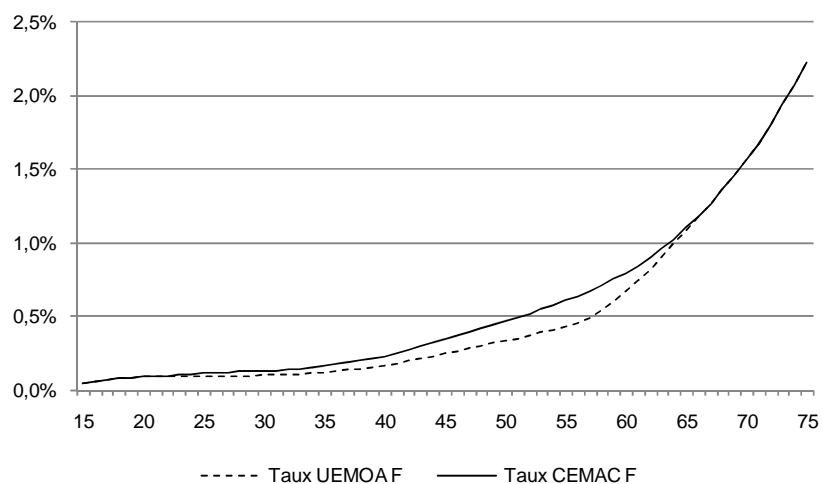
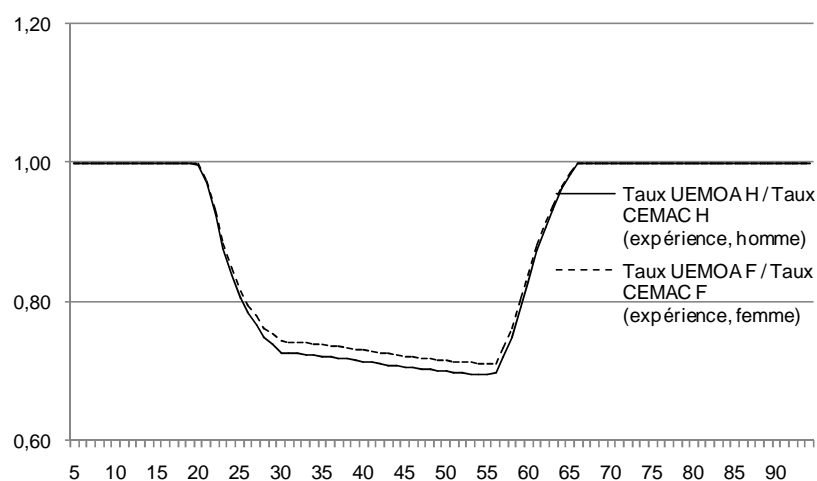


Figure 14 - Présentation des tables UEMOA F et CEMAC F

En complément, le graphique suivant présente les rapports entre les taux de mortalité UEMOA et CEMAC, issus des tables d'expérience construites (pour mémoire, ces graphiques sont à comparer au graphique de la Figure 12).

Figure 15 - Rapport des taux d'expérience UEMOA et CEMAC (H et F)

On note que l'allure des écarts observés avec les données d'expérience est relativement cohérente avec celle des écarts issus des données de l'OMS (représentant la population générale).

2.4.2. Rapprochements des tables UEMOA et CEMAC avec les taux bruts

En complément, les graphiques suivants présentent pour chaque segment les taux ajustés des tables d'expérience construites ainsi que les taux bruts de mortalité associés.

Figure 16 - Comparaison des taux UEMOA H avec les taux bruts

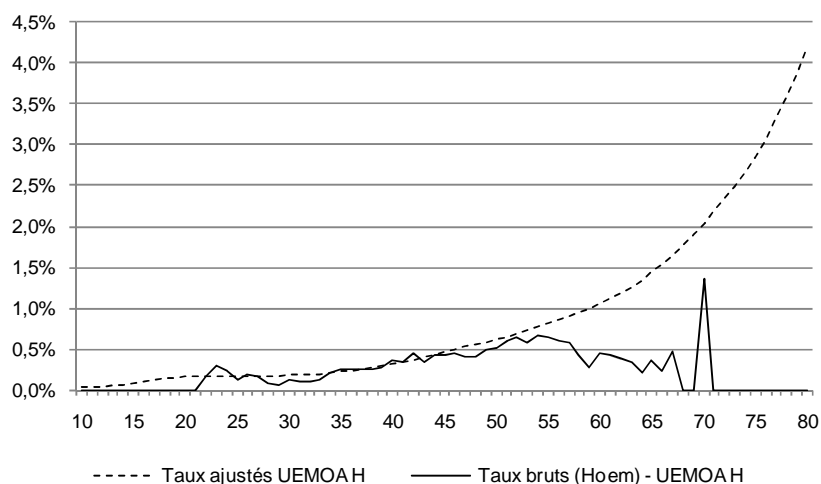


Figure 17 - Comparaison des taux UEMOA F avec les taux bruts

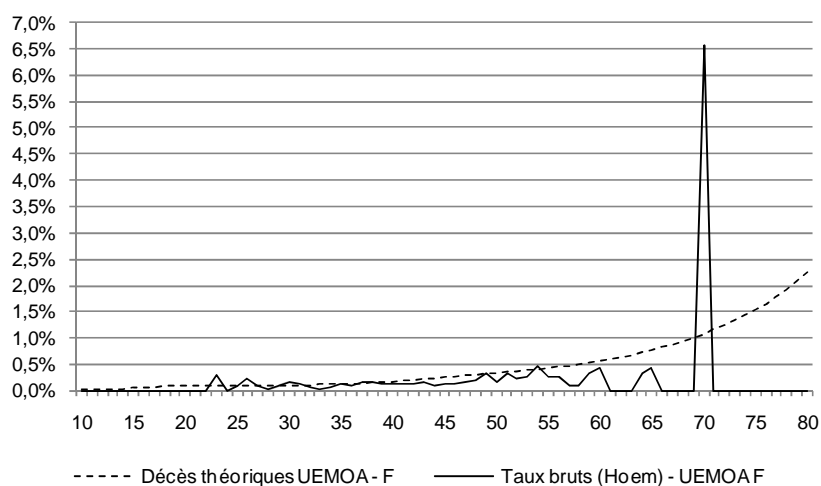
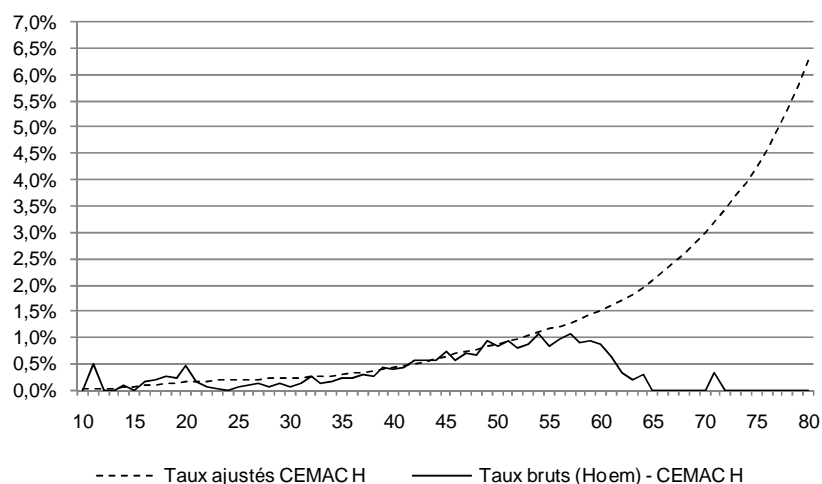
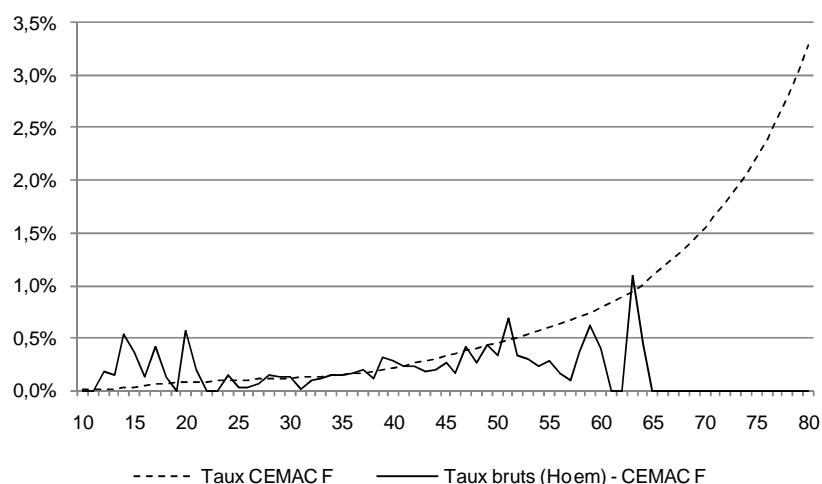


Figure 18 - Comparaison des taux CEMAC H avec les taux bruts**Figure 19 - Comparaison des taux CEMAC F avec les taux bruts**

De manière générale, il ressort qu'entre 30 et 55 ans pour l'UEMOA et entre 35 et 55 ans pour la CEMAC, les taux de décès issus des tables d'expérience construites sont cohérents avec les taux bruts issus des données.

3. Contrôle des tables d'expérience construites et comparaisons

Cette section présente des statistiques permettant d'apprécier la pertinence des tables de mortalité d'expérience construites. Aussi, cette section présente des statistiques facilitant le contrôle de la cohérence entre les tables construites et les données d'expérience d'une part, et entre les tables construites et les données de l'OMS (population générale) d'autre part. En complément, une

comparaison entre les tables construites et les tables réglementaires est également réalisée.

3.1. Cohérence des tables construites avec les observations

Cette sous-section s'attache à vérifier la cohérence entre le nombre de décès prédits à partir des tables construites et le nombre de décès observés dans les données d'expérience.

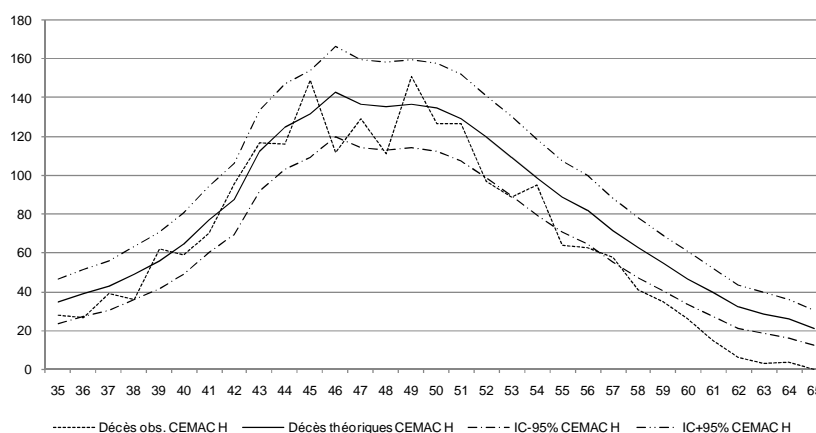
3.1.1. Contrôle des tables d'expérience CEMAC

La cohérence entre la réalisation et la prédiction des décès est réalisé dans ce paragraphe pour les hommes, puis pour les femmes de la CEMAC.

Comparaison réalisation / prédiction pour la table CEMAC H

Le graphique suivant présente la comparaison des décès par âge pour les hommes de la CEMAC (les intervalles de confiance à 95 % associés sont obtenus en effectuant l'approximation usuelle d'une loi binomiale par une loi normale).

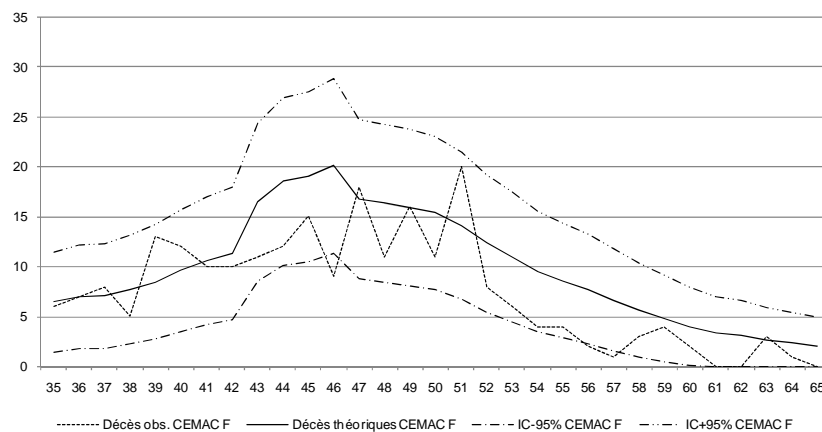
Figure 20 - Comparaison réalisation / prédiction décès (CEMAC, homme)



Le graphique permet de constater que pour la population masculine de la sous-région CEMAC, les décès observés sont comparables aux décès théoriques issus de la table d'expérience construite, hormis au-delà de 55 ans où l'écart devient anormalement important, probablement du fait d'une erreur dans le décompte des décès dans les données transmises.

Comparaison réalisation / prédiction pour la table CEMAC F

Le graphique suivant présente la comparaison pour les femmes de la CEMAC.

Figure 21 - Comparaison réalisation / prédiction décès (CEMAC, femme)

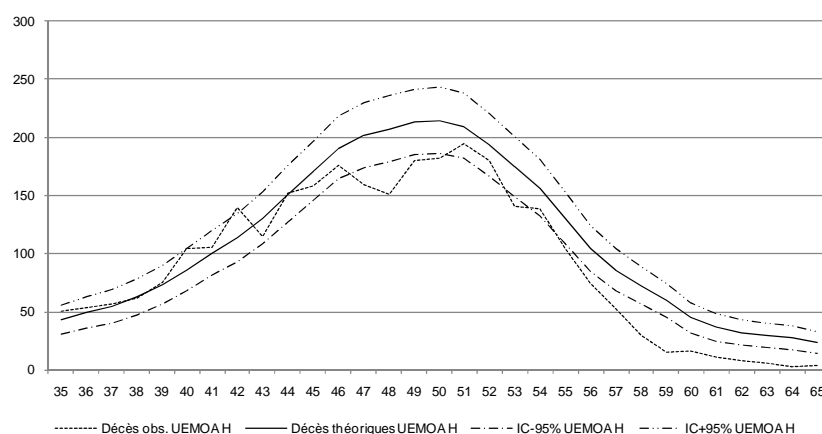
L'interprétation de ce graphique pour la population féminine est toutefois limitée, compte tenu de la faible exposition au risque de cette population : on observe un nombre de décès par âge réduit et un intervalle de confiance important.

3.1.2. Contrôle des tables d'expérience UEMOA

Dans ce paragraphe, la cohérence entre la réalisation et la prédiction des décès est réalisée pour les hommes, puis pour les femmes de l'UEMOA.

Comparaison réalisation / prédiction pour la table UEMOA H

Le graphique suivant présente la comparaison des décès par âge pour les hommes de la sous-région UEMOA.

Figure 22 - Comparaison réalisation / prédiction décès (UEMOA, homme)

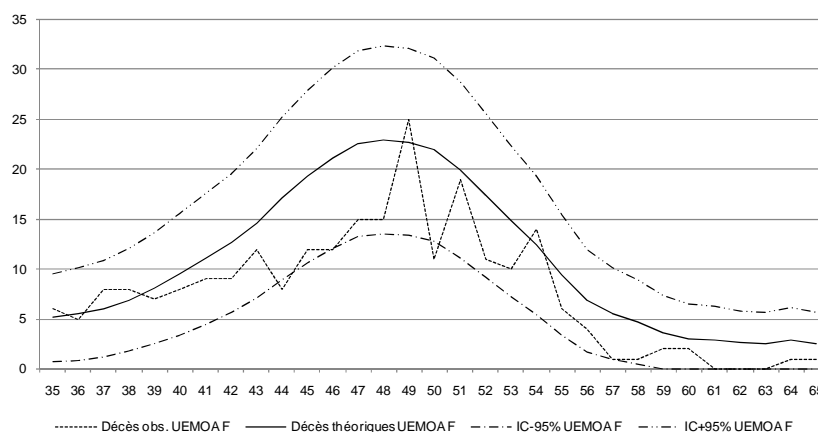
À l'image de ce qui avait été observé pour les hommes de la CEMAC, le graphique ci-dessous pour l'UEMOA permet de constater que les décès observés sont comparables aux décès théoriques issus de la table d'expérience

construite, hormis au-delà de 55 ans où l'écart devient anormalement important, probablement du fait d'une erreur dans le décompte des décès dans les données transmises (il apparaît également un écart important autour de 48 ans, vraisemblablement au titre des fluctuations d'échantillonnage).

Comparaison réalisation / prédiction pour la table UEMOA F

Le graphique suivant présente la comparaison pour la population féminine de l'UEMOA.

Figure 23 - Comparaison réalisation / prédiction décès (UEMOA, femme)



L'interprétation de ce graphique pour la population féminine de la sous-région UEMOA est toutefois limitée par la faible exposition au risque de cette population : on observe un nombre de décès par âge réduit et un intervalle de confiance important.

3.2. Cohérence des tables construites avec les tables OMS

L'objet de cette sous-section est de confronter les taux des tables construites (population assurée) aux taux des tables issus des statistiques de l'OMS (population générale). Cette confrontation est réalisée au global d'une part, et uniquement pour les plus de 54 ans (fermeture) d'autres part.

3.2.1. Contrôle global des tables construites avec les tables OMS

Une population assurée correspond à une population relativement privilégiée par rapport à la population générale, ce qui s'illustre par des taux de mortalité plus faibles pour la population assurée que pour la population générale.

C'est vrai en France, où l'on observe généralement un abattement compris entre 25 % et 60 %. À titre d'exemple, il apparaît ainsi qu'entre 30 et 70 ans, l'abattement entre les tables du moment TH/TF00-02 (basées sur la population générale française) et les tables « du moment » reconstituées en 2002 à partir

des tables prospectives TGH/TGF05 (basées sur une population d'assurés rentiers français) est compris entre 30 % et 45 % pour les hommes et entre 50 % et 70 % pour les femmes (*cf.* Planchet [2006]), sachant que les rentiers ont des taux de mortalité plus faibles que les autres assurés des contrats d'assurance vie.

Dans ce paragraphe, on montre que c'est également vrai au sein de la zone CIMA. À cet effet, on compare directement le nombre de décès théoriques obtenu à partir des tables construites (population assurée) à celui obtenu à partir des tables OMS (population générale). On note que dans les deux cas, l'exposition au risque retenue est celle de la population assurée du segment correspondant (ainsi par exemple, pour le segment UEMOA H, dans les deux cas les décès théoriques sont calculés à partir de l'exposition des hommes assurés de l'UEMOA). Les résultats sont présentés dans le tableau ci-dessous.

Tableau 6 – Décès théoriques (tables d'expérience et tables OMS)

	Décès théoriques				Total
	UEMOA H	UEMOA F	CEMAC H	CEMAC F	
Tables OMS	13 515	1 945	7 904	1 878	25 241
Tables d'expérience	4 473	465	2 881	372	8 190
Ecart	-67%	-76%	-64%	-80%	-68%

Il ressort du tableau ci-dessus que la mortalité de la population assurée est effectivement plus faible que celle de la population générale.

En particulier, au sein de la zone CIMA il apparaît qu'au global les taux obtenus correspondent à un abattement de près de 70 % par rapport aux taux de la population générale issus des statistiques de l'OMS. Cet abattement élevé est cohérent avec une population assurée fortement privilégiée par rapport à la population générale (pour mémoire, en France l'abattement est en moyenne de l'ordre de 40 % sur l'ensemble des portefeuilles, et s'élèvent à près de 70 % sur des portefeuilles d'assurés « haut de gamme »), compte tenu du faible taux de pénétration du marché de l'assurance vie dans la région (égal à moins de 0,3 % dans la zone selon Kone [2007], contre plus de 7 % en France).

3.2.2. Contrôle de la fermeture des tables construites avec les tables OMS

Comme indiqué en 2.3, les tables construites sont fermées à 105 ans avec un modèle de Gompertz entre 91 et 104 ans.

En termes de comparaison avec les taux de l'OMS, on obtient les résultats présentés dans les tableaux suivants pour les taux de décès aux grands âges.

Tableau 7 – Fermetures tables d'expérience et tables OMS - UEMOA

Comparaison des taux de décès quinquennaux (UEMOA)						
Tranche d'âges	UEMOA H (OMS)	UEMOA H (expérience)	Ecart UEMOA H	UEMOA F (OMS)	UEMOA F (expérience)	Ecart UEMOA F
'55-59'	13%	4%	-73%	9%	2%	-80%
'60-64'	17%	6%	-65%	12%	3%	-74%
'65-69'	23%	9%	-60%	18%	5%	-73%
'70-74'	32%	13%	-61%	27%	7%	-75%
'75-79'	44%	18%	-60%	38%	10%	-75%
'80-84'	59%	26%	-56%	53%	14%	-73%
'85-89'	75%	37%	-51%	69%	21%	-69%
'90-94'	84%	55%	-34%	79%	34%	-57%
'95-99'	88%	87%	-1%	86%	66%	-23%
100-105'	100%	100%	0%	100%	94%	-6%

Tableau 8 – Fermetures tables d'expérience et tables OMS - CEMAC

Comparaison des taux de décès quinquennaux (CEMAC)						
Tranche d'âges	CEMAC H (OMS)	CEMAC H (expérience)	Ecart CEMAC H	CEMAC F (OMS)	CEMAC F (expérience)	Ecart CEMAC F
'55-59'	13%	5%	-62%	10%	3%	-74%
'60-64'	17%	7%	-61%	12%	3%	-71%
'65-69'	23%	9%	-60%	18%	5%	-73%
'70-74'	32%	13%	-60%	27%	7%	-75%
'75-79'	44%	18%	-60%	39%	10%	-75%
'80-84'	59%	26%	-56%	53%	14%	-73%
'85-89'	75%	37%	-50%	70%	21%	-69%
'90-94'	84%	55%	-34%	80%	34%	-58%
'95-99'	88%	87%	-1%	85%	66%	-23%
100-105'	100%	100%	0%	100%	94%	-6%

Malgré l'absence de données au-delà de 55 ans, il ressort à ces âges que les niveaux d'abattements entre les taux de mortalité issus des tables OMS (population générale) et ceux issus des tables d'expérience (population assurée) sont cohérents avec les niveaux observés au global (*cf.* Tableau 6 du paragraphe précédent), c'est-à-dire observés sur la population pour laquelle nous disposons d'une exposition au risque exploitable.

3.3. Comparaison des tables construites avec les tables TD/TV CIMA

Aujourd'hui, au sein de la zone CIMA, la table TD CIMA est la table réglementaire pour calculer le tarif des assurances en cas de décès et la table TV CIMA est la table réglementaire pour calculer le tarif des assurances en cas de vie. Pour mémoire, on rappelle ici la table TD CIMA correspond à la table française PM 60-64 construite à partir de la population générale française masculine observée entre 1960 et 1964. De même, la table TV CIMA correspond à la table française PF 60-64 construite à partir de la population générale française féminine observée entre 1960 et 1964. Ces tables réglementaires TD CIMA et TV CIMA sont des tables unisexes.

Sur ce même principe, on considère ici que les tables « hommes » sont envisagées pour les risques en cas de décès et que les tables « femmes » sont envisagées pour les risques en cas de vie, quel que soit le sexe de l'assuré.

Aussi, les comparaisons des tables construites avec les tables réglementaires ne sont pas réalisées par sexe, mais en considérant d'une part les tables envisagées au titre du risque décès (soit les tables construites « hommes » et la TD CIMA) et d'autre part les tables envisagées au titre du risque vie (soit les tables construites « femmes » et la TV CIMA).

Par ailleurs, en termes d'outils, les comparaisons entre tables sont réalisées à partir de l'indice comparatif de mortalité (SMR, *Standardized mortality ratio*), qui correspond à une méthode de standardisation des taux de décès (à partir d'une population de référence) et qui permet d'obtenir des coefficients représentant les écarts relatifs de niveau de mortalité entre deux tables. Cet indice, ou SMR, est égal au rapport entre le nombre de décès théoriques à partir de deux tables calculé à partir d'une même population de référence (en l'occurrence ici la population assurée totale de la zone CIMA). Il est généralement exprimé sur une base 100.

3.3.1. Comparaison des tables construites avec les tables réglementaires (risques décès)

Dans un premier temps, la comparaison est effectuée pour les tables relatives au risque en cas de décès. En retenant la population assurée totale de la CIMA comme population de référence pour le calcul des SMR, on obtient les résultats présentés dans le tableau suivant (le numérateur du rapport est en colonne).

Tableau 9 – SMR des tables CEMAC H, UEMOA H et TD CIMA

Tables décès	CEMAC H	UEMOA H	TD CIMA
CEMAC H	100	75	101
UEMOA H		100	135
TD CIMA			100

Il ressort que les décès issus de la table réglementaire pour les garanties en cas de décès (TD CIMA) majorent de 1 % les décès issus de la table CEMAC H et de 35 % ceux issus de la table UEMOA H. En outre, on observe que les décès issus de la table UEMOA H minorent de 25 % ceux de la table CEMAC H.

3.3.2. Comparaison des tables construites avec les tables réglementaires (risques vie)

Dans un second temps, la comparaison est effectuée pour les tables relatives au risque en cas de vie. En retenant également la population assurée totale de la CIMA comme population de référence pour le calcul des SMR, on obtient les résultats présentés dans le tableau suivant.

Tableau 10 – SMR des tables CEMAC F, UEMOA F et TV CIMA

Tables vie	CEMAC F	UEMOA F	TV CIMA
CEMAC F	100	77	92
UEMOA F		100	119
TV CIMA			100

Il ressort que les décès issus de la table réglementaire pour les garanties en cas de vie (TV CIMA) minorent de 8 % les décès issus de la table CEMAC F et majorent de 19 % ceux issus de la table UEMOA F. En outre, on observe que les décès issus de la table UEMOA F minorent de 23 % ceux de la table CEMAC F.

4. Synthèse et conclusion (et présentation de défis techniques complémentaires sur la mortalité)

Les données transmises par les organismes assureurs impliqués dans l'étude ont permis de construire un jeu de 4 tables de mortalité, croisement du sexe (H / F) et de la sous-région (CEMAC / UEMOA) selon la démarche suivante⁵ :

- ajustement des taux de mortalité d'expérience pour les hommes de la zone CEMAC (à partir des données de la plage 35-55 ans) et de l'UEMOA (à partir des données de la plage 30-55 ans) par une technique de positionnement par rapport à une table de mortalité externe ;
- déduction des taux de mortalité d'expérience des femmes pour les mêmes sous-régions par un positionnement dans le cadre du modèle multiplicatif de Cox ;
- construction des tables des sous-régions UEMOA et CEMAC en appliquant une extrapolation des taux de mortalité ajustés en tenant compte, pour les âges faibles et les âges élevés, des écarts de mortalité entre la CEMAC et l'UEMOA mesurés par les statistiques de l'OMS ;
- fermeture des tables à 105 ans avec un modèle de Gompertz entre 91 et 104 ans, dans une perspective de cohérence avec les indicateurs de l'OMS.

On observe sur les données transmises un comportement anormal des décès après 55 ans, dont le nombre baisse de manière inexplicable. Il peut s'agir d'un mécanisme lié au fonctionnement des contrats (suppression de la garantie décès) ou d'une anomalie dans les données. Dans le présent travail, cette plage d'âge n'a donc pas été exploitée.

Au global toutefois, l'approche retenue permet d'obtenir un jeu de tables finalement assez proche des tables réglementaires actuelles.

⁵ En pratique toutefois, la CIMA envisage de retenir une paire de table unique CIMA H et CIMA F (construites en pondérant les tables par sous-région par les expositions au risque) afin de faciliter l'implémentation des tables d'expérience par les organismes assureurs, cf. CIMA [2011].

Par ailleurs, les faibles effectifs disponibles pour la population féminine donnent lieu à une incertitude plus grande que pour les hommes quant au niveau issu de l'expérience. Toutefois, le niveau du coefficient d'abattement de cette population par rapport à la mortalité masculine est classique et raisonnable (de l'ordre de 50 %).

On peut également retenir que le niveau des abattements à appliquer à la population générale (dont la mortalité est supposée mesurée par les études de l'OMS) pour se recalculer sur la mortalité de la population assurée est cohérent avec ce que l'on observe par exemple en France sur des populations d'assurés sur des contrats « haut de gamme », donc des populations significativement privilégiées d'un point de vue socioculturel.

Dans l'ensemble, on peut donc considérer que, malgré des données de qualité améliorable, les références de mortalité obtenues ici conduisent à des résultats explicables et relativement cohérents avec les références dont on peut disposer.

À l'issue de ce travail, on dispose ainsi de références fiables pour apprécier le niveau de la mortalité de la population assurée, par sexe (hommes / femmes) et par sous-région (UEMOA / CEMAC). Deux pistes d'analyses complémentaires apparaissent toutefois utiles dans le cadre d'une modélisation plus fine des risques associés à la durée de vie humaine dans le contexte de la zone CIMA/FANAF.

La première consiste à tenir compte des différences potentielles de mortalité entre pays de la zone. En effet, dans le présent travail, la prise en compte de l'hétérogénéité s'est limitée à une ventilation par sous-région, afin de limiter les risques potentiels liés à la réduction des échantillons de données. Dans le cadre d'une analyse complémentaire, il semble toutefois souhaitable de tenir compte d'une ventilation par pays et de disposer d'outils adaptés à cet effet.

La seconde piste d'analyse complémentaire consiste à tenir compte des évolutions de la durée de vie humaine au fil des années. En effet, dans le présent travail, seule la mortalité actuelle est modélisée, et dans un contexte où la CIMA souhaite favoriser l'instauration de régimes de retraite au sein du marché, une réflexion s'impose sur la modélisation de la mortalité future. Dans le cadre d'une analyse complémentaire, il serait donc également souhaitable de disposer d'outils adaptés pour modéliser la mortalité prospective.

Les travaux des parties II et III ci-après s'attachent à fournir des réponses à chacun de ces défis techniques complémentaires, malgré les limites actuelles dans les données d'expérience au sein de la région.

Bibliographie

- CIMA [2011] « Note sur la confection des nouvelles tables de mortalité », *CIMA*.
- Hill C., Com-Nougé C., Kramar A., Moreau T., O'Quigley J., Senoussi R., Chastang C. [1990] « Analyse statistique des données de survie », *Inserm, Médecine – Sciences, Flammarion*.
- Kamega A., Planchet F. [2011], « Hétérogénéité : mesure du risque d'estimation dans le cas d'une modélisation intégrant des facteurs observables », *Bulletin Français d'Actuariat, Vol. 11, No. 21*.
- Klein J. P., Moeschberger M. L. [2005] « Survival Analysis – Techniques for Censored and Truncated Data », *Springer, 2nd edition*.
- Kone M.G.K. [2007] « État des lieux de l'assurance vie dans les pays de la CIMA », *CIMA*.
- Planchet F. [2006], « Tables de mortalité d'expérience pour les portefeuilles de rentiers – (Tables TGH 05 et TGF 05) – Notice de présentation », *Institut des actuaires*.
- Planchet F., Kamega A., Ziguélé M. [2010], « Confection des tables de mortalité réglementaires pour la zone CIMA », *CIMA/WINTER & Associés (confidentiel)*.
- Planchet F., Théron P. [2006] « Modèles de Durée - Applications actuarielles », *Economica*.
- Quashie A., Denuit M. [2005] « Modèles d'extrapolation de la mortalité aux grands âges », *Institut des Sciences Actuarielles et Institut de Statistique, Université Catholique de Louvain, WP*.

Annexe A : retraitement des données

À la réception des mises à jour faisant suite au premier traitement des données (cf. 1.1.2), un second traitement est effectué sur les données des 32 assureurs vie *a priori* retenus pour les statistiques. Le volume de données initial pour ces 32 assureurs vie est de 387 247 enregistrements pour la zone CEMAC (7 assureurs) et 928 699 enregistrements pour la zone UEMOA (25 assureurs).

Le retraitement mené a conduit à la suppression d'enregistrements susceptibles de biaiser l'analyse de la mortalité, conformément aux éléments présentés ci-après.

A.1. Liste des traitements

Doublons

Les données sont retraitées des doublons établis à partir du matricule de l'assuré, de sa date de naissance et de son sexe. En pratique, seul un enregistrement est retenu en présence de doublons :

- si les informations renseignées pour les enregistrements en doublons sont strictement identiques, le choix de l'enregistrement retenu est effectué par son ordre d'apparition dans le fichier ;
- si des informations diffèrent, notamment au niveau des dates d'entrées et/ou de sortie (il s'agit alors d'assurées ayant renouvelé leur contrat ou en ayant souscrits plusieurs), l'enregistrement retenu est celui avec la date d'entrée renseignée la plus ancienne (dans ce cas, la date de sortie affectée à cet enregistrement est la plus tardive parmi l'ensemble des dates de sortie des doublons concernés, ce qui permet de cumuler les expositions au risque).

Suite à ce retraitement, 23 959 enregistrements ont été supprimés pour les données de la sous-région CEMAC et 39 843 enregistrements ont été supprimés pour la sous-région UEMOA.

Date d'entrée postérieure ou égale à la date de sortie

Lorsque la date d'entrée renseignée est postérieure ou égale à la date de sortie, les enregistrements sont exclus du périmètre des données.

Suite à ce retraitement, 9 256 enregistrements ont été supprimés pour les données de la sous-région CEMAC et 2 680 enregistrements ont été supprimés pour la sous-région UEMOA.

Matricules non renseignés

Lorsque le matricule (ou de manière générale l'identifiant de l'assuré) n'est pas renseigné, les enregistrements sont exclus du périmètre des données.

Suite à ce retraitement, 10 enregistrements ont été supprimés pour les données de la sous-région CEMAC et aucun enregistrement n'a été supprimé pour la sous-région UEMOA.

Date de sortie erronée

Lorsque la date de sortie (pour un motif autre que le décès ici) renseignée est erronée (format inexploitable) les enregistrements sont exclus du périmètre des données.

Suite à ce retraitement, 1 enregistrement a été supprimé pour les données de la sous-région CEMAC et 3 enregistrements ont été supprimés pour la sous-région UEMOA.

Date de naissance non renseignée, erronée ou incohérente

Lorsque la date de naissance est non renseignée, erronée ou incohérente (format inexploitable, date de naissance postérieure à la date d'entrée, etc.) les enregistrements sont exclus du périmètre des données.

Suite à ce retraitement, 29 692 enregistrements ont été supprimés pour les données de la sous-région CEMAC et 920 enregistrements ont été supprimés pour la sous-région UEMOA.

Par ailleurs, les enregistrements pour lesquels la date de naissance renseignée est incomplète sont maintenus dans le périmètre en retenant les conventions suivantes :

- lorsque le mois et l'année de naissance sont renseignés mais que le jour ne l'est pas, la date de naissance retenue est le 15 du mois et de l'année renseignés ;
- lorsque l'année de naissance est renseignée mais que le mois et le jour ne le sont pas, la date de naissance retenue est le 01 juillet de l'année renseignée.

En d'autres termes, il est supposé qu'en l'absence de précision, la naissance a lieu au milieu de l'année ou au milieu du mois de naissance renseigné.

Date de sortie non renseignée, erronée ou incohérente

Lorsque la date de sortie en cas de décès est non renseignée ou incohérente (date de décès postérieure à la date d'envoi du fichier) les enregistrements sont exclus du périmètre des données.

Suite à ce retraitement, 1 enregistrement a été supprimé pour les données de la sous-région CEMAC et 17 enregistrements ont été supprimés pour la sous-région UEMOA.

Sexe non renseigné ou erroné

Lorsque le code sexe n'est pas renseigné ou est erroné les enregistrements sont exclus du périmètre des données.

Suite à ce retraitement, 5 019 enregistrements ont été supprimés pour les données de la sous-région CEMAC et 7 513 enregistrements ont été supprimés pour la sous-région UEMOA.

A.2. Détail des retraitements par assureur

Les tableaux suivants présentent le détail des retraitements individuels réalisés pour chaque assureur.

Tableau 11 – Retraitement des données (CEMAC)

Traitement	CEMAC							Total CEMAC
	Assureur C1	Assureur C2	Assureur C3	Assureur C4	Assureur C5	Assureur C6	Assureur C7	
Nombre de lignes avant retraitements	73 812	172 096	50 899	2 844	18 912	13 981	54 703	387 247
Lignes en doublons	11 851	0	1	1 075	3 800	2 658	4 574	23 959
Incohérence date d'entrée et date de sortie	534	103	5 042	0	0	167	3 410	9 256
Matricule non renseigné	10	0	0	0	0	0	0	10
Date de sortie erronée	1	0	0	0	0	0	0	1
Date de naissance non renseignée, erronée ou incohérente	210	62	0	0	0	10	29 410	29 692
Date de sortie pour décès non renseignée ou incohérente	0	0	1	0	0	0	0	1
Sexe non renseigné ou erroné	0	4 612	0	0	397	1	9	5 019
Nombre de lignes supprimées	12 515	4 768	5 043	1 075	4 197	2 725	32 598	62 921
Nombre de lignes après retraitements	61 297	167 328	45 856	1 769	14 715	11 256	22 105	324 326

Tableau 12 – Retraitement des données (UEMOA, 1/3)

Traitement	UEMOA									
	Assureur U1	Assureur U2	Assureur U3	Assureur U4	Assureur U5	Assureur U6	Assureur U7	Assureur U8	Assureur U9	Assureur U10
Nombre de lignes avant retraitements	3 293	196 324	63 415	3 251	3 900	10 455	9 518	4 961	34 899	67 845
Lignes en doublons	0	0	0	9	3	0	563	175	1 135	7 230
Incohérence date d'entrée et date de sortie	0	0	14	0	0	0	1	958	86	0
Matricule non renseigné	0	0	0	0	0	0	0	0	0	0
Date de sortie erronée	0	0	0	0	0	3	0	0	0	0
Date de naissance non renseignée, erronée ou incohérente	1	6	19	0	0	0	12	0	307	0
Date de sortie pour décès non renseignée ou incohérente	0	0	0	0	0	0	0	0	0	0
Sexe non renseigné ou erroné	3	0	0	0	0	7 439	9	0	0	0
Nombre de lignes supprimées	4	6	33	9	3	7 442	576	1 092	1 223	7 230
Nombre de lignes après retraitements	3 289	196 318	63 382	3 242	3 897	3 013	8 942	3 869	33 676	60 615

Tableau 13 – Retraitement des données (UEMOA, 2/3)

Traitement	UEMOA							
	Assureur U11	Assureur U12	Assureur U13	Assureur U14	Assureur U15	Assureur U16	Assureur U17	Assureur U18
Nombre de lignes avant retraitements	39 901	77 721	1 574	26 859	43 152	10 449	41 070	72 619
Lignes en doublons	0	237	130	0	2 767	102	0	14 328
Incohérence date d'entrée et date de sortie	3	1	4	3	0	0	41	3
Matricule non renseigné	0	0	0	0	0	0	0	0
Date de sortie erronée	0	0	0	0	0	0	0	0
Date de naissance non renseignée, erronée ou incohérente	5	210	3	0	313	8	9	1
Date de sortie pour décès non renseignée ou incohérente	0	0	0	0	0	9	0	0
Sexe non renseigné ou erroné	0	0	0	0	0	0	0	8
Nombre de lignes supprimées	8	448	137	3	3 078	69	50	14 340
Nombre de lignes après retraitements	39 893	77 273	1 437	26 856	40 074	10 380	41 020	58 279

Tableau 14 – Retraitement des données (UEMOA, 3/3)

Traitement	UEMOA							Total UEMOA
	Assureur U19	Assureur U20	Assureur U21	Assureur U22	Assureur U23	Assureur U24	Assureur U25	
Nombre de lignes avant retraitements	75 413	14 444	17 028	63 546	4 568	32 109	10 385	928 699
Lignes en doublons	10 984	1 486	0	0	694	0	0	39 843
Incohérence date d'entrée et date de sortie	472	2	0	1 087	1	1	3	2 680
Matricule non renseigné	0	0	0	0	0	0	0	0
Date de sortie erronée	0	0	0	0	0	0	0	3
Date de naissance non renseignée, erronée ou incohérente	5	3	1	16	1	0	0	920
Date de sortie pour décès non renseignée ou incohérente	0	0	0	8	0	0	0	17
Sexe non renseigné ou erroné	0	0	0	3	51	0	0	7 513
Nombre de lignes supprimées	11 461	1 490	1	1 114	731	1	3	50 552
Nombre de lignes après retraitements	63 952	12 954	17 027	62 432	3 837	32 108	10 382	878 147

Annexe B : mise à jour du périmètre

Les tableaux suivants présentent les statistiques pour les assureurs supprimés du périmètre des données en raison de la faible exposition au risque (inférieure à 1 000), de l'absence de décès observé ou de l'existence d'une forte hétérogénéité inexplicée au sein d'un pays.

Tableau 15 – Exclusion pour faible exposition (CEMAC, femme)

Assureur	Exposition 2003-2006	Âge moyen	Taux de décès moyen	Taux de décès moyen (borne inf. à 95%)	Taux de décès moyen (borne sup. à 95%)	Âge moyen au décès
Assureur C4	448	39 ans	0,6698%	0,0000%	1,4277%	45 ans

Tableau 16 – Exclusion pour faible exposition (UEMOA, homme)

Assureur (sexe)	Exposition 2003-2006	Âge moyen	Taux de décès moyen	Taux de décès moyen (borne inf. à 95%)	Taux de décès moyen (borne sup. à 95%)	Âge moyen au décès
Assureur U16	13	38 ans	0,0000%	0,0000%	0,0000%	s.o.
Assureur U4	16	41 ans	0,0000%	0,0000%	0,0000%	s.o.
Assureur U6	916	41 ans	0,0000%	0,0000%	0,0000%	s.o.
Assureur U13	56	37 ans	0,0000%	0,0000%	0,0000%	s.o.
Assureur U23	0	s.o.	s.o.	s.o.	s.o.	s.o.

Tableau 17 – Exclusion pour faible exposition (UEMOA, femme)

Assureur (sexe)	Exposition 2003-2006	Âge moyen	Taux de décès moyen	Taux de décès moyen (borne inf. à 95%)	Taux de décès moyen (borne sup. à 95%)	Âge moyen au décès
Assureur U16	4	38 ans	0,0000%	0,0000%	0,0000%	s.o.
Assureur U1	370	39 ans	0,0000%	0,0000%	0,0000%	s.o.
Assureur U4	5	40 ans	0,0000%	0,0000%	0,0000%	s.o.
Assureur U6	283	39 ans	0,0000%	0,0000%	0,0000%	s.o.
Assureur U13	29	39 ans	0,0000%	0,0000%	0,0000%	s.o.
Assureur U23	0	s.o.	s.o.	s.o.	s.o.	s.o.

Tableau 18 – Exclusion pour absence de décès (CEMAC, homme)

Assureur	Exposition 2003-2006	Âge moyen	Taux de décès moyen	Taux de décès moyen (borne inf. à 95%)	Taux de décès moyen (borne sup. à 95%)	Âge moyen au décès
Assureur C6	12 860	44 ans	0,0000%	0,0000%	0,0000%	s.o.

Tableau 19 – Exclusion pour absence de décès (CEMAC, femme)

Assureur	Exposition 2003-2006	Âge moyen	Taux de décès moyen	Taux de décès moyen (borne inf. à 95%)	Taux de décès moyen (borne sup. à 95%)	Âge moyen au décès
Assureur C6	3 904	42 ans	0,0000%	0,0000%	0,0000%	s.o.

Tableau 20 – Exclusion pour absence de décès (UEMOA, homme)

Assureur	Exposition 2003-2006	Âge moyen	Taux de décès moyen	Taux de décès moyen (borne inf. à 95%)	Taux de décès moyen (borne sup. à 95%)	Âge moyen au décès
Assureur U1	1 065	38 ans	0,0000%	0,0000%	0,0000%	s.o.

Tableau 21 – Exclusion pour faible taux de décès et forte hétérogénéité par pays (UEMOA, homme)

Assureur	Exposition 2003-2006	Âge moyen	Taux de décès moyen	Taux de décès moyen (borne inf. à 95%)	Taux de décès moyen (borne sup. à 95%)	Âge moyen au décès
Assureur U17	14 420	39 ans	0,9293%	0,7719%	1,0866%	44 ans
Assureur U18	97 752	40 ans	0,0481%	0,0343%	0,0618%	46 ans
Assureur U19	23 174	41 ans	0,0820%	0,0451%	0,1189%	47 ans
Assureur U20	17 163	36 ans	0,3088%	0,2257%	0,3919%	47 ans
Assureur U11	49 260	44 ans	0,0406%	0,0228%	0,0584%	49 ans
Assureur U12	266 334	48 ans	0,0267%	0,0205%	0,0329%	50 ans
Assureur U14	51 630	44 ans	0,3060%	0,2583%	0,3537%	47 ans
Assureur U15	108 288	45 ans	0,0887%	0,0709%	0,1064%	52 ans
Total	628 022	45 ans	0,0952%	0,0876%	0,1029%	48 ans

Tableau 22 – Exclusion pour faible taux de décès et forte hétérogénéité par pays (UEMOA, femme)

Assureur	Exposition 2003-2006	Âge moyen	Taux de décès moyen	Taux de décès moyen (borne inf. à 95%)	Taux de décès moyen (borne sup. à 95%)	Âge moyen au décès
Assureur U17	4 727	39 ans	0,2539%	0,1102%	0,3975%	40 ans
Assureur U18	16 553	40 ans	0,0302%	0,0037%	0,0567%	43 ans
Assureur U19	10 524	39 ans	0,0950%	0,0361%	0,1539%	45 ans
Assureur U20	1 811	42 ans	0,3314%	0,0662%	0,5965%	47 ans
Assureur U11	17 435	44 ans	0,0229%	0,0005%	0,0454%	43 ans
Assureur U12	31 193	44 ans	0,0224%	0,0058%	0,0391%	49 ans
Assureur U14	10 113	44 ans	0,1681%	0,0882%	0,2480%	44 ans
Assureur U15	14 944	43 ans	0,0268%	0,0005%	0,0530%	44 ans
Total	107 300	43 ans	0,0606%	0,0459%	0,0753%	44 ans

PARTIE II

MORTALITÉ ACTUELLE : MODÉLISATION DE L'HÉTÉROGÉNÉITÉ ET MESURE DES RISQUES ASSOCIÉS DANS LA ZONE

Cette deuxième partie cherche à prendre en compte l'hétérogénéité entre pays dans la mortalité d'expérience au sein de la région. Cet exercice conduit à segmenter la population en sous-populations, et donc à réduire la taille des échantillons dans les modèles. Avec ces échantillons réduits, il apparaît alors des fluctuations d'échantillonnages plus importantes et donc une plus grande incertitude dans les estimations des paramètres des modèles (risque d'estimation). Une attention particulière est alors accordée dans cette deuxième partie au choix de modèles permettant de tenir compte de l'hétérogénéité, tout en limitant le niveau de l'incertitude liée aux fluctuations d'échantillonnage.

Chapitre 3

Mesure du risque d'estimation en présence d'hétérogénéité¹

Sommaire

CHAPITRE 3 MESURE DU RISQUE D'ESTIMATION EN PRÉSENCE D'HÉTÉROGÉNÉITÉ	93
1. Introduction.....	94
2. Intervalles et bandes de confiance	95
2.1.Intervalles de confiance sur les taux bruts (Hoem).....	96
2.2.Intervalles et bandes de confiance sur la fonction de survie (Kaplan Meier) 96	
2.3.Intervalles et bandes de confiance : illustrations.....	97
3. Mesure du risque d'estimation sur les taux ajustés	99
3.1.Méthode 1 : simulation directe des taux bruts	100
3.2.Méthode 2 : simulation des résidus.....	102
3.3.Mesure sur les taux ajustés : illustrations.....	103
4. Mesure du risque d'estimation sur les provisions	108
4.1.Provisions déterministes	108
4.2.Provisions stochastiques	109
4.3.Mesure sur les provisions : illustrations.....	111
5. Évolution du risque d'estimation en présence d'hétérogénéité.....	114
5.1.Identification et mesure de l'hétérogénéité	115
5.2.Évolution du risque d'estimation sur les taux de décès ajustés.....	118
5.3.Évolution du risque d'estimation sur les provisions	120
6. Synthèse et conclusion	122
Notations des taux de décès.....	124
Bibliographie.....	125

¹ Ce chapitre s'appuie principalement sur l'article Kamega et Planchet [2010], relatif à la mesure du risque d'estimation associé à une table d'expérience.

Ce chapitre présente des outils de mesure du risque d'estimation à partir d'un modèle de Brass (modèle à référence externe). Ces outils sont dans un premier temps appliqués à la mesure de ce risque dans le cadre de la construction d'une loi d'expérience à partir d'une population globale (i. e. sans tenir compte de l'hétérogénéité). Dans un second temps, ces outils sont appliqués à la mesure du risque d'estimation dans le cadre de la construction d'une loi d'expérience à partir d'une sous-population de la population globale, segmentée au titre de la prise en compte de l'hétérogénéité. Dans ce second scénario il apparaît que le risque d'estimation augmente significativement : la modélisation de l'hétérogénéité à partir d'un modèle indépendant pour chaque sous-population n'est donc pas adaptée car elle majore significativement le risque d'estimation.

1. Introduction

Comme indiqué dans Planchet [2009], l'évaluation des provisions techniques à partir de la meilleure estimation (*best estimate*) des flux de trésorerie nécessite de prendre en compte l'expérience du portefeuille assuré. À cet effet, la détermination d'une loi d'expérience est nécessaire, que celle-ci soit déterminée complètement à l'aide des données disponibles ou calibrée par un positionnement par rapport à une référence externe (*cf.* Planchet et Leroy [2009]). Il s'avère alors indispensable de prendre en compte l'hétérogénéité de la population observée afin de limiter le biais associé (*cf.* Dreesbeke et al. [1989]) et de conférer ainsi à la loi d'expérience une plus grande robustesse face aux déformations futures de la structure de la population.

En pratique, quelle que soit la sophistication du modèle proposé, la prise en compte de l'hétérogénéité conduit de fait à segmenter la population d'expérience en sous-populations. Cependant, plus la segmentation est fine et plus les risques systématiques associés à la construction de la loi d'expérience sont importants (principalement en raison de la baisse du volume d'information et de la multiplication des modèles, *cf.* Alho [2007]). En d'autres termes, plus on segmente et plus la volatilité autour des tables ajustées augmente, et il arrive un moment où la segmentation devient inefficace.

Au surplus, à l'échelle internationale, la mutation du secteur conduit à porter une attention nouvelle aux risques non mutualisables portés par les organismes assureurs, du fait de leur poids potentiel dans la queue de la distribution et du recours fréquent à un quantile d'ordre élevé dans différentes évaluations (pour le calcul du capital de solvabilité dans le dispositif Solvabilité 2 en Europe, pour le calcul du coût des risques non couvrables (CRNHR) des MCEV en Europe ou pour le calcul de l'ajustement pour risque dans la norme IFRS sur les contrats d'assurance).

Une réflexion s'impose alors sur le choix du niveau de segmentation optimal. À cet effet, la quantification du risque d'estimation, c'est-à-dire le risque systématique associé à l'utilisation d'une loi de durée dont on sait qu'elle n'est pas exacte du fait des erreurs d'échantillonnage affectant sa construction, peut être un critère utile.

Dans ce cadre, ce chapitre présente des outils de mesure du risque d'estimation associé à la construction de tables de mortalité d'expérience à partir d'un modèle à référence externe (modèle de Brass). L'approche proposée ici ne s'appuie pas sur le cadre général du maximum de vraisemblance, qui propose pourtant des outils naturels de mesure du risque d'estimation, car il n'est pas très bien adapté au contexte des tables d'expérience (en pratique, on est souvent amené à recourir à des techniques qui ne relèvent pas du maximum de vraisemblance).

De manière générale, les outils de mesure de risques présentés ici s'avèrent particulièrement utiles lorsque les données d'expérience disponibles se révèlent peu volumineuses ou lacunaires.

Plan

L'étude s'attache donc à mesurer le risque d'estimation (sections 2, 3 et 4), et à quantifier son évolution en cas de baisse de la taille de l'échantillon suite à une segmentation dans un contexte d'hétérogénéité (section 5).

Dans le détail, la deuxième section de l'étude présente les intervalles de confiance associés aux erreurs d'échantillonnage, tant au niveau de l'estimation des taux de décès bruts qu'au niveau de l'estimation de la fonction de survie. Dans la troisième et la quatrième section, on s'intéresse ensuite aux mesures du risque d'estimation pour des taux de décès ajustés et dans le cadre d'une problématique de provisionnement (ce qui peut être utile pour définir ensuite une fonction de perte). Enfin, dans la cinquième section on segmente la population afin de tenir compte de l'hétérogénéité et on étudie l'évolution du risque d'estimation lors du passage d'une modélisation de la population globale à une modélisation indépendante d'une sous-population.

Données

Les illustrations numériques présentées dans cette étude s'appuient sur les données d'expérience utilisées pour la construction des tables de mortalité réglementaires dans la zone CIMA/FANAF pour les contrats d'assurance en cas de vie et en cas de décès. Les données retenues s'appuient plus précisément sur les données de la population masculine des pays de la sous-région UEMOA, représentés ici par la Côte d'Ivoire, le Mali et le Togo.

Pour mémoire, les données utilisées ont été collectées en 2009 et sont présentées en détail dans le deuxième chapitre. On retient ici qu'elles sont observées sur les années 2003 à 2006 et comptent des troncatures à gauche (relatives aux entrées après le 01/01/2003) et des censures à droite (relatives aux sorties avant le 31/12/2006 pour un motif autre que le décès).

Dans les sections 2 à 4, toutes les applications numériques sont réalisées à partir de la population d'expérience UEMOA globale. Puis, dans la section 5 on distingue les données par pays afin de prendre en compte l'hétérogénéité à partir d'un modèle indépendant pour chaque pays.

2. Intervalles et bandes de confiance

Cette section présente les intervalles et bandes de confiance, respectivement pour les taux bruts et pour la fonction de survie.

2.1. Intervalles de confiance sur les taux bruts (Hoem)

Les taux de décès bruts, c'est-à-dire les taux de décès observés, présentent des irrégularités directement liées aux fluctuations d'échantillonnage. La description des estimations de taux de décès observés par les seules valeurs estimées est donc insuffisante : il est nécessaire de les décrire par des indicateurs numériques complémentaires. À cet effet, on peut retenir les intervalles de confiance (pour mémoire, dans ce contexte les intervalles de confiance permettent d'encadrer les taux de décès théoriques réels au titre des fluctuations d'échantillonnage).

En d'autres termes, on cherche ici un intervalle de confiance pour q_x (taux de décès réel et théorique à l'âge x dans la population) à partir de \hat{q}_x (taux de décès observé à l'âge x dans un échantillon exposé au risque de taille R_x). À cet effet, on retient l'estimateur de taux bruts de Hoem, qui considère entre autres que les décès suivent une loi binomiale et que l'on a, pour tout $0 < t < s < 1$, l'approximation suivante : ${}_{s-t}q_{x+t} \approx (s-t)q_x$ (cf. Planchet et Thérond [2006] pour une présentation complète des hypothèses de calculs de cet estimateur). Deux approches sont envisageables pour la détermination des intervalles de confiance dans ce cas.

Si R_x est grand (c'est-à-dire $R_x \hat{q}_x > 5$ et $R_x(1 - \hat{q}_x) > 5$, cf. Saporta [2006]), on a :

$$Q_x \sim N\left(q_x; \sigma_x = \sqrt{\frac{q_x(1-q_x)}{R_x}}\right),$$

et on déduit que l'expression approchée des bornes de l'intervalle de confiance de q_x est :

$$\hat{q}_x \pm u_{\alpha/2} \sqrt{\frac{\hat{q}_x(1-\hat{q}_x)}{R_x}},$$

où $u_{\alpha/2}$ désigne le quantile d'ordre $\alpha/2$ de la loi normale centrée réduite et $1 - \alpha$ désigne le niveau de confiance.

Si R_x n'est pas assez grand, on peut calculer l'intervalle de confiance à distance finie à partir d'une procédure itérative intégrant la probabilité associée au nombre de décès (cf. Planchet et Thérond [2006] pour plus de précisions sur cette alternative). Ce cas nous intéresse peu en pratique.

2.2. Intervalles et bandes de confiance sur la fonction de survie (Kaplan Meier)

L'estimateur de la fonction de survie Kaplan Meier présente l'avantage de ne retenir aucune hypothèse sur la répartition des décès sur $[x, x+1]$. En outre,

il permet également d'obtenir des intervalles de confiance et à cet effet, on s'appuie sur l'estimateur de Greenwood de la variance de l'estimateur de Kaplan Meier :

$$\hat{V}[\hat{S}(x)] = \hat{S}(x)^2 \gamma(x)^2,$$

où $\hat{S}(x)$ représente l'estimateur de la fonction de survie à l'âge x et

$$\gamma(x) = \sqrt{\sum_{i=1}^x \frac{d_i}{R_i(R_i - d_i)}}, \text{ avec } d_i \text{ le nombre de décès observés à l'âge } i.$$

Dans la littérature, on retrouve généralement deux types d'intervalles de confiance pour la fonction de survie (*cf.* par exemple Klein et Moeschberger [2005]) : les intervalles de confiance ponctuels, c'est-à-dire les intervalles de confiance à un âge x fixé, et les bandes de confiance, c'est-à-dire les intervalles de confiance pour tous les âges compris entre x_m et x_M .

Dans un contexte d'analyse des taux de décès, et en notant que $q_x = 1 - \frac{S(x+1)}{S(x)}$, on comprend que le choix des bandes confiance pour la

fonction de survie est plus approprié. Aussi, on présente ici les bandes de confiance linéaires pour la fonction de survie (ces bandes présentant l'avantage d'être symétriques, contrairement à celles issues de transformations logarithmiques par exemple, *cf.* Klein et Moeschberger [2005]). Nair [1984] définit ces bandes de confiance par :

$$\hat{S}(x) \pm c_\alpha(a_{x_m}, a_{x_M}) \hat{S}(x) \gamma(x),$$

où $a_X = \frac{p \times \gamma(X)^2}{1 + p \times \gamma(X)^2}$ (p étant la taille de l'échantillon) et $c_\alpha(a_{x_m}, a_{x_M})$

désigne le coefficient de confiance (issu de tables présentées en annexe de Klein et Moeschberger [2005] par exemple). On note que les bandes de confiance linéaires ci-dessus présentent l'avantage d'être proportionnelles aux intervalles de confiance ponctuels linéaires (pour mémoire ces intervalles ponctuels s'écrivent $\hat{S}(x) \pm u_{\alpha/2} \hat{S}(x) \gamma(x)$, où $u_{\alpha/2}$ désigne le quantile d'ordre $\alpha/2$ de la loi normale centrée réduite et $1 - \alpha$ désigne le niveau de confiance, *cf.* par exemple Planchet et Thérond [2006] ou Martinussen et Scheike [2006]), contrairement aux bandes de confiance linéaires de Hall et Wellner [1980].

2.3. Intervalles et bandes de confiance : illustrations

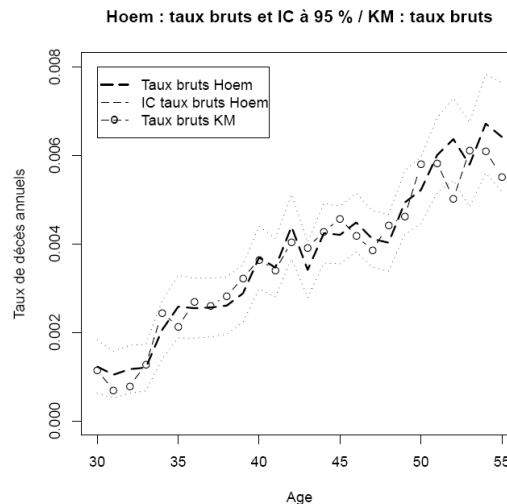
Pour les applications numériques, on retient $x_m = 30$ et $x_M = 55$ et on considère à ce stade les données d'expérience globale de l'UEMOA. Par ailleurs, on note qu'avec les données de l'étude et pour un niveau de confiance

à $1-\alpha=95\%$, $c_\alpha(a_{x_m}, a_{x_M})$ est majoré par 2,543 (avec $a_{x_m}=0,717$ et $a_{x_M}=0,761$, cf. Klein et Moeschberger [2005]). Aussi, pour l'illustration des bandes de confiance de la fonction de survie présentée en 2.3.2, on suppose que $c_\alpha(a_{x_m}, a_{x_M})=2,543$.

2.3.1. Illustrations et comparaisons pour les taux bruts

Le graphique ci-dessous présente les taux de décès bruts estimés selon la méthode de Hoem, les intervalles de confiance à 95 % associés, ainsi que les taux bruts déduits de l'estimation de la fonction de survie selon Kaplan Meier.

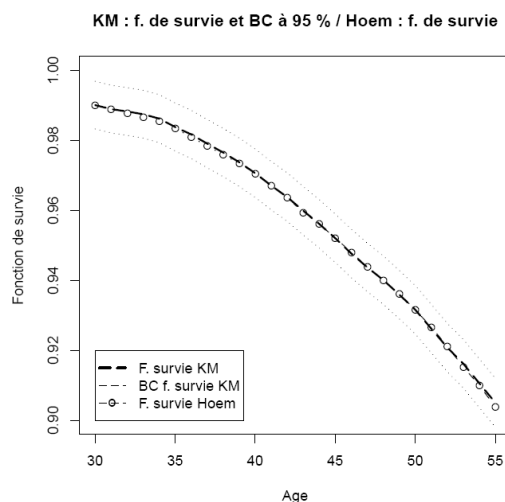
Figure 24 - Taux de décès bruts et intervalles de confiance (UEMOA)



On note que les deux méthodes d'estimation des taux bruts sont comparables et que les taux bruts estimés selon l'approche Kaplan Meier sont compris dans les bornes de l'intervalle de confiance des taux bruts estimés par Hoem (hormis à 52 ans).

2.3.2. Illustrations et comparaisons pour la fonction de survie

Le graphique ci-dessous présente la fonction de survie estimée selon la méthode de Kaplan Meier, les bandes de confiance à 95 % associées, ainsi que la « fonction de survie » déduite de l'estimation des taux bruts selon Hoem (à noter que dans un souci de comparabilité des deux fonctions de survie, il est supposé ici qu'à l'âge initial $x_m=30$, la fonction de survie déduite de Hoem est égale à celle issue de Kaplan Meier).

Figure 25 - Fonctions de survie et bandes de confiance (UEMOA)

Il ressort de cette illustration que les deux fonctions de survie sont très proches, ce qui est logique au regard de la Figure 24.

3. Mesure du risque d'estimation sur les taux ajustés

Cette section présente les mesures du risque systématique lié aux fluctuations d'échantillonnage à travers les taux de décès ajustés (il s'agit ainsi du risque d'estimation). À cet effet, on s'appuie sur un modèle de Brass, c'est-à-dire un modèle à référence externe tel que :

$$y_x = a \times z_x + b + \varepsilon_x,$$

avec $y_x = \ln\left(\frac{\hat{q}_x}{1 - \hat{q}_x}\right)$, $z_x = \ln\left(\frac{q_x^{ref}}{1 - q_x^{ref}}\right)$ et $\varepsilon_x \sim N(0; \sigma_\varepsilon)$. On note par ailleurs $\theta = (a, b)$.

Ce chapitre présente le risque d'estimation au titre des fluctuations d'échantillonnage, c'est-à-dire au titre des données uniquement. Dans ce contexte, on suppose qu'il n'y a pas de risque de modèle, c'est-à-dire que l'on fait l'hypothèse que le modèle est correctement spécifié. Cette hypothèse revient à considérer que les erreurs ε_x présentés *supra* ne correspondent pas à l'erreur de spécification du modèle, mais à l'erreur autour de $\hat{y}_x = \ln\left(\frac{q_x(\hat{\theta})}{1 - q_x(\hat{\theta})}\right) = \hat{a} \times z_x + \hat{b}$ au titre des fluctuations d'échantillonnage (la mesure du risque d'estimation présentée dans la sous-section 3.2 ci-après est largement basée sur cette interprétation).

On note par ailleurs que la détermination des paramètres $\theta = (a, b)$ est au cœur du dispositif de mesure du risque d'estimation (*cf.* les sous-sections 3.1 et 3.2). Une attention particulière doit donc être accordée à la méthode et à la qualité de l'estimation des paramètres. L'estimation par maximum de vraisemblance peut être biaisée, car la taille de l'échantillon est limitée à

$n = x_M - x_m + 1$ observations. En revanche, l'estimation par les moindres carrés n'est pas biaisée, et c'est donc cette approche qui est retenue ici. On rappelle toutefois que l'estimation issue de cette approche n'est pas robuste (étant donnée que l'ajout ou la suppression d'observations dans les données peut modifier sensiblement l'estimation ; *cf.* Meucci [2007] pour une illustration des mesures de robustesse et une démonstration de l'absence de robustesse des estimations par moindres carrés) et qu'il convient ainsi d'analyser les résultats avec prudence.

Pour mémoire les taux bruts présentent des irrégularités directement liées aux erreurs d'échantillonnage (*cf.* section 2). Ainsi, les deux approches présentées ici pour mesurer le risque d'estimation, au titre des fluctuations d'échantillonnage, sur les taux de décès ajustés consistent à générer de manière aléatoire, directement ou indirectement, des taux bruts (suivant la loi adéquate), afin d'en déduire l'impact sur l'estimation des paramètres $\theta = (a, b)$ du modèle de Brass, puis l'impact sur les taux de décès ajustés. La première approche s'appuie sur la simulation directe des taux bruts (*cf.* 3.1) alors que dans la seconde, les « taux de décès bruts simulés » sont issus de la simulation des résidus du modèle (*cf.* 3.2). Dans les deux cas, la technique de simulation retenue consiste à considérer une méthode de Monte-Carlo pour simuler la distribution d'une loi normale (à partir de la simulation d'une distribution de loi normale centrée et réduite, elle-même déduite d'une simulation de distribution de loi uniforme entre 0 et 1).

3.1. Méthode 1 : simulation directe des taux bruts

On se place dans le cas des taux bruts estimés selon l'approche de Hoem. La première étape consiste alors à générer k simulations ($k \in [1, K]$) des taux de décès bruts, pour tous les âges x ($x \in [x_m, x_M]$), selon

$$Q_x \sim N \left(\hat{q}_x, \sqrt{\frac{\hat{q}_x (1 - \hat{q}_x)}{R_x}} \right).$$

L'étape suivante consiste à calculer pour chaque simulation k une estimation des paramètres $\theta^k = (a^k, b^k)$. Pour mémoire, le modèle est de la forme :

$$y_x^k = a^k z_x + b^k + \varepsilon_x^k,$$

où $y_x^k = \ln\left(q_x^k / (1 - q_x^k)\right)$ et q_x^k est une réalisation de Q_x . En retenant un ajustement par la méthode des moindres carrés ordinaires et en notant $n = x_M - x_m + 1$, on a ainsi (cf. par exemple Bourbonnais [2004]) :

$$\hat{a}^k = \frac{\sum_{x=x_m}^{x_M} y_x^k \times z_x - n \times \bar{y}^k \times \bar{z}}{\sum_{x=x_m}^{x_M} z_x^2 - n \times \bar{z}^2},$$

et

$$\hat{b}^k = \bar{y}^k - \hat{a}^k \times \bar{z},$$

où la notation \bar{v} représente la moyenne arithmétique de la série v .

Pour la troisième étape, on note que :

$$\ln\left(q_x\left(\hat{\theta}^k\right) / \left(1 - q_x\left(\hat{\theta}^k\right)\right)\right) = \hat{y}_x^k = \hat{a}^k z_x + \hat{b}^k,$$

et on en déduit k réalisations (au titre des fluctuations d'échantillonnage) des taux ajustés (on parle alors de taux simulés) :

$$q_x\left(\hat{\theta}^k\right) = \frac{\exp\left(\hat{a}^k z_x + \hat{b}^k\right)}{1 + \exp\left(\hat{a}^k z_x + \hat{b}^k\right)}.$$

Le risque d'estimation peut alors être mesuré par le coefficient :

$$c\left(\psi_x\right) = \frac{\psi_x}{q_x\left(\hat{\theta}\right)},$$

où $q_x\left(\hat{\theta}\right) = \frac{\exp\left(\hat{a} \times z_x + \hat{b}\right)}{1 + \exp\left(\hat{a} \times z_x + \hat{b}\right)}$ correspond au taux de décès ajusté à partir de

l'estimation du taux de décès brut observé à l'âge x (on parle alors de taux

ajusté) et $\psi_x = \sqrt{E\left[\left(q_x\left(\hat{\theta}^k\right) - q_x\left(\hat{\theta}\right)\right)^2\right]}$. Ce coefficient constitue une mesure

de dispersion, en pourcentage et au titre du risque d'estimation, des taux de décès simulés autour du taux de décès ajusté. On peut noter au passage qu'on utilise ici comme taux de décès de « base » le taux obtenu avec la valeur *best estimate* des paramètres, qui ne coïncide pas avec sa vision *best estimate* du fait de la non linéarité de la transformation qui lie ces deux grandeurs.

3.2. Méthode 2 : simulation des résidus

Conformément aux indications présentées en introduction de la section 3, les résidus du modèle correspondent dans notre étude à une estimation des erreurs au titre des fluctuations d'échantillonnage (pour mémoire, les résidus du modèle sont $e_x = y_x - \hat{a} \times z_x - \hat{b}$, où $\hat{\theta} = (\hat{a}, \hat{b})$ sont les paramètres du modèle estimés par la méthode des moindres carrés ordinaires). Sur ces bases, cette seconde approche pour mesurer le risque d'estimation s'appuie sur la simulation des résidus (de laquelle on déduit des réalisations de taux bruts, puis des réalisations de taux ajustés au titre des fluctuations d'échantillonnage – on parle alors de taux simulés –). Par souci de cohérence avec les travaux de la sous-section précédente, les taux de décès ajustés $q_x(\hat{\theta})$, desquels on déduit les « vrais » résidus observés du modèle de Brass, sont déterminés à partir des taux bruts estimés selon l'approche de Hoem (on note toutefois que les taux bruts estimés à partir de la fonction de survie de Kaplan Meier auraient pu être retenus ici).

Pour mesurer le risque d'estimation à partir de la simulation des résidus, la première étape consiste à tester l'adéquation de la distribution des « vrais » résidus observés du modèle à la loi normale. Dans le cas d'échantillons de petites tailles (c'est-à-dire avec $n = x_M - x_m + 1 \leq 50$, cf. Rakotomalala [2008]), ce qui peut être fréquent dans le cas de tables instantanées, le test de Shapiro-Wilk est particulièrement robuste (pour mémoire, ce test est basé sur le rapport entre une combinaison linéaire des étendues successives et l'écart-type). À mesure que le nombre d'observations retenues pour l'ajustement du modèle de référence augmente, on privilégiera le test de d'Agostino, qui devient particulièrement efficace à partir de $n \geq 20$, cf. Rakotomalala [2008] (pour mémoire, le test d'Agostino est basé sur les coefficients d'aplatissement et d'asymétrie).

Dès que la compatibilité à la loi normale est vérifiée, dans une deuxième étape on calcule empiriquement la moyenne et l'écart-type de la distribution

des résidus, respectivement notés $\hat{\mu}_\varepsilon$ et $\hat{\sigma}_\varepsilon$ (on a donc $\hat{\mu}_\varepsilon = \frac{1}{n} \sum_{i=x_m}^{x_M} e_i$ et

$\hat{\sigma}_\varepsilon = \sqrt{\frac{1}{n-1} \sum_{i=x_m}^{x_M} (e_i - \hat{\mu}_\varepsilon)^2}$). On utilise alors ces paramètres pour générer des

erreurs aléatoires selon la loi normale, et on a donc pour différents scénarios k ($k \in [1, K]$) :

$$y_x^k = \hat{a} \times z_x + \hat{b} + e_x^k,$$

où e_x^k est une réalisation de $\hat{\varepsilon}_x^k$ et $\hat{\varepsilon}_x^k \sim N(\hat{\mu}_\varepsilon; \hat{\sigma}_\varepsilon)$. Les « taux de décès bruts » associés aux k scénarios sont alors ici :

$$q_x^k = \frac{\exp(\hat{a} \times z_x + \hat{b} + e_x^k)}{1 + \exp(\hat{a} \times z_x + \hat{b} + e_x^k)}.$$

Pour chaque scénario k , on calcule à partir de ces « taux bruts » une estimation des paramètres $\theta^k = (a^k, b^k)$ en retenant un ajustement par la méthode des moindres carrés ordinaires (à l'image de ce qui est présenté dans la sous-section 3.1). On en déduit alors k réalisations des taux ajustés (on parle alors de taux simulés) au titre du risque d'estimation :

$$q_x^k(\hat{\theta}^k) = \frac{\exp(\hat{a}^k z_x + \hat{b}^k)}{1 + \exp(\hat{a}^k z_x + \hat{b}^k)}.$$

À l'image de la sous-section 3.1, on calcule le coefficient :

$$c(\psi_x) = \frac{\psi_x}{q_x(\hat{\theta})},$$

afin d'avoir une mesure de dispersion, en pourcentage et au titre du risque d'estimation, des taux de décès simulés autour du taux de décès ajusté (avec

$$\text{désormais } \psi_x = \sqrt{E \left[\left(q_x^k(\hat{\theta}^k) - q_x(\hat{\theta}) \right)^2 \right]}.$$

3.3. Mesure sur les taux ajustés : illustrations

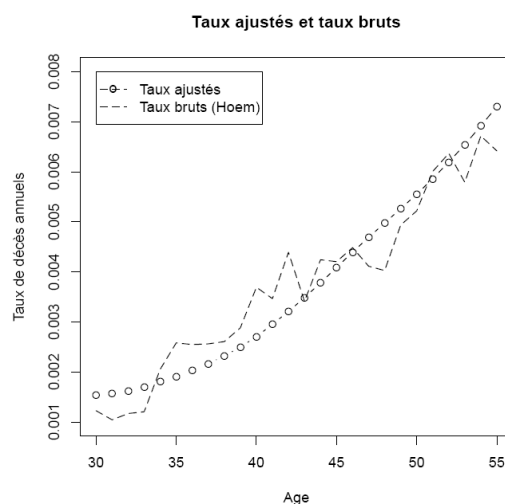
On présente ici une illustration et une comparaison des résultats obtenus selon les méthodes retenues pour la simulation des taux bruts au titre des fluctuations d'échantillonnage.

Dans l'ensemble des applications numériques, on s'appuie sur $K = 15\,000$ simulations des taux bruts et on considère que pour chaque simulation k les réalisations de la loi normale centrée réduite $N(0;1)$, desquels on déduit les réalisations appropriées pour chaque méthode, sont exactement les mêmes dans les deux méthodes. On note en outre que dans le cadre de la méthode décrite en 3.1, méthode qui pour mémoire se base sur la simulation directe des taux bruts, les taux bruts sont en pratique générés sous la contrainte $q_x^k > 0$ (lorsque cette condition n'est pas vérifiée pour une simulation k , celle-ci est annulée et reconduite pour les deux méthodes).

Par ailleurs, en pratique l'ajustement du modèle de Brass est réalisé en considérant la table réglementaire française TH00-02 (décès) comme table de référence (à l'image de ce qui a été fait dans le chapitre sur la construction des tables d'expérience). Cet ajustement conduit aux estimations suivantes : $\hat{a} = 0,799$ et $\hat{b} = -1,082$ (avec des probabilités de rejeter à tort l'hypothèse H_0 de nullité du paramètre selon le test de Student respectivement égales à $4,0 \cdot 10^{-12}$ et à $0,7\%$). Aussi, il ressort que le R^2 ajusté est égal à $86,5\%$. Enfin, on note que la probabilité associée au test Shapiro-Wilk sur les résidus est égale à $55,8\%$; en d'autres termes, on accepte l'hypothèse H_0 de normalité des résidus observés.

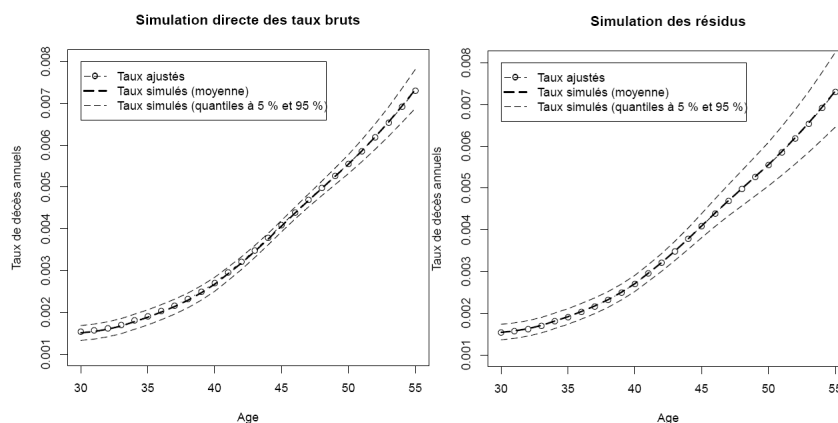
À titre d'information, le graphique suivant présente les taux ajustés et les taux bruts (estimés selon l'approche de Hoem).

Figure 26 - Taux de décès ajustés et taux de décès bruts (UEMOA)



3.3.1. Illustrations et comparaisons des taux de décès

Les graphiques suivants présentent les taux de décès ajustés issus des estimations $\hat{\theta} = (\hat{a}, \hat{b})$, ainsi que, pour les deux méthodes présentées, les taux de décès simulés au titre des fluctuations d'échantillonnage (dans ce cas, on présente les moyennes et les quantiles à 5% et à 95%).

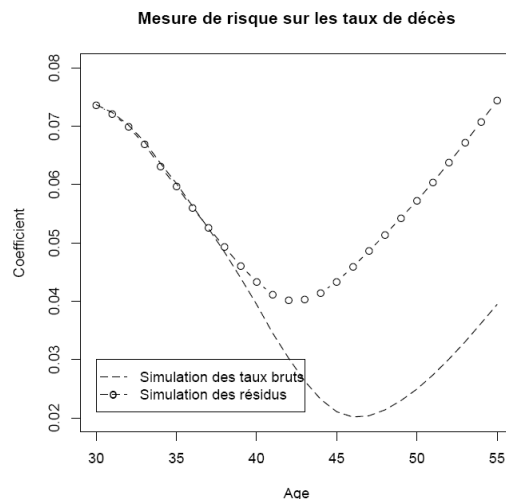
Figure 27 - Taux de décès ajustés et simulés (UEMOA)

Il ressort que la dispersion des taux de décès simulés issus de la simulation des résidus est plus importante que celle des taux simulés issus de la simulation directe des taux bruts.

Pour préciser ces observations sur les taux de décès, les paragraphes ci-dessous présentent une comparaison des mesures de risques entre les deux méthodes retenues.

3.3.2. Illustrations et comparaisons des mesures de risques sur les taux

Pour quantifier le risque d'estimation selon la méthode de simulation retenue, on présente ci-dessous le coefficient $c(\psi_x)$, introduit dans les sous-sections 3.1 et 3.2 et représentant la dispersion en pourcentage des taux simulés (au titre du risque d'estimation dû aux fluctuations d'échantillonnage) autour des taux ajustés. Dans ce contexte, le graphique suivant présente les valeurs de ce coefficient par méthode et par âge.

Figure 28 - Risque d'estimation sur les taux (par âge) (UEMOA)

Avant de commenter ce graphique, on présente ci-après une moyenne arithmétique du coefficient $c(\psi_x)$ pour tous les âges $x \in [x_m, x_M]$ de l'étude et par méthode de simulation des taux bruts.

Tableau 23 - Risque d'estimation sur les taux (moyenne) (UEMOA)

Méthode de simulation des taux bruts (risque d'estimation)	Mesure du risque d'estimation (en % des taux ajustés)
Méthode 1 : simulation « directe » des taux bruts (risque systématique)	$c(\psi) = 4,07 \%$
Méthode 2 : simulation des résidus (risque systématique)	$c(\psi) = 5,59 \%$

Il ressort de ces résultats que la mesure du risque d'estimation issue de la simulation des résidus majore en moyenne de plus de 35 % celle issue de la simulation directe des taux bruts.

Ces écarts s'expliquent par les hypothèses retenues. En effet, la méthode de simulation des résidus repose largement sur l'hypothèse d'absence d'erreur de spécification du modèle (ce qui revient à considérer que l'erreur du modèle est uniquement due aux fluctuations d'échantillonnage). Bien que cette hypothèse soit justifiée, elle ne reflète pas la réalité et en pratique, il existe une erreur de spécification du modèle. En d'autres termes, pour la méthode de simulation des résidus, les mesures de risque d'estimation présentées dans le graphique et le tableau ci-dessus incluent également une part de risque de modèle (d'où l'écart de mesure de risque observé entre les deux méthodes).

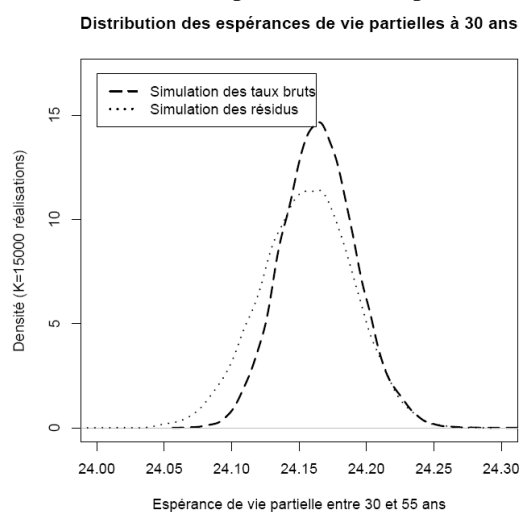
En particulier, et pour revenir à la Figure 28 on note que les valeurs de mesures de risques par âge présentées sont très liées à la structure du portefeuille d'expérience, au modèle et à son calibrage. En effet, d'une part

dans le cas de la simulation directe des taux bruts, les mesures de risques sont directement liées à l'exposition au risque et à l'écart entre les taux bruts observés et les taux ajustés. D'autre part, dans le cas de la simulation des résidus, on note que les « taux bruts » simulés ne sont pas directement estimés mais sont issus d'une transformation non linéaire qui est donc de nature à introduire un biais (dans ce cas, les résultats sont donc à apprécier avec prudence).

3.3.3. Illustrations et comparaisons des tables associées aux taux de décès

Outre la présentation et la comparaison des taux de décès, il est nécessaire de comparer les tables de mortalité associées, et à cet effet il convient d'utiliser une fonctionnelle propre à chaque table qui lui associe un nombre positif. L'espérance de vie est de ce point de vue une fonctionnelle naturelle (dans le cadre de l'assurance, le montant de la provision en est une autre, *cf.* section 4 ci-après). Dans ce contexte, le graphique suivant présente, pour chaque méthode de simulation, les estimations des fonctions de densité (par l'estimateur à noyau) des espérances de vie résiduelles (entre 30 et 55 ans) établies à partir des taux de décès simulés.

Figure 29 - Distribution des espérances de vie partielles (UEMOA)



Dans la continuité des observations et des conclusions des paragraphes 3.3.1 et 3.3.2, il apparaît sur le graphique ci-dessus que la distribution de l'espérance de vie partielle issue de la simulation directe des taux bruts présente une queue moins épaisse que celle issue de la simulation des résidus. En d'autres termes, les tables de mortalité issues des taux simulés à partir de la simulation directe des taux bruts présentent une volatilité plus faible (au titre du risque d'estimation) que les tables issues des taux simulés à partir de la simulation des résidus.

4. Mesure du risque d'estimation sur les provisions

Pour étudier le risque systématique au titre des fluctuations d'échantillonnage (risque d'estimation) sur les tables de mortalité, on se place désormais dans un contexte d'assurance et on choisit le montant des provisions comme variable d'intérêt.

On considère plus précisément des provisions relatives à des engagements temporaires au décès (la durée de l'engagement est de d années, avec $d \geq 1$) et évaluées à partir des taux de décès simulés (l'aléa sur les taux de décès utilisé ici est établi à partir de la simulation directe des taux de décès bruts, car elle permet une meilleure mesure du risque lié aux fluctuations d'échantillonnage, cf. paragraphe 3.3.2). À cet effet, on travaille d'une part avec un montant d'engagement déterministe (dans ce cas, les taux de décès sont les seules sources d'aléas) et d'autre part avec un montant d'engagement aléatoire (ici, les sources d'aléas sont les taux de décès et les durées de vie restantes des assurés, représentant respectivement le risque systématique et le risque mutualisable).

4.1. Provisions déterministes

Lorsque les provisions sont déterministes, l'assureur estime en 0 les flux probables des prestations à payer en t pour un individu d'âge x par (pour la réalisation k des taux de décès simulés) :

$$F_x^k(t) = \begin{cases} C \times q_x(\hat{\theta}^k), & t = 0 \\ C \times q_{x+t}(\hat{\theta}^k) \times \prod_{j=0}^{t-1} (1 - q_{x+j}(\hat{\theta}^k)), & 1 \leq t < d \end{cases}$$

où C désigne le montant de la prestation.

En faisant l'hypothèse de décès en milieu d'année, on en déduit que le montant de la provision pour la réalisation k est :

$$L_0^k = \sum_{t=0}^{d-1} F_x^k(t) \times (1 + r_{t+1})^{-t - \frac{1}{2}},$$

où r_t représente le taux d'actualisation des flux de l'échéance t .

L'impact du risque d'estimation sur les provisions peut alors être mesuré par le coefficient :

$$c(\Upsilon) = \frac{\Upsilon}{L_0},$$

où L_0 correspond à la provision calculée à partir des taux de décès ajustés

$q_x(\hat{\theta})$ et $\Upsilon = \sqrt{E \left[\left(L_0^k - L_0 \right)^2 \right]}$. Ce coefficient permet d'avoir en pourcentage une mesure de dispersion (au titre du risque d'estimation) autour de la provision calculée à partir des taux de décès ajustés.

On note enfin $\bar{L}_0 = \frac{1}{K} \sum_{j=1}^K L_0^j$, la moyenne des provisions déterministes

calculées avec les taux de décès simulés.

4.2. Provisions stochastiques

Les provisions stochastiques visent à introduire un aléa sur la date de survenance du sinistre (risque mutualisable), au-delà de l'aléa sur le taux de décès (risque systématique). En pratique, cette approche nécessite de simuler des durées de vie restantes, et donc le moment du sinistre, selon la loi des taux de décès ajustés (la méthode retenue ici pour quantifier le risque mutualisable est une méthode classique s'appuyant sur des méthodes de simulations ; lorsque de nombreux facteurs de risques sont à prendre en compte, cette méthode peut toutefois devenir délicate à intégrer et on peut dans ce cas retenir par exemple l'approche « semi-analytique » proposée dans Guibert [2010]). Aussi, avant de présenter les outils de mesure du risque d'estimation dans ce contexte (cf. 4.2.3), on présente la méthode de simulation des durées de vie restantes retenue (cf. 4.2.1) et la méthode de calcul des provisions stochastiques dans un cas simplifié (c'est-à-dire dans un cas où l'on considère dans un premier temps les taux de décès ajustés et non simulés, cf. 4.2.2).

4.2.1. Simulation des durées de vie restantes

La méthode retenue pour la simulation des durées de vie restantes dans le cas du modèle de Brass est inspirée de l'approche retenue dans Delwarde et Denuit [2006] dans le cas du modèle d'Heligman-Pollard.

On génère ainsi pour différents scénarios une variable aléatoire V de loi uniforme sur $[0, 1]$ et on ne considère que la partie entière de la durée de vie restante à l'âge x , notée T_x . La variable T_x est ainsi un entier qui satisfait :

$$T_x p_x \geq V > T_{x+1} p_x,$$

soit :

$$\prod_{j=0}^{T_x-1} p_{x+j} \geq V > \prod_{j=0}^{T_x} p_{x+j},$$

c'est-à-dire que T_x est la plus grande durée entière précédant le moment du décès.

4.2.2. Provisions stochastiques et taux de décès ajustés

Dans un premier temps, on évalue les provisions stochastiques en considérant les taux de décès ajustés (il n'y a ainsi aucune illustration du risque d'estimation ici), notés pour mémoire $q_x(\hat{\theta})$ (les probabilités de survie utilisées pour simuler les durées de vie restantes sont donc $p_x(\hat{\theta}) = 1 - q_x(\hat{\theta})$).

La démarche retenue ici est comparable à celle présentée dans Planchet et al. [2006].

On considère désormais que le flux aléatoire de prestation à payer en t est :

$$\tilde{F}_x(t) = \begin{cases} C, & T_x = t < d \\ 0, & \text{sinon} \end{cases},$$

en supposant $E[\tilde{F}_x(t)] = F_x(t)$. Le montant aléatoire de l'engagement en 0 s'écrit alors (par souci de comparabilité avec la provision déterministe, il est supposé que les décès sont observés en milieu d'année) :

$$\Lambda = \tilde{F}_x(t) \times \left(1 + r_{T_x+1}\right)^{-T_x - 1/2},$$

et on suppose que $E[\Lambda] = L_0$. Sur la base de g ($g \in [1, G]$) réalisations de T_x , on obtient alors des réalisations $\lambda_1, \lambda_2, \dots, \lambda_G$ de Λ .

On note \hat{L}_0 , la moyenne des provisions stochastiques calculées sur ces bases (on a ainsi $\hat{L}_0 = \bar{\lambda} = \frac{1}{G} \sum_{i=1}^G \lambda_i$).

4.2.3. Provisions stochastiques et taux de décès simulés

Ici, on considère deux sources d'aléas : une première liée aux taux de décès stochastiques Π (Π représentant les taux de décès $q_x(\hat{\theta}^k)$ pour une réalisation k) et une seconde liée à la loi conditionnelle des engagements stochastiques $\Lambda | \Pi$. Ici également, la démarche retenue est comparable à celle présentée dans Planchet et al. [2006].

Commençons par présenter la loi conditionnelle des engagements aléatoires. Les probabilités de survie utilisées pour simuler les durées de vie restantes, notées T_x^k pour une réalisation k des taux de décès, sont $p_x(\hat{\theta}^k) = 1 - q_x(\hat{\theta}^k)$. En pratique, le flux aléatoire de prestation à payer en t pour une simulation k est donc :

$$\tilde{F}_x^k(t) = \begin{cases} C, & T_x^k = t < d \\ 0, & \text{sinon} \end{cases},$$

en considérant $E\left[\tilde{F}_x^k(t)\right] = F_x^k(t)$. Pour une simulation k donnée des taux de décès ajustés, le montant aléatoire de l'engagement en 0 s'écrit alors :

$$\Lambda^k = \tilde{F}_x^k(t) \times \left(1 + r_{T_x^k+1}\right)^{-T_x^k - 1/2},$$

en supposant $E\left[\Lambda^k\right] = L_0^k$. Sur la base de g ($g \in [1, G]$) réalisations de T_x^k , on obtient alors des réalisations $\lambda_1^k, \lambda_2^k, \dots, \lambda_G^k$ de Λ^k . En outre, pour une simulation k donnée des taux de décès, on note que la moyenne empirique de la provision sur l'ensemble des réalisations de T_x^k est $\hat{L}_0^k = \bar{\lambda}^k = \frac{1}{G} \sum_{i=1}^G \lambda_i^k$.

En intégrant désormais la première source d'aléa, qui pour mémoire correspond aux fluctuations d'échantillonnage sur les taux de décès (et qui représente ainsi le risque d'estimation, non mutualisable), la provision stochastique moyenne issue des taux de décès simulés s'écrit

$$\bar{\hat{L}}_0 = \frac{1}{K} \sum_{j=1}^K \hat{L}_0^j = \frac{1}{K} \sum_{j=1}^K \bar{\lambda}^j = \frac{1}{K} \sum_{j=1}^K \left(\frac{1}{G} \sum_{i=1}^G \lambda_i^j \right).$$

4.3. Mesure sur les provisions : illustrations

Les applications numériques relatives aux mesures du risque d'estimation sur les provisions sont réalisées en considérant un contrat temporaire décès dans deux scénarios : un premier dans lequel la durée du contrat est $d = 5$, et un second dans lequel la durée est $d = 20$. Dans chaque scénario, on compare les résultats obtenus avec les provisions déterministes et stochastiques.

Dans les deux scénarios on considère un assuré d'âge $x = 31$, un capital assuré égal à $C = 1$ et la courbe d'actualisation retenue est la courbe de l'Institut des Actuaire au 31/12/2009².

Concernant le nombre de simulations, pour la provision stochastique calculée à partir des taux ajustés, on retient $G = 37,5 \cdot 10^6$ simulations des durées de vie restantes. Pour la provision stochastique calculée à partir des taux simulés (au titre du risque d'estimation), pour chaque simulation k des taux de décès on retient $G = 2\,500$ simulations des durées de vie restantes (pour mémoire, le nombre de simulations des taux de décès est $K = 15\,000$, ce qui conduit donc à un total de 37,5 millions de simulations).

² Courbe disponible sur le site de l'Institut des Actuaire : http://www.institutdesactuaire.com/gene/link.php?doc_link=../docs/2010003190313_I A20091231.xls.

4.3.1. Illustrations et comparaisons sur une garantie temporaire décès de 5 ans

Pour quantifier le risque d'estimation dans les différentes approches de calcul des provisions, on présente dans le tableau ci-dessous les résultats synthétiques des estimations et des simulations dans quatre scénarios : calcul de provisions déterministe avec les taux ajustés, calcul de provisions stochastiques avec les taux ajustés, calcul de provisions déterministes avec les taux simulés et calcul de provisions stochastiques avec les taux simulés.

Tableau 24 - Risque d'estimation sur les provisions (5 ans) (UEMOA)

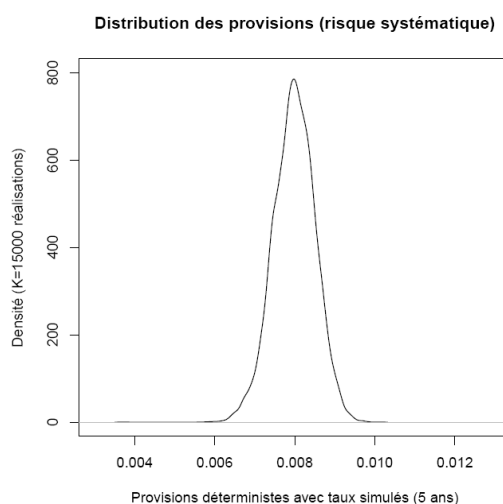
Statistique	Provision déterministe et taux ajustés	Provision stochastique et taux ajustés	Provision déterministe et taux simulés (risque d'estimation)	Provision stochastique et taux simulés (risque d'estimation)
Moyenne	$L_0 = 8,12.10^{-3}$	$\hat{L}_0 = 8,11.10^{-3}$	$\bar{L}_0 = 7,98.10^{-3}$	$\tilde{L}_0 = 7,99.10^{-3}$

Il ressort que la prise en compte du risque d'estimation impacte légèrement la provision calculée (évolutions comprises entre 1,5 % et 1,7 % lors de la prise en compte du risque systématique d'estimation).

Concernant la dispersion de la distribution de la provision déterministe calculée à partir des taux simulés (risque d'estimation), il apparaît que le coefficient est égal à $c(Y) = 6,64 \%$.

La dispersion de la distribution peut également être représentée par une estimation de la fonction de densité (par l'estimateur à noyau) de la provision déterministe calculée à partir des taux simulés.

Figure 30 – Distribution des provisions (5 ans) (UEMOA)



Cette illustration permet de disposer d'éléments additionnels sur la mesure du risque systématique d'estimation sur les provisions.

4.3.2. Illustrations et comparaisons sur une garantie temporaire décès de 20 ans

On présente ci-après les résultats obtenus dans le cadre d'une provision d'une garantie temporaire décès de 20 ans (et non plus 5 ans comme dans le paragraphe précédent). Nous obtenons les résultats synthétiques présentés dans les tableaux suivants.

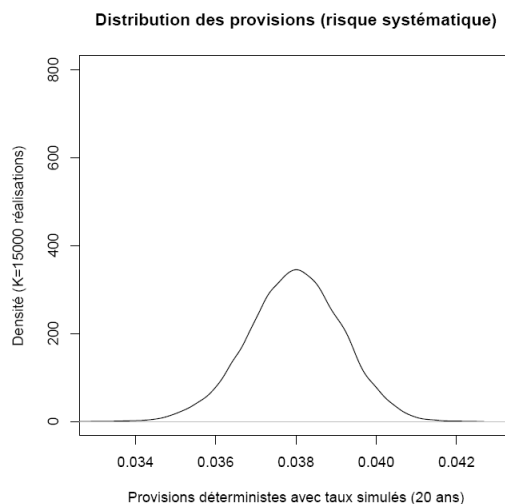
Tableau 25 - Risque d'estimation sur les provisions (20 ans) (UEMOA)

Statistique	Provision déterministe et taux ajustés	Provision stochastique et taux ajustés	Provision déterministe et taux simulés (risque d'estimation)	Provision stochastique et taux simulés (risque d'estimation)
Moyenne	$L_0 = 3,83.10^{-2}$	$\hat{L}_0 = 3,84.10^{-2}$	$\bar{L}_0 = 3,80.10^{-2}$	$\tilde{L}_0 = 3,80.10^{-2}$

Il ressort lorsque la durée de l'engagement est de 20 ans que la prise en compte du risque systématique a un impact limité sur la provision calculée (évolutions comprises entre 0,8 et 1,0 % lors de la prise en compte du risque systématique), et plus faible que celui observé lorsque la durée de l'engagement est de 5 ans.

Concernant la dispersion de la distribution de la provision déterministe calculée à partir des taux simulés, il apparaît que le coefficient est égal $c(\Upsilon) = 3,18 \%$. La dispersion obtenue ici pour la provision de 20 ans apparaît ainsi deux fois plus faible que celle observée pour la provision de 5 ans.

En complément, on présente ci-après une estimation de la fonction de densité de la provision déterministe calculée à partir des taux simulés.

Figure 31 - Distribution des provisions (20 ans)

L'échelle en ordonnée de ce graphique est identique à celle de la Figure 30, et on note que le montant moyen de la provision est plus de quatre fois supérieur à celui associé à la Figure 30.

5. Évolution du risque d'estimation en présence d'hétérogénéité

Les travaux et illustrations des sections précédentes ont été réalisés à partir des données de toute la population de la sous-région UEMOA. Ces données incluent les données de trois pays, présentées dans le tableau ci-dessous (il s'agit de données sur la plage d'âge 30-55 ans).

Tableau 26 - Statistiques ventilés par pays (UEMOA)

Homme (pop. ass.)	Exposition	Âge moyen	Taux décès moyen	Taux décès moyen (borne inf. à 95%)	Taux décès moyen (borne sup. à 95%)
CI	549 656	43,9 ans	0,40%	0,38%	0,41%
ML	12 114	42,5 ans	0,22%	0,14%	0,31%
TG	133 779	43,2 ans	0,42%	0,39%	0,46%
UEMOA (CI-ML-TG)	695 549	43,8 ans	0,40%	0,38%	0,41%

Dans le cadre d'une modélisation de la mortalité d'expérience, il est toutefois nécessaire de tenir compte de l'éventuelle hétérogénéité au sein de la population, afin de limiter le biais associé lorsque la structure de la population évolue (comme indiqué en introduction de ce chapitre). En pratique ici, il s'agit ainsi de tenir compte des différences de mortalité entre pays. À cet effet, dans le cadre de ce chapitre on se limite à une démarche simple qui consiste à modéliser le comportement de chaque pays de manière indépendante.

L'objectif de cette section 5 est alors de présenter l'évolution du risque systématique d'estimation lorsque l'on passe de la population UEMOA à celle

d'un pays. Le pays retenu à cet effet est le Togo (TG), pays intermédiaire entre la Côte d'Ivoire (CI) et le Mali (ML) en termes d'exposition au risque.

Dans ce contexte, dans un premier temps on teste et on mesure l'hétérogénéité au sein de la population UEMOA (*cf.* sous-section 5.1). Dans un second temps, on segmente cette population pour tenir compte de l'hétérogénéité à partir d'un modèle indépendant pour chaque sous-population, puis on déduit l'évolution du risque d'estimation au niveau des taux de décès et des provisions (*cf.* respectivement sous-sections 5.2 et 5.3). On note que dans les illustrations présentées dans ces sous-sections 5.2 et 5.3, les méthodes statistiques utilisées sont celles présentées et utilisées dans les sections précédentes (en particulier, la méthode de simulation des taux bruts retenue est la simulation directe). En outre, toutes ces illustrations présentent une comparaison des résultats obtenus à partir des données du Togo et à partir de celles de l'UEMOA (dans ce dernier cas, les illustrations sont alors identiques à celles présentées dans les sections précédentes).

5.1. Identification et mesure de l'hétérogénéité

Pour apprécier les différences de niveau dans la mortalité de la population assurée, deux niveaux d'analyses statistiques sont retenus : un premier dans lequel on teste l'identité des distributions de survie pour tous les couples de pays, et un second dans lequel on mesure les différences de mortalité entre pays. Cette sous-section présente les outils utilisés et les résultats obtenus dans ce cadre.

5.1.1. Présentation des outils pour tester et mesurer l'hétérogénéité

L'appréciation de l'hétérogénéité au sein de la population est donc réalisée en deux temps : un premier dans lequel on la teste et un second dans lequel on la mesure, le cas échéant. Pour ce faire, plusieurs outils sont disponibles.

Pour le premier point, on travaille ici avec des informations ligne à ligne et on peut alors s'appuyer sur des tests non paramétriques comparant les fonctions de survie de différentes sous-populations à partir de données individuelles. En pratique, ces tests s'appuient sur l'apparition ou non d'un décès et sur l'exposition au risque de chaque groupe au fil des âges. Les tests de référence sont alors les tests du log-rank et de Gehan (qui représentent des adaptations au cas des données censurées des tests de rang, respectivement les tests de Savage et de Wilcoxon, *cf.* par exemple Hill et al. [1990], Dreesbeke et al. [1989] ou encore Planchet et Thérond [2006]). Lorsque l'on compare plus de deux sous-populations, les tests de référence sont les tests de Tarone et Ware et de Breslow (qui représentent respectivement les généralisations des tests du log-rank et de Gehan, *cf.* Dreesbeke et al. [1989]). Dans cette étude, on compare les sous-populations deux à deux et on utilise à cet effet le test du log-rank.

Concernant la mesure de l'hétérogénéité entre sous-populations, elle est réalisée à partir d'une estimation du risque relatif de décès instantané pour le test du log-rank (*cf.* Hill et al. [1990]). Cette approche permet d'obtenir des coefficients représentant les écarts relatifs de niveau de mortalité entre deux sous-populations.

Identification de l'hétérogénéité : présentation du test du log-rank

La méthode retenue pour le test de comparaison des sous-populations assurées est donc la méthode du log-rank, sous l'hypothèse H0 d'identité dans la distribution de survie. Dans ce test, pour lequel on suppose, sous H0, que les décès suivent une loi hypergéométrique (*cf.* par exemple Hill et al. [1990] pour une présentation détaillée du principe), on considère la statistique :

$$\frac{\left(D_j^{obs} - D_j^{th}\right)^2}{\sigma^2},$$

qui suit asymptotiquement un $\chi^2(1)$, avec j une des deux sous-populations de

la comparaison (j est égal à 1 ou à 2), $D_j^{obs} = \sum_{t_i=1}^T d_{ij}$, $D_j^{th} = \sum_{t_i=1}^T d_i \frac{r_{ij}}{r_i}$ et

$$\sigma^2 = \sum_{t_i=1}^T d_i \frac{(r_i - d_i)}{(r_i - 1)} \frac{r_{i1} r_{i2}}{r_i^2} \quad (T \text{ représente le nombre d'âges sous observation, } r_i$$

et d_i représentent respectivement l'exposition au risque et le nombre de décès en t_i au titre des deux sous-populations 1 et 2, r_{ij} et d_{ij} représentent respectivement l'exposition au risque et le nombre de décès en t_i au titre de la sous-population j uniquement).

En pratique, on utilise la forme approchée du test du log-rank, dont la statistique est :

$$\frac{\left(D_1^{obs} - D_1^{th}\right)^2}{D_1^{th}} + \frac{\left(D_2^{obs} - D_2^{th}\right)^2}{D_2^{th}}.$$

Aussi, lorsque cette statistique approchée est supérieure à la statistique de Khi-deux (à un degré de liberté) au seuil de 5 %, soit à $\chi_{\alpha}^2(1) = 3,841$, on rejette H0. Pour mémoire, cette statistique approchée est toujours inférieure à celle du log-rank exacte (*cf.* Hill et al. [1990]). Aussi, à chaque fois que cette statistique conduit à rejeter H0, la statistique exacte du log-rank aurait *a fortiori* conduit à la même conclusion.

Mesure de l'hétérogénéité : présentation du risque relatif (RR) associé au test du log-rank

Afin de quantifier la différence de mortalité entre deux sous-populations, on s'appuie sur une estimation du rapport des risques instantanés de décès. L'estimateur retenu est l'estimateur de Tarone [1981] (également nommé estimateur de Mantel-Haenszel modifié), dont l'emploi est recommandé en présence d'ex-æquo (c'est-à-dire lorsque l'on suppose qu'il se produit plus

d'un décès à chaque instant t_i); sur ce point on pourra consulter Hill et al. [1990].

Sur une base 100, cet estimateur est déterminé par (pour deux sous-populations 1 et 2) :

$$100 \times \frac{\sum_{t_i=1}^T d_{i2} \frac{r_{i1}}{r_i}}{\sum_{t_i=1}^T d_{i1} \frac{r_{i2}}{r_i}},$$

avec les mêmes notations que celles du test du log-rank. L'estimateur de Tarone [1981] considère que le nombre de décès à chaque instant suit une loi de Poisson.

5.1.2. Présentation des résultats sur l'analyse de l'hétérogénéité

Sur la base des outils ci-dessus, la matrice suivante présente les risques relatifs (RR) entre couples de pays de la zone UEMOA (représentée ici par la Côte d'Ivoire, le Mali et le Togo). Les combinaisons pour lesquelles on accepte l'hypothèse d'identité des distributions de survie au seuil de 5 % sont en gris foncé et en gras.

Tableau 27 – RR et tests de comparaison par pays (UEMOA)

Homme (pop. ass.)	CI	ML	TG
CI	100	61	110
ML		100	181
TG			100

De manière générale, il apparaît que les résultats ci-dessus sont cohérents avec les statistiques descriptives présentées dans le Tableau 26.

Concernant les tests d'hétérogénéité, il ressort que pour tous les couples de la zone UEMOA, les distributions sont significativement différentes entre elles pour les populations assurées masculines.

Par ailleurs, concernant les mesures de l'hétérogénéité, il apparaît pour la population assurée que :

- le risque de décès du Togo majore significativement de 10 % celui de la Côte d'Ivoire, et de 81 % celui du Mali ;
- le risque de décès du Mali minore significativement de 39 % celui de la Côte d'Ivoire.

Au final, il apparaît donc bien que la population assurée de l'UEMOA est hétérogène, et la prise en compte d'une segmentation de la population est donc nécessaire.

5.2. Évolution du risque d'estimation sur les taux de décès ajustés

En pratique, il s'agit ici d'étudier l'évolution du risque d'estimation sur les taux de décès et sur les espérances de vie partielles (entre 30 et 55 ans) lorsque l'on souhaite segmenter la population UEMOA pour tenir compte des différences par pays (pour mémoire, le pays retenu pour les illustrations est le Togo).

5.2.1. Comparaison des taux de décès uniquement

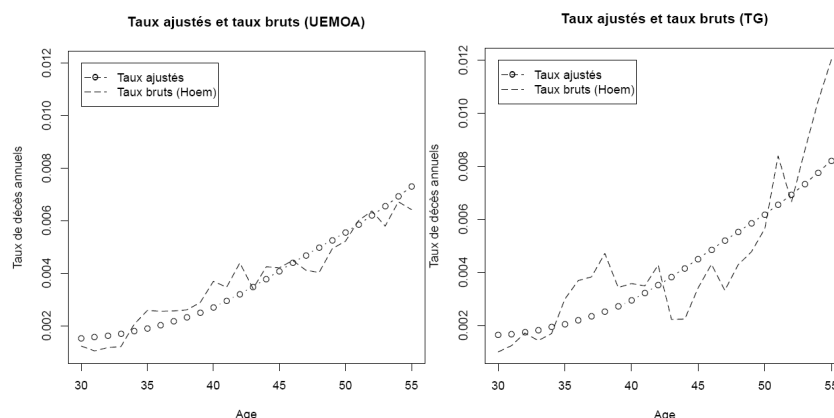
Dans un premier temps, on décrit les statistiques associées à l'ajustement des taux de décès.

Dans le cas des données de l'UEMOA, les statistiques relatives à l'ajustement sont présentées dans la section 3.3.

Dans le cas des données du Togo, l'ajustement conduit aux estimations suivantes : $\hat{a} = 0,824$ et $\hat{b} = -0,841$ (avec des probabilités de rejeter à tort l'hypothèse H_0 de nullité du paramètre selon le test de Student respectivement égales à $8,6 \cdot 10^{-8}$ et à 19,9 %). Aussi, il ressort que le R^2 ajusté est égal à 69,2 %. Enfin, on note que la probabilité associée au test Shapiro-Wilk sur les résidus est égale à 71,7 % ; en d'autres termes, on accepte l'hypothèse H_0 de normalité des résidus observés. De manière générale, il apparaît que ces statistiques sont moins satisfaisantes que celles issues de l'ajustement avec les données de l'UEMOA.

Sur ces bases, les graphiques ci-dessous présentent les taux ajustés et les taux bruts pour les données de l'UEMOA et du Togo.

Figure 32 - Taux de décès ajustés et taux de décès bruts (UEMOA et TG)

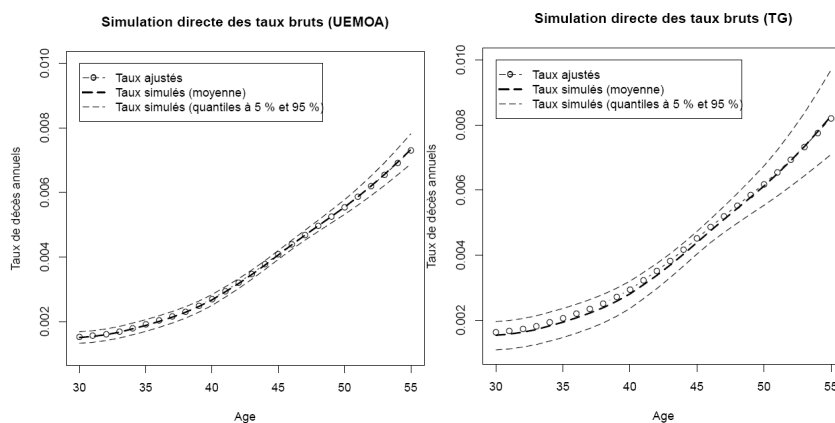


Il apparaît sur ces graphiques que les fluctuations d'échantillonnage sont logiquement plus importantes avec les données du Togo.

Pour confirmer ce premier résultat, les graphiques suivants présentent les taux de décès ajustés issus des estimations $\hat{\theta} = (\hat{a}, \hat{b})$, ainsi que, pour les deux populations présentées, les taux de décès simulés au titre des fluctuations

d'échantillonnage (dans ce cas, on présente les moyennes et les quantiles à 5 % et à 95 %).

Figure 33 - Taux de décès ajustés et simulés (UEMOA et TG)



Il ressort que la dispersion des taux de décès simulés à partir des données du Togo est plus importante que celle des taux simulés à partir des données de l'UEMOA.

Pour quantifier l'écart de dispersion, le tableau ci-dessous présente une moyenne arithmétique du coefficient $c(\psi_x)$ pour tous les âges $x \in [x_m, x_M]$.

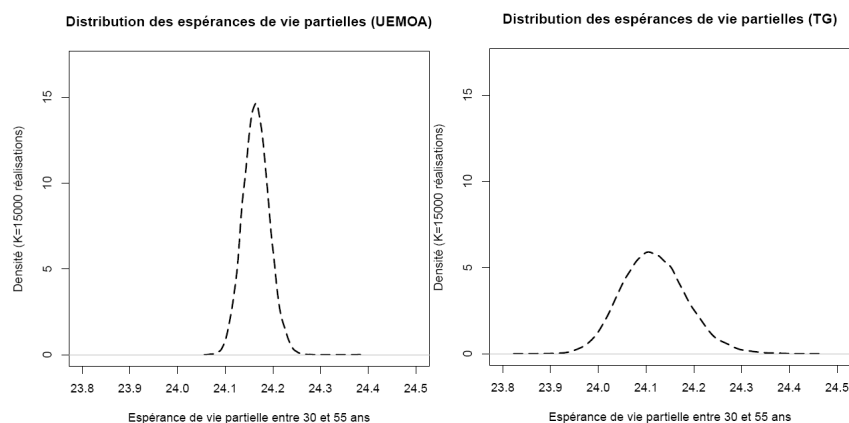
Tableau 28 - Risque d'estimation sur les taux (moyenne) (UEMOA et TG)

Population	Mesure du risque d'estimation (en % des taux ajustés)
UEMOA (CI-ML-TG)	$c(\psi) = 4,07 \%$
TG	$c(\psi) = 10,03 \%$

La moyenne du coefficient $c(\psi_x)$, qui pour mémoire représente la dispersion en pourcentage des taux simulés (au titre du risque d'estimation dû aux fluctuations d'échantillonnage) autour des taux ajustés, augmente de près de 150 %.

5.2.2. Comparaison des espérances de vie partielles

Outre la présentation et la comparaison des taux de décès, il est nécessaire de comparer les tables de mortalité associées, et à cet effet on utilise les espérances de vie partielles. Dans ce contexte, le graphique suivant présente, pour chaque population, les estimations des fonctions de densité des espérances de vie résiduelles établies à partir des taux de décès simulés.

Figure 34 - Distribution des esp. de vie partielles (UEMOA et TG)

Dans la continuité des observations et des conclusions ci-dessus, il apparaît sur ces graphiques que la distribution de l'espérance de vie partielle issue de la population UEMOA présente une queue moins épaisse que celle issue de la population du Togo. En d'autres termes, les tables de mortalité issues des taux simulés à partir des données UEMOA présentent une volatilité plus faible que celles issues des taux simulés à partir des données du Togo.

5.3. Évolution du risque d'estimation sur les provisions

Cette sous-section s'attache à quantifier le risque d'estimation dans les provisions. À ce titre, on présente dans le tableau ci-dessous les résultats synthétiques des estimations et des simulations dans un scénario de calcul de provisions déterministes avec les taux ajustés d'une part, et avec les taux simulés d'autre part (le calcul de provisions stochastiques n'est donc pas repris dans ce tableau).

Tableau 29 - Risque d'estimation sur provisions (20 ans) (UEMOA et TG)

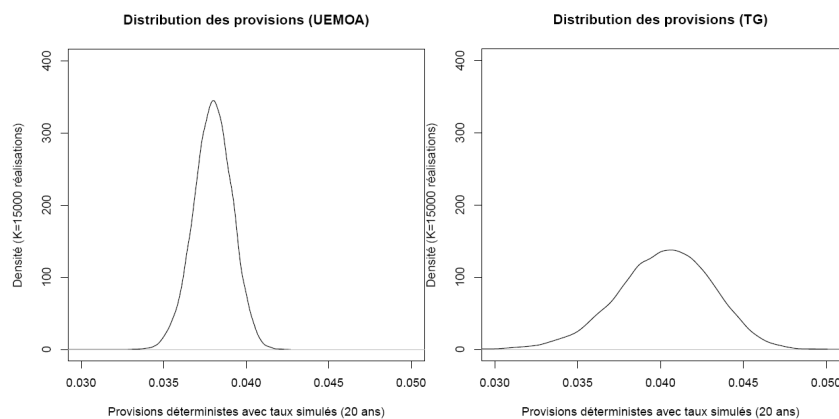
Statistique	UEMOA		TG	
	Provision déterministe et taux ajustés	Provision déterministe et taux simulés (risque d'estimation)	Provision déterministe et taux ajustés	Provision déterministe et taux simulés (risque d'estimation)
Moyenne	$L_0 = 3,83.10^{-2}$	$\bar{L}_0 = 3,80.10^{-2}$	$L_0 = 4,18.10^{-2}$	$\bar{L}_0 = 4,02.10^{-2}$
Quantile à 0,5 %	Sans objet	$3,48.10^{-2}$	Sans objet	$3,17.10^{-2}$
Quantile à 5 %	Sans objet	$3,60.10^{-2}$	Sans objet	$3,53.10^{-2}$
Quantile à 95 %	Sans objet	$3,99.10^{-2}$	Sans objet	$4,47.10^{-2}$
Quantile à 99,5 %	Sans objet	$4,08.10^{-2}$	Sans objet	$4,70.10^{-2}$
Coefficient	Sans objet	$c(Y) = 3,18 \%$	Sans objet	$c(Y) = 7,91 \%$

Il ressort qu'avec les données du Togo, la prise en compte du risque systématique impacte significativement la provision calculée (baisse de 3,8 % lors de la prise en compte du risque systématique), alors que pour mémoire l'impact est limité avec les données UEMOA (baisse de 0,8 %).

Concernant les quantiles, il ressort qu'à partir des données du Togo les quantiles à 0,5 % et à 99,5 % présentent respectivement un écart de -21 % et 17 % par rapport à la provision moyenne, contre un écart respectif de -8 % et 7 % avec les données UEMOA.

Enfin, il apparaît que le coefficient $c(Y)$ relatif à la modélisation à partir des données du Togo majore de près de 150 % celui obtenu à partir des données de l'UEMOA.

En complément au Tableau 29, on présente ci-après une estimation de la fonction de densité de la provision déterministe calculée à partir des taux simulés à partir des données de l'UEMOA et du Togo.

Figure 35 - Distribution des provisions (20 ans) (UEMOA et TG)

Ces graphiques constituent ainsi une forme d'illustration des troisième et cinquième colonnes du Tableau 29, relatives à la provision déterministe calculée à partir des taux simulés de l'UEMOA et du Togo. Ils confirment la présence d'un risque systématique d'estimation plus important dans le cas d'une modélisation à partir des données du Togo.

6. Synthèse et conclusion

Dans le cadre de travaux sur le traitement de l'hétérogénéité, cette étude permet de présenter des mesures du risque d'estimation, c'est-à-dire du risque systématique associé à l'utilisation d'une loi de durée dont on sait qu'elle n'est pas bonne du fait des erreurs d'échantillonnage affectant sa construction. En particulier, il s'agit ici de présenter l'évolution du risque d'estimation en cas de baisse de l'échantillon d'expérience au titre d'une prise en compte de l'hétérogénéité à partir d'une modélisation indépendante de chaque sous-population. Cette analyse est réalisée en quatre étapes.

Les trois premières étapes introduisent et présentent les outils de mesure de risque d'estimation retenus. Ainsi, dans un premier temps les intervalles et bandes de confiance au titre du risque d'échantillonnage sont présentés pour les taux de décès bruts et pour la fonction de survie. Ces outils, présentés pour les estimateurs de Hoem (taux bruts) et de Kaplan-Meier (fonction de survie), constituent une première mesure du risque lié aux erreurs d'échantillonnage.

Dans un deuxième temps, les analyses portent sur les taux de décès ajustés à partir du modèle de Brass (modèle à référence externe). À ce niveau, deux approches ont été testées pour simuler l'impact des erreurs d'échantillonnage sur les taux de décès ajustés. La première porte sur la simulation directe des taux de décès bruts (desquels on déduit les paramètres du modèle de Brass pour chaque réalisation des taux bruts), et la seconde porte sur la simulation indirecte des taux bruts puisqu'elle consiste à simuler des résidus du modèle ajusté (desquels on déduit des réalisations de « taux de décès bruts », puis des paramètres du modèle de Brass pour chaque réalisation de ces « taux bruts »). Il apparaît ainsi que la seconde méthode inclut un risque lié à la spécification au modèle, au-delà du seul risque d'estimation lié aux fluctuations

d'échantillonnage, qui se traduit par une mesure de risque plus importante, tant au niveau des taux de décès qu'au niveau des tables de mortalité (dans ce dernier cas, l'espérance de vie partielle a été utilisée comme fonctionnelle). Cette seconde méthode présente toutefois l'avantage de ne s'appuyer sur aucune loi pour les taux bruts et peut ainsi se révéler particulièrement intéressante lorsque la loi des taux bruts n'est pas connue.

Dans un troisième temps, les analyses portent sur le calcul de provision (fonctionnelle naturelle dans le cas de l'assurance), dans le cadre d'une garantie temporaire décès. Cette partie permet de présenter certains résultats évidents, tels que l'évolution de la moyenne des provisions en cas de prise en compte du risque systématique (pour mémoire, la moyenne des provisions est en revanche inchangée en cas de prise en compte du risque mutualisable) ou la réduction (en pourcentage des provisions) du risque systématique lorsque la durée de couverture augmente.

Les trois étapes ci-dessus permettent ainsi de présenter des outils de mesure du risque d'estimation.

Toutes les applications numériques associées à ces trois étapes ont été réalisées à partir des données d'expérience de la zone UEMOA, qui incluent les données d'expérience de la Côte d'Ivoire, du Togo et du Mali. En pratique toutefois, il apparaît que cette population UEMOA est hétérogène (les tests statistiques effectués confirment en effet que les différences de mortalité par pays sont significatives). Il convient alors de tenir compte de cette hétérogénéité dans la modélisation de la mortalité d'expérience. Aussi, la quatrième étape de ce travail s'attache à tenir compte de cette hétérogénéité, à partir d'une modélisation indépendante de chaque pays, et à en présenter les conséquences sur l'évolution du risque d'estimation. À cet effet, on utilise les données du Togo, pays dont l'exposition au risque représente 20 % de l'exposition de la zone UEMOA. Il apparaît alors que l'impact du risque d'estimation sur la provision *best estimate* est plus de 4 fois plus important avec les données du Togo qu'avec les données de l'UEMOA (il est égal à 3,8 % pour le Togo, contre 0,8 % pour la zone UEMOA).

La prise en compte de l'hétérogénéité par la méthode retenue (modélisation indépendante de chaque sous-population) n'est donc pas appropriée, car elle fait apparaître un risque d'estimation significatif, qui conduit à biaiser la provision *best estimate* et à épaissir la queue de distribution des engagements (ce qui majore le capital de solvabilité requis dans le projet Solvabilité 2, le CRNHR dans le référentiel MCEV et l'ajustement pour risque dans le projet norme IFRS sur les contrats d'assurance). La segmentation du portefeuille peut donc présenter un coût significatif en cas de modélisation indépendante de chaque population. Pour affiner la segmentation et tenir compte de l'hétérogénéité, il convient donc de retenir une approche plus adaptée. En pratique, il apparaîtra dans le chapitre suivant que la prise en compte d'un modèle intégrant l'hétérogénéité à partir de facteurs observables est une solution envisageable.

Notations des taux de décès

Les principales notations retenues ici pour les taux de décès annuels à l'âge x sont les suivantes :

- q_x : taux de décès réel théorique ;
- Q_x : variable aléatoire des taux de décès observés (taux bruts) ;
- \hat{q}_x : réalisation des taux de décès observés (taux bruts) ;
- q_x^{ref} : taux de décès (ajustés) issus de la table de référence externe retenue ;
- $q_x(\hat{\theta})$: taux de décès ajustés à partir des taux observés \hat{q}_x ;
- q_x^k : réalisation k des taux de décès bruts (réalisation à partir de la simulation directe de la variable aléatoire Q_x ou à partir de la simulation des résidus, *cf.* respectivement 3.1 et 3.2) ;
- $q_x(\hat{\theta}^k)$: taux de décès simulés à partir de la réalisation des taux bruts q_x^k (réalisation à partir de la simulation directe) ;
- $q_x^k(\hat{\theta}^k)$: taux de décès simulés à partir de la réalisation des taux bruts q_x^k (réalisation à partir de la simulation des résidus).

Bibliographie

- Alho J. M. [2007] « Méthodes utilisées pour établir les projections relatives à la mortalité – Distributions prédictives de la mortalité future », *ISSA, Quinzième Conférence internationale des actuaires et statisticiens de la sécurité sociale*.
- Bourbonnais R. [2004] *Économétrie*, Dunod, 5^e édition.
- Delwarde A., Denuit M. [2006] « Construction de tables de mortalité périodiques et prospectives », *Economica*.
- Droesbeke J.-J., Fichet B., Tassi P. [1989] (éditeurs) « Analyse statistique des durées de vie », *Economica*.
- Guibert Q. [2010] *Analyse de la solvabilité d'un régime de retraite supplémentaire*, Mémoire d'actuariat, ISFA.
- Hall W. J., Wellner J. A. [1980], « Confidence bands for a survival curve from censored data », *Biometrika*, Vol. 67, No. 1.
- Hill C., Com-Nougé C., Kramar A., Moreau T., O'Quigley J., Senoussi R., Chastang C. [1990] « Analyse statistique des données de survie », *Inserm, Médecine – Sciences, Flammarion*.
- Kamega A., Planchet F. [2010], « Mesure du risque d'estimation associé à une table d'expérience », *Cahiers de recherche de l'ISFA, WP2136*.
- Klein J. P., Moeschberger M. L. [2005] « Survival Analysis – Techniques for Censored and Truncated Data », *Springer, 2nd edition*.
- Martinussen T., Scheike T. H. [2006] « Dynamic Regression Models for Survival Data », *Springer*.
- Meucci A. [2007] « Risk and Asset Allocation », *Springer, Reprint of the 2007 edition*.
- Nair V. N. [1984], « Confidence Bands for Survival Functions With Censored Data: A Comparative Study », *Technometrics*, Vol. 26, No. 3.
- Planchet F. [2009] « Provisionnement *best estimate* et risque arrêt de travail », *La Tribune de l'Assurance, n. 140*.
- Planchet F., Juillard M., Faucillon L. [2006] « Quantification du risque systématique de mortalité pour un régime de rentes en cours de service », *Assurance et gestion des risques, Vol. 75*.
- Planchet F., Leroy G. [2009] « Quel niveau de segmentation pertinent ? », *La Tribune de l'Assurance, n. 142*.
- Planchet F., Thérond P. [2006] « Modèles de Durée - Applications actuarielles », *Economica*.
- Rakotomalala R. [2008] « Tests de normalité – Techniques empiriques et tests statistiques », *Université Lumière Lyon 2, Version 2.0*.

Saporta G. [2006] « Probabilités, analyses des données et statistique », *Technip*,
2^{ème} édition.

Tarone R. E. [1981] « On Summary Estimators of Relative Risk », *Journal of
Chronic Diseases*, Vol. 34.

Chapitre 4

Réduction du risque d'estimation en présence d'hétérogénéité¹

Sommaire

CHAPITRE 4 RÉDUCTION DU RISQUE D'ESTIMATION EN PRÉSENCE D'HÉTÉROGÉNÉITÉ	127
1. Introduction.....	128
2. Choix du modèle à variables explicatives observables.....	130
2.1.Modèles à variables explicatives et problèmes de dimension.....	130
2.2.Choix de modèles single-index (SIM) : Cox et Lin et Ying	131
3. Modèle de Cox : ajustement et simulation des taux de décès.....	133
3.1.Cox : évaluation des taux de décès annuels ajustés	133
3.2.Cox : évaluation des taux de décès annuels simulés	15
4. Modèle de Lin et Ying : ajustement et simulation des taux de décès	16
4.1.Lin et Ying : évaluation des taux de décès annuels ajustés.....	16
4.2.Lin et Ying : évaluation des taux de décès annuels simulés	19
5. Comparaison des ajustements des modèles de Cox et de Lin et Ying et <i>backtesting</i> sur l'hétérogénéité.....	141
5.1.Comparaison des ajustements de Cox et de Lin et Ying	141
5.2.Backtesting sur l'hétérogénéité (comparaison des modèles d'hétérogénéité)	145
6. Comparaison des risques d'estimation entre les modèles indépendants pour chaque sous-population et ceux intégrant l'hétérogénéité à partir de facteurs observables	148
6.1.Mesure du risque d'estimation sur les taux ajustés	149
6.2.Mesure du risque d'estimation sur les tables	150
7. Synthèse et conclusion	153
Bibliographie.....	156

¹ Ce chapitre s'appuie principalement sur l'article Kamega et Planchet [2011], relatif à la mesure du risque d'estimation dans le cadre de la prise en compte de l'hétérogénéité à partir d'une modélisation intégrant des facteurs observables.

Ce chapitre présente une mesure du risque d'estimation pour deux exemples de modèles intégrant l'hétérogénéité à partir de variables explicatives observables. En particulier, il s'agit ici de montrer que le choix de tels modèles pour prendre en compte l'hétérogénéité permet de limiter le niveau du risque d'estimation.

1. Introduction²

Comme le rappelle Vaupel [2002], toutes les populations sont potentiellement hétérogènes : deux individus de même âge et de même sexe dans une population peuvent présenter deux risques de décès très différents.

En pratique, il apparaît que la modélisation du risque de décès fournit des résultats différents au niveau agrégé (c'est-à-dire en considérant la population au global) et désagrégé (c'est-à-dire en segmentant la population en sous-populations pour tenir compte de l'hétérogénéité), ce qui traduit un biais d'hétérogénéité (*cf.* par exemple Dreesbeke et al. [1989]). Ce résultat est souvent expliqué par le phénomène « mobile-stable », selon lequel les individus du segment avec un taux de sortie élevé sortent en premier et augmentent ainsi, lorsque le temps s'écoule, la proportion d'individus du segment avec un taux de sortie faible. Ainsi, on comprend que lorsque la répartition de la population par segment reste stable dans le temps, on peut faire l'économie d'une modélisation de l'hétérogénéité. En revanche, dès que la répartition de la population évolue, comme souvent lors de l'évaluation d'une provision technique, une prise en compte de l'hétérogénéité est indispensable, afin de disposer d'une loi d'expérience robuste.

Une réflexion s'impose alors sur la démarche à retenir pour modéliser les lois de survie d'une population hétérogène, compte tenu notamment des problématiques de choix de segmentation optimale (*cf.* Planchet et Leroy [2009]) et de risque d'estimation (*cf.* Kamega et Planchet [2010]). On cite ici trois approches possibles.

La première démarche consiste à modéliser le comportement de chaque sous-population de manière indépendante. Les modèles de durée associés à cette démarche peuvent toutefois rapidement se heurter à des problèmes d'insuffisance de données, qui accentuent les problématiques de choix de segmentation optimale et majorent le risque d'estimation (*cf.* chapitre précédent).

La deuxième démarche consiste à se tourner vers des modèles de durée intégrant des facteurs d'hétérogénéité observables à partir de variables explicatives. Ici, la population est considérée dans son ensemble et on s'attache à mesurer l'effet des variables explicatives (qui définissent les segments) sur le phénomène observé (en l'occurrence la fonction de hasard).

La troisième démarche consiste à se tourner vers des modèles de durée intégrant des facteurs d'hétérogénéité inobservables (ou résiduels, *cf.* par exemple Delwarde et Denuit [2006]). À cet effet, on peut s'appuyer sur des modèles de fragilité, qui permettent de rendre compte de l'hétérogénéité dans

² *Cf.* également conclusion du chapitre précédent.

les risques de décès individuels. En pratique, on distingue le modèle de fragilité classique de Vaupel, qui considère les différences de niveau de mortalité par individu et s'appuie sur l'hypothèse de proportionnalité, et le modèle de fragilité combiné de Barbi, qui considère les différences de niveau par individu et les différences de pente par groupe d'individus et ne s'appuie donc pas sur l'hypothèse de proportionnalité (ces modèles sont présentés dans Vaupel et al. [1979] et dans Barbi et al. [2003]).

Dans ce chapitre on s'intéresse aux modèles de la deuxième catégorie, qui intègrent des facteurs d'hétérogénéité observables à partir de variables explicatives. En particulier, le travail présenté ici permet d'apprécier l'évolution du risque d'estimation lors du passage de la première à la seconde démarche.

Plan et données

L'étude s'attache alors à justifier le choix du modèle d'hétérogénéité retenu (section 2) et à mesurer le risque d'estimation associé (sections 3 à 6).

Les illustrations numériques présentées s'appuient sur les données d'expérience utilisées pour la construction des tables de mortalité réglementaires dans la zone CIMA/FANAF pour les contrats d'assurance en cas de vie et en cas de décès. Ces données couvrent plus précisément la population des pays de la sous-région UEMOA, représentés ici par la Côte d'Ivoire, le Mali et le Togo.

On considère une hétérogénéité par pays, au-delà des différences liées au sexe et à l'âge. Aussi, les tableaux suivants présentent des statistiques (exposition, âge moyen et taux de décès annuel moyen) sur les données relatives aux populations assurées par pays pour les femmes et les hommes de la sous-région UEMOA et âgées de 30 à 55 ans.

Tableau 30 - Statistiques ventilées par pays (UEMOA - Femme)

Femme (pop. ass.)	Exposition	Âge moyen	Taux décès moyen	Taux décès moyen (borne inf. à 95%)	Taux décès moyen (borne sup. à 95%)
CI	117 199	43,2 ans	0,19%	0,17%	0,22%
ML	3 499	41,7 ans	0,11%	0,00%	0,23%
TG	22 882	42,2 ans	0,07%	0,04%	0,11%
UEMOA (CI-ML-TG)	143 580	43,0 ans	0,17%	0,15%	0,19%

Tableau 31 - Statistiques ventilées par pays (UEMOA - Homme)

Homme (pop. ass.)	Exposition	Âge moyen	Taux décès moyen	Taux décès moyen (borne inf. à 95%)	Taux décès moyen (borne sup. à 95%)
CI	549 656	43,9 ans	0,40%	0,38%	0,41%
ML	12 114	42,5 ans	0,22%	0,14%	0,31%
TG	133 779	43,2 ans	0,42%	0,39%	0,46%
UEMOA (CI-ML-TG)	695 549	43,8 ans	0,40%	0,38%	0,41%

Pour mémoire, ces données ont été collectées en 2009 et sont présentées en détail dans le deuxième chapitre. On retient ici qu'elles sont observées sur les années 2003 à 2006 et comptent des troncatures à gauche (relatives aux entrées après le 01/01/2003) et des censures à droite (relatives aux sorties avant le 31/12/2006 pour un motif autre que le décès).

2. Choix du modèle à variables explicatives observables

Dans cette section, on présente les étapes du processus du choix du modèle retenu : de la présentation de la problématique de dimension associée aux modèles à variables explicatives au choix des modèles de Cox et de Lin et Ying.

2.1. Modèles à variables explicatives et problèmes de dimension

En statistique mathématiques, lorsqu'un phénomène peut être expliqué par plusieurs variables explicatives, on peut se tourner vers des régressions purement paramétriques telles que les régressions linéaires. L'avantage est que dans ce cas, on peut facilement trouver des estimateurs convergents. L'inconvénient de ces modèles est qu'ils s'appuient sur de nombreuses hypothèses pour le comportement du phénomène observé et présentent donc un risque important de ne pas être fidèles aux données d'expérience.

Une alternative consiste alors à se tourner vers des régressions non paramétriques, qui s'appuient sur un nombre limité d'hypothèses pour le comportement du phénomène observé et sont donc moins contraignants. Ces modèles présentent toutefois un inconvénient bien connu sous le terme de « fléau de la dimension » (du mathématicien Richard Bellman), relatif à l'augmentation explosive du volume de données associée à l'ajout de dimensions supplémentaires dans un espace mathématique. Selon ce fléau, les estimateurs non paramétriques d'une fonction de régression se comportent mal lorsque le nombre de variables est important (sur ce point, on pourra par exemple consulter Viallon [2006]).

Le modèle additif d'Aalen est un exemple de modèle non paramétrique. Ce modèle suppose que l'intensité d'un processus de Poisson $N(t)$ ($t \in [0, \tau], \tau < \infty$) de dimension n (n représentant le nombre d'individus sous risque) prend la forme suivante (une description complète de ce modèle est disponible dans Martinussen et Scheike [2006] et Klein et Moeschberger [2005]) :

$$\lambda(t) = Y(t) X^T(t) \beta(t),$$

où $Y(t)$ est un indicateur de risque (pour tout $i \in [1, n]$, $Y_i(t)$ est égal à 1 si l'individu i est sous risque à la date t , et est égal à 0 sinon), $\beta(t)$ représente le vecteur du coefficient de base et des coefficients des p variables et $X(t)$ représente la matrice du terme constant de base et des variables de dimension

p (la première colonne de $X(t)$ est ainsi égale à l'unité). Si ce modèle présente une grande souplesse, il peut dans certains cas être trop sophistiqué, notamment lorsque le volume de données disponible est limité, et faire l'objet d'importantes limites opérationnelles, notamment en raison du « fléau de la dimension ».

Il est donc nécessaire de réduire la dimension de ces modèles non paramétriques. À cet effet, la méthode qui est considérée ici est la méthode retenue par Lopez [2007] dans sa thèse sur la réduction de dimension en présence de données censurées : il s'agit de la régression *single index*.

2.2. Choix de modèles single-index (SIM) : Cox et Lin et Ying

Les modèles de régression *single index* (SIM) sont définis par :

$$m(z) = E(Y | Z = z) = f(\delta^T z),$$

où Y représente la variable à expliquer de dimension 1, Z représente les variables explicatives de dimension p , m représente une fonction inconnue telle que $m : \mathbb{R}^p \rightarrow \mathbb{R}$, f représente une fonction de lien inconnue telle que $f : \mathbb{R} \rightarrow \mathbb{R}$ et $\delta \in \Theta \subset \mathbb{R}^p$ est un paramètre inconnu de dimension finie.

Si f est connue, le problème devient purement paramétrique ; et d'un autre côté si δ est connu, le problème devient non paramétrique mais de dimension 1. De manière générale, les modèles SIM sont souvent présentés comme un compromis raisonnable entre une modélisation purement paramétrique et une modélisation purement non paramétrique (comme précisé par exemple dans l'ouvrage de McCullagh et Nelder [1989]).

Lopez [2007] propose donc le choix des modèles SIM pour limiter le problème de dimension et montre que l'on peut raisonnablement estimer ces modèles semi-paramétriques en présence de censure. Pour modéliser l'hétérogénéité à partir d'un modèle prenant en compte des variables explicatives, les modèles semi-paramétriques de type SIM semblent donc être adaptés. Il convient désormais de s'interroger sur le choix du modèle SIM à retenir.

En pratique, il apparaît que le modèle multiplicatif de Cox [1972] et le modèle additif de Lin et Ying [1994] sont des cas particuliers du modèle SIM, dans lesquels les hypothèses ne portent pas sur l'espérance conditionnelle mais sur le taux de hasard instantané conditionnel. En effet, le modèle multiplicatif de Cox peut s'écrire :

$$\lambda(t | Z = z) = \lambda_0(t) e^{\delta^T z},$$

où λ_0 est une fonction supposée inconnue et δ est un paramètre à estimer (on retrouve ainsi facilement l'écriture d'un modèle SIM). De même, le modèle additif de Lin et Ying peut s'écrire :

$$\lambda(t | Z = z) = \lambda_0(t) + \gamma^T z,$$

où λ_0 est une fonction supposée inconnue et γ est un paramètre à estimer (le modèle de Lin et Ying est un cas particulier du modèle d'Aalen, dans lequel γ remplace $\gamma(t)$). En outre, il apparaît que le modèle multiplicatif de Cox et le modèle additif de Lin et Ying se rejoignent lorsque que λ_0 est constant (donc invariant dans le temps) et que le terme exponentiel dans le modèle multiplicatif (soit $e^{\delta^T z}$) est remplacé par l'expression linéaire $\{1 + \delta^T z\}$; dans ce cas $\gamma = \lambda_0 \delta$.

On limite ici notre choix de modèle SIM aux deux modèles ci-dessus : le modèle de Cox et le modèle de Lin et Ying (en pratique, de nombreux modèles alternatifs auraient pu être retenus, dont des extensions au modèle multiplicatif de Cox et d'autres extensions au modèle additif d'Aalen, cf. Martinussen et Scheike [2006]).

Le modèle de Cox est le plus couramment utilisé compte tenu de ses bonnes propriétés qui ont été largement étudiées. Cependant, comme le rappellent Hill et al. [1990], dans ce modèle le rapport des risques instantanés pour deux sous-populations de caractéristiques z_1 et z_2 ne dépend que de z_1 et z_2 et pas du temps :

$$\lambda(t | Z = z_1) / \lambda(t | Z = z_2) = e^{\delta^T z_1} / e^{\delta^T z_2}.$$

Le modèle de Cox repose ainsi sur l'hypothèse de proportionnalité des taux de décès instantanés entre les différents segments, ce qui est contraignant.

Dans le modèle de Lin et Ying, c'est la différence absolue des risques instantanés pour deux sous-populations de caractéristiques z_1 et z_2 qui ne dépend que de z_1 et z_2 et pas du temps :

$$\lambda(t | Z = z_1) - \lambda(t | Z = z_2) = \gamma^T z_1 - \gamma^T z_2.$$

Cette hypothèse est également contraignante.

En pratique, on est donc amené à faire un choix entre une contrainte sur les différences relatives (hypothèse de proportionnalité) et une contrainte sur les différences absolues. Au-delà de tout test statistique sur les hypothèses (cf. références dans les sections suivantes, notamment pour le modèle de Cox et la validation de l'hypothèse de proportionnalité), le choix peut être guidé par « avis d'expert », en tenant compte du contexte de l'étude.

Selon les statistiques de l'OMS³, sur la population générale il apparaît qu'entre 30 et 54 ans les différences relatives des taux de décès au sein de la zone CIMA/FANAF lorsque l'âge augmente sont plus stables que les différences absolues (sur ce point, des références sont également disponibles dans le deuxième chapitre). Dans ce contexte, une contrainte sur les différences relatives semblerait plus appropriée.

Mais d'un autre côté, selon les mêmes statistiques de l'OMS, il apparaît qu'au-delà de 55 ans, les différences dans les taux de décès diminuent lorsque l'âge augmente (sur ce point, des références sont également disponibles dans le deuxième chapitre). Dans ce contexte, une contrainte sur les différences absolues semblerait plus appropriée.

En d'autres termes, selon le cadre d'exploitation des données, le modèle de Cox ou le modèle de Lin et Ying peut être plus ou moins approprié. Aussi, dans cette étude, on utilise ces deux modèles pour mesurer le risque d'estimation dans le cadre d'un modèle intégrant l'hétérogénéité à partir de facteurs observables (en particulier, le modèle de Lin et Ying est comparé au modèle de Cox). En outre, on rappelle ici que les modèles de Cox et de Lin et Ying sont également adaptés en présence de données tronquées et censurées (*cf.* Klein et Moeschberger [2005] sur ce point), comme c'est le cas ici.

3. Modèle de Cox : ajustement et simulation des taux de décès

Le modèle de Cox est un modèle classique en analyse de survie et a ainsi été largement étudié. Ici, on se concentre sur son utilisation dans le cadre de la mesure du risque d'estimation.

L'approche retenue ici pour illustrer le risque d'estimation consiste à générer directement et de manière aléatoire des taux bruts (suivant la loi adéquate), afin d'en déduire l'impact sur l'estimation des paramètres requis pour l'estimation des taux ajustés.

Dans ce contexte, dans un premier temps cette section présente l'évaluation des taux de décès annuels ajustés (*cf.* 3.1), et dans un second temps elle présente la simulation des taux de décès en vue de la mesure du risque d'estimation (*cf.* 3.2).

3.1. Cox : évaluation des taux de décès annuels ajustés

Le modèle de Cox, qui pour mémoire peut s'écrire $\lambda(t | Z = z) = \lambda_0(t) e^{\delta^T z}$, permet de mesurer l'effet multiplicatif des variables explicatives, en l'occurrence ici des pays, sur la survie.

À cet effet, on estime le paramètre δ par la méthode du maximum de vraisemblance. On considère en particulier la vraisemblance partielle de Cox, qui se calcule par le produit des probabilités conditionnelles d'observer à l'instant t_i ($i \in [1; D]$) un (ou plusieurs) décès connaissant la composition de la

³ *Cf.* http://apps.who.int/whosis/database/life_tables/life_tables.cfm

population sous risque à cet instant (Hill et al. [1990] présentent une justification de cette démarche).

3.1.1. Estimation de Cox (en absence et en présence d'ex-æquo)

Ce paragraphe présente l'estimation proposée par Cox, en absence et en présence d'ex-æquo.

Estimation de Cox en absence d'ex-æquo

Lorsque l'on suppose qu'il ne se produit qu'un seul décès à chaque instant t_i , Cox [1972] indique que la probabilité conditionnelle que ce soit le sujet de caractéristique $z_{(i)}$ qui meurt en t_i , sachant que l'on avait le groupe R_i de sujets à risque, est :

$$\frac{\exp\{\delta^T z_{(i)}\}}{\sum_{j \in R_i} \exp\{\delta^T z_{j(i)}\}},$$

où $z_{j(i)}$ représente les caractéristiques du $j^{\text{ème}}$ individu sous risque en t_i . La vraisemblance partielle de Cox se calcule ainsi comme le produit de ces contributions, et la log-vraisemblance s'écrit alors :

$$L(\delta) = \sum_{i=1}^D \delta^T z_{(i)} - \sum_{i=1}^D \ln \left(\sum_{j \in R_i} \exp\{\delta^T z_{j(i)}\} \right).$$

Estimation de Cox en présence d'ex-æquo

Lorsque l'on suppose qu'il se produit plusieurs décès à chaque instant t_i , Cox [1972] fournit une nouvelle spécification de son modèle multiplicatif dans un cas discret :

$$\frac{\lambda(t|Z=z)}{1-\lambda(t|Z=z)} = \frac{\lambda_0(t)}{1-\lambda_0(t)} e^{\delta^T z}.$$

En effet, si Cox suppose qu'en temps continu (lorsqu'il ne se produit qu'un seul décès à chaque instant t_i) son modèle s'écrit $\lambda(t|Z=z) = \lambda_0(t) e^{\delta^T z}$, en temps discret (lorsqu'il se produit plusieurs décès à chaque instant t_i) il suppose que les taux de risque instantanés ne sont pas suffisamment proches de 0 pour considérer que $1-\lambda(t|Z=z) = 1$ et $1-\lambda_0(t) = 1$.

Avec cette nouvelle spécification, la contribution à la vraisemblance des d_i décès en t_i , sachant que l'on avait le groupe R_i de sujets à risque, s'écrit :

$$\frac{\exp\{\delta^T s_{(i)}\}}{\sum_{j \in (R_i; d_i)} \exp\{\delta^T s_{j(i)}\}},$$

où $s_{(i)}$ représente la somme des $z_{(i)}$, pour tous les individus morts en t_i , et la notation du dénominateur signifie que la somme est prise sur tous les ensembles d'individus d_i bien distincts tirés dans R_i . La log-vraisemblance s'écrit alors :

$$L(\delta) = \sum_{i=1}^D \delta^T s_{(i)} - \sum_{i=1}^D \ln \left(\sum_{j \in (R_i; d_i)} \exp\{\delta^T s_{j(i)}\} \right).$$

3.1.2. Estimation de Breslow (en présence d'ex-æquo)

Une des spécificités du risque décès, est que la fréquence de survenance du risque est faible et que l'exposition au risque est relativement élevée. Aussi, dans ce cas, le nombre de combinaisons possibles dans la somme du dénominateur $\sum_{j \in (R_i; d_i)} \exp\{\delta^T s_{j(i)}\}$ dans l'estimateur de Cox en présence

d'ex-æquo est particulièrement important et limite la mise en œuvre de l'estimation (notamment lorsque l'estimation s'inscrit dans le cadre de simulations, comme c'est le cas ici).

Dans ce contexte, on peut se tourner vers les approximations de Breslow et d'Elfron (une comparaison des approches de Cox, Breslow et Elfron à partir d'exemples chiffrés simples est présentée dans Klein et Moeschberger [2005]). On retient ici la simplification de Breslow, très utilisée en pratique, selon laquelle la contribution à la vraisemblance des d_i décès en t_i , sachant que l'on avait le groupe R_i de sujets à risque, s'écrit :

$$\exp\{\delta^T s_{(i)}\} / \left[\sum_{j \in R_i} \exp\{\delta^T z_{j(i)}\} \right]^{d_i},$$

avec les notations définies ci-dessus. La log-vraisemblance s'écrit alors :

$$L(\delta) = \sum_{i=1}^D \delta^T s_{(i)} - \sum_{i=1}^D d_i \times \ln \left(\sum_{j \in R_i} \exp\{\delta^T z_{j(i)}\} \right).$$

3.1.3. Statistiques de test

Dans le cadre de cette étude, on se limite pour les illustrations au test de significativité globale et au test de significativité des paramètres. On suppose ici que ces tests sont réalisés à partir de la statistique du rapport des vraisemblances qui suit une distribution du Khi-deux sous l'hypothèse nulle H_0 (cette statistique compte parmi les plus utilisées dans ce contexte, avec la statistique de Wald, cf. par exemple Klein et Moeschberger [2005] pour une présentation illustrée).

Dans le cas du test de significativité globale, on teste l'hypothèse de nullité simultanée de l'ensemble des paramètres, et on considère donc $H_0 : \delta = 0$, soit

(p étant la dimension de Z représentant les variables explicatives du modèle) :

$$\chi_L^2(p) = 2 \left[L(\hat{\delta}) - L(0) \right].$$

Dans le cas du test de significativité des paramètres, on teste l'hypothèse de nullité chaque paramètre δ_j (avec $j=1, \dots, p$ et $\delta = (\delta_1, \dots, \delta_p)$), et on considère donc $H_0 : \delta_j = 0$, soit :

$$\chi_{L_j}^2(1) = 2 \left[L(\hat{\delta}) - L(\tilde{\delta} \setminus \tilde{\delta}_j, \delta_j = 0) \right],$$

où l'expression « $\tilde{\delta} \setminus \tilde{\delta}_j, \delta_j = 0$ » représente l'estimation des paramètres δ_g (avec $g \in (1, \dots, p) \setminus j$) en fixant $\delta_j = 0$.

Seuls ces deux tests de significativité du modèle et des paramètres sont réalisés ici, dans le cadre de notre étude. On note toutefois qu'en pratique, l'utilisation du modèle de Cox renvoie à de nombreux tests statistiques complémentaires. La littérature est abondante à ce sujet (une revue détaillée de la littérature sur les principaux tests statistiques existants sur le modèle de Cox est présentée dans les ouvrages de Hill et al. [1990], Therneau et Grambsch [2000] et Martinussen et Scheike [2006]). En se limitant aux tests les plus classiques et portant uniquement sur l'hypothèse de taux de hasard proportionnels (en pratique, les tests portent également sur la fonction de lien du modèle, la forme des variables du modèle, les propriétés des résidus, etc.), on énumère ici trois approches. La première (approche graphique de Kay) consiste à considérer un modèle stratifié et à illustrer l'évolution des différences des logarithmes des fonctions de hasard cumulés des strates lorsque l'âge augmente : si elles sont à peu près constantes, l'hypothèse de risque proportionnel est supposée adéquate pour la stratification retenue. La deuxième (approche de Therneau et Grambsch [2000]) consiste à considérer une extension du modèle de Cox en considérant un paramètre dépendant du temps : si le paramètre dépend significativement du temps, l'hypothèse de proportionnalité n'est pas appropriée. La troisième (approche de Lin et al. [1993]) consiste à tester l'hypothèse de proportionnalité à partir des résidus cumulés du modèle.

3.1.4. Évaluation des taux de décès annuels ajustés

Une fois les paramètres estimés, on en déduit les taux de décès ajustés pour chaque sous-population.

En pratique, dans un premier temps on détermine les taux de décès ajustés pour la sous-population ivoirienne (sous-population de base dans nos applications) à partir du modèle à référence externe de Brass. À cet effet, les taux bruts sont estimés selon l'approche de Hoem et les taux de référence sont ceux des tables réglementaires françaises TH/TF00-02 (décès). Ces taux ajustés

sont notés $q_{x,CI}(\hat{\theta})$, où $\hat{\theta} = (\hat{a}, \hat{b})$ sont les paramètres du modèle de Brass (cf. chapitre précédent).

Dans un second temps on en déduit les taux de décès ajustés du Mali et du Togo, à partir des paramètres du modèle de Cox, par les relations suivantes (en retenant l'hypothèse que les taux de décès instantanés sont constants entre deux âges entiers) :

$$q_{x,ML}(\hat{\theta}; \hat{\delta}_{ML}) = 1 - \left(1 - q_{x,CI}(\hat{\theta})\right)^{\exp(\hat{\delta}_{ML})},$$

et

$$q_{x,TG}(\hat{\theta}; \hat{\delta}_{TG}) = 1 - \left(1 - q_{x,CI}(\hat{\theta})\right)^{\exp(\hat{\delta}_{TG})},$$

où $\hat{\delta} = (\hat{\delta}_{TG}; \hat{\delta}_{ML})$ sont les paramètres estimés du modèle de Cox.

3.2. Cox : évaluation des taux de décès annuels simulés

On se place dans le cas des taux bruts estimés selon l'approche de Hoem pour chaque sous-population. Ces taux bruts sont notés $\hat{q}_{x,h}$ pour le pays h .

La technique de simulation retenue consiste à considérer une méthode de Monte-Carlo pour simuler la distribution d'une loi normale (à partir de la simulation d'une distribution de loi normale centrée et réduite, elle-même déduite d'une simulation de distribution de loi uniforme entre 0 et 1). Ainsi, pour chaque pays h , on génère dans une première étape k simulations ($k \in [1, K]$) des taux de décès bruts, pour tous les âges x ($x \in [x_m, x_M]$),

$$\text{selon } Q_{x,h} \sim N \left(\hat{q}_{x,h}; \sqrt{\frac{\hat{q}_{x,h}(1-\hat{q}_{x,h})}{R_{x,h}}} \right) \text{ (on reprend ainsi pour chaque pays } h$$

la méthode de simulation directe des taux bruts présentée au chapitre précédent).

Dans une seconde étape, on en déduit pour chaque simulation k et pour tous les âges x la valeur approchée du nombre de décès par pays par $\hat{d}_{x,h}^k \approx A \left(\hat{q}_{x,h}^k \times R_{x,h} \right)$, où $A(\cdot)$ représente l'arrondi à l'entier le plus proche,

$\hat{q}_{x,h}^k$ représente une réalisation k de $Q_{x,h}$ et $R_{x,h}$ représente le groupe de sujets à risque à la l'âge x pour le pays h .

Les étapes suivantes consistent à estimer les K réalisations (au titre des fluctuations d'échantillonnage) des taux ajustés (on parle alors de taux simulés), et à cet effet on distingue le cas de la Côte d'Ivoire du cas du Mali et du Togo.

Dans le cas de la Côte d'Ivoire, les taux ajustés sont obtenus à partir du modèle de Brass. Pour chaque simulation k on a donc :

$$q_{x,CI}(\hat{\theta}^k) = \frac{\exp(\hat{a}^k z_x + \hat{b}^k)}{1 + \exp(\hat{a}^k z_x + \hat{b}^k)},$$

où pour mémoire $z_x = \ln(q_x^{ref} / (1 - q_x^{ref}))$ et $\hat{\theta}^k = (\hat{a}^k, \hat{b}^k)$ est estimée par la méthode des moindres carrés pour chaque simulation k .

Dans le cas du Mali et du Togo, les taux ajustés sont obtenus à partir du modèle de Cox. Ainsi, pour chaque simulation k on estime les paramètres $\delta^k = (\delta_{ML}^k; \delta_{TG}^k)$ dudit modèle avec la méthode de Breslow ; dans ce cas, la log-vraisemblance s'écrit alors :

$$L(\delta^k) = \sum_{i=1}^D (\delta^k)^T s_{(i)}^k - \sum_{i=1}^D \hat{d}_i^k \times \ln \left(\sum_{j \in R_i} \exp \left\{ (\delta^k)^T z_{j(i)} \right\} \right),$$

où $\hat{d}_i^k = \sum_h \hat{d}_{i,h}^k$ (avec $h = CI, ML, TG$ dans notre étude). Pour chaque

simulation k , on déduit ensuite des paramètres estimés $\hat{\delta}^k = (\hat{\delta}_{ML}^k; \hat{\delta}_{TG}^k)$ les taux ajustés du Mali et du Togo par les relations :

$$q_{x,ML}(\hat{\theta}^k; \hat{\delta}_{ML}^k) = 1 - \left(1 - q_{x,CI}(\hat{\theta}^k) \right)^{\exp(\hat{\delta}_{ML}^k)},$$

et

$$q_{x,TG}(\hat{\theta}^k; \hat{\delta}_{TG}^k) = 1 - \left(1 - q_{x,CI}(\hat{\theta}^k) \right)^{\exp(\hat{\delta}_{TG}^k)}.$$

4. Modèle de Lin et Ying : ajustement et simulation des taux de décès

À l'image des travaux présentés pour le modèle de Cox, on se concentre ici sur l'utilisation du modèle de Lin et Ying dans le cadre de la mesure du risque d'estimation.

On présente ainsi d'une part les taux ajustés (cf. 4.1), et d'autre part les taux simulés (cf. 4.2).

4.1. Lin et Ying : évaluation des taux de décès annuels ajustés

Le modèle de Lin et Ying, qui pour mémoire peut s'écrire $\lambda(t | Z = z) = \lambda_0(t) + \gamma^T z$, est un cas particulier du modèle additif d'Aalen. Il permet de mesurer le risque additionnel dû à l'effet des variables explicatives du modèle, en l'occurrence ici des pays, en termes absolu (pour mémoire, le modèle multiplicatif de Cox mesure plutôt l'excès de risque en termes relatif). À cet effet, on estime le paramètre γ à partir d'une formule explicite.

4.1.1. Estimation de Klein et Moeschberger

Lin et Ying [1994] et Klein et Moeschberger [2005] montrent qu'à partir de la décomposition martingale du processus de Poisson, l'estimation des coefficients du modèle est :

$$\hat{\gamma} = A^{-1}B,$$

$$\text{où } A = \sum_{i=1}^D \sum_{j \in R_i} (z_{j(i)} - \bar{z}_{j(i)})^T (z_{j(i)} - \bar{z}_{j(i)}), \quad B = \sum_{i=1}^D \sum_h d_{i,h} (z_{(i),h} - \bar{z}_{(i)})$$

$$\bar{z}_{(i)} = \frac{1}{R_i} \sum_{j \in R_i} z_{j(i)} \quad (\text{avec les mêmes notations que celles de la section 3}$$

relative au modèle de Cox).

4.1.2. Statistiques de test

À l'image des travaux réalisés pour le modèle de Cox, dans le cadre de cette étude on se limite pour les illustrations au test de significativité globale et au test de significativité des paramètres.

Selon les mêmes auteurs, la significativité globale du modèle peut être appréciée à partir de la statistique de Wald qui suit une distribution du Khi-deux à p degrés de libertés (p étant la dimension de Z représentant les variables explicatives du modèle) sous l'hypothèse $H_0 : \gamma = 0$, soit :

$$\chi_{W}^2 = \hat{\gamma}^T \hat{V}^{-1} \hat{\gamma},$$

$$\text{où } \hat{V} = A^{-1}CA^{-1}, \quad \text{avec } C = \sum_{i=1}^D \sum_h d_{i,h} (z_{(i),h} - \bar{z}_{(i)})^T (z_{(i),h} - \bar{z}_{(i)}).$$

De même, dans le cas du test de significativité des paramètres, on teste l'hypothèse de nullité chaque paramètre γ_j (avec $j = 1, \dots, p$ et $\gamma = (\gamma_1, \dots, \gamma_p)$), et on considère donc $H_0 : \gamma_j = 0$, soit :

$$\chi_{W_j}^2 = \hat{\gamma}_j^2 / \hat{V}_{jj}.$$

4.1.3. Évaluation des taux de décès annuels ajustés

Une fois les paramètres estimés, on en déduit les taux de décès ajustés pour chaque sous-population.

En pratique, dans un premier temps on détermine les taux de décès ajustés pour la sous-population ivoirienne (pour mémoire il s'agit de la sous-population de base) à partir de l'approche retenue pour les travaux sur le modèle de Cox (cf. paragraphe 3.1.4). Ces taux ajustés sont notés $q_{x,CI}(\hat{\theta})$, où $\hat{\theta} = (\hat{a}, \hat{b})$ sont les paramètres du modèle de Brass.

Dans un second temps on en déduit les taux de décès ajustés du Mali et du Togo, à partir des paramètres du modèle de Lin et Ying, par la relation suivante (sous l'hypothèse que les taux de décès instantanés sont constants entre deux âges entiers) :

$$q_{x,ML}(\hat{\theta}; \hat{\gamma}_{ML}) = 1 - \left(1 - q_{x,CI}(\hat{\theta})\right) \exp(-\hat{\gamma}_{ML}),$$

et

$$q_{x,TG}(\hat{\theta}; \hat{\gamma}_{TG}) = 1 - \left(1 - q_{x,CI}(\hat{\theta})\right) \exp(-\hat{\gamma}_{TG}),$$

où $\hat{\gamma} = (\hat{\gamma}_{TG}; \hat{\gamma}_{ML})$ sont les paramètres estimés du modèle de Lin et Ying.

4.2. Lin et Ying : évaluation des taux de décès annuels simulés

Les premières étapes de la simulation des taux de décès à partir du modèle de Lin et Ying sont identiques à celles réalisées pour les simulations à partir du modèle de Cox (cf. sous-section 3.2).

Pour chaque simulation k ($k \in [1, K]$) des taux de décès bruts, et pour tous les âges x ($x \in [x_m, x_M]$), on dispose ainsi d'une simulation du nombre de décès par pays par $\hat{d}_{x,h}^k$ et des taux de décès simulés pour la Côte d'Ivoire $q_{x,CI}(\hat{\theta}^k)$.

Il convient alors de déterminer, pour chaque simulation k , les taux de décès du Mali et du Togo. À cet effet, on s'appuie sur le modèle de Lin et Ying et on estime pour la simulation k les paramètres $\gamma^k = (\gamma_{ML}^k; \gamma_{TG}^k)$ dudit modèle. À cet effet, on utilise la relation :

$$\hat{\gamma} = A^{-1}B^k,$$

où $B^k = \sum_{i=1}^D \sum_h \hat{d}_{i,h}^k (z_{(i),h} - \bar{z}_{(i)})$. Pour chaque simulation k , on en déduit ensuite les taux de décès simulés du Mali et du Togo par les relations :

$$q_{x,ML}(\hat{\theta}^k; \hat{\gamma}_{ML}^k) = 1 - \left(1 - q_{x,CI}^k(\hat{\theta}^k) \right) \exp(-\hat{\gamma}_{ML}^k),$$

et

$$q_{x,TG}(\hat{\theta}^k; \hat{\gamma}_{TG}^k) = 1 - \left(1 - q_{x,CI}^k(\hat{\theta}^k) \right) \exp(-\hat{\gamma}_{TG}^k).$$

5. Comparaison des ajustements des modèles de Cox et de Lin et Ying et *backtesting* sur l'hétérogénéité

Cette section présente les résultats des ajustements des modèles intégrant l'hétérogénéité à partir de facteurs observables et présente un *backtesting* sur la capacité de ces modèles à prendre en compte l'hétérogénéité, en comparaison avec une approche consistant à retenir un modèle indépendant pour chaque sous-population.

5.1. Comparaison des ajustements de Cox et de Lin et Ying

Cette sous-section présente les résultats des ajustements des modèles de Cox et de Lin et Ying respectivement. Une comparaison de ces résultats est également présentée.

5.1.1. Résultats des ajustements de Cox

On considère la population UEMOA, représentée ici par la Côte d'Ivoire, le Mali et le Togo (cf. Tableau 30 et Tableau 31 ci-avant).

Les résultats de l'estimation de $\hat{\delta} = (\hat{\delta}_{TG}; \hat{\delta}_{ML})$ à partir des deux approches envisagées en cas de présence d'ex-æquo (approche de Cox ou de Breslow), sont présentés dans le tableau suivant (pour les femmes uniquement). Ces estimations sont réalisées à partir des décès et des expositions au risque par âge.

Tableau 32 – Modèle de Cox : comparaison ajustement (UEMOA – F)

Statistique	Cox (avec ex-æquo)	Breslow (avec ex-æquo)
Minimum(*)	1747	2146
Itérations	8	8
χ_L^2 (<i>p-value</i> modèle)	18,02 ($p = 0,01.10^{-2}$)	18,07 ($p = 0,01.10^{-2}$)
$\hat{\delta}_{ML}$ (valeur initiale)	-0,4638 (0)	-0,4709 (0)
$\exp(\hat{\delta}_{ML})$	0,6289	0,6244
$\chi_{L_{ML}}^2$ (<i>p-value</i> paramètre)	1,03 ($p = 31,01.10^{-2}$)	1,07 ($p = 30,10.10^{-2}$)
$\hat{\delta}_{TG}$ (valeur initiale)	-0,9199 (0)	-0,9198 (0)
$\exp(\hat{\delta}_{TG})$	0,3986	0,3986
$\chi_{L_{TG}}^2$ (<i>p-value</i> paramètre)	17,49 ($p = 2,89.10^{-5}$)	17,51 ($p = 2,86.10^{-5}$)

(*) car en pratique on minimise l'opposé de la log-vraisemblance.

Il apparaît que les estimations de Cox et de Breslow conduisent à des résultats comparables dans cet exemple.

En retenant désormais l'approche de Breslow uniquement, le tableau ci-après présente les résultats des estimations pour la population féminine et masculine.

Tableau 33 – Modèle de Cox : ajustement Breslow (UEMOA – F. et H.)

Statistique	Femmes	Hommes
χ_L^2 (<i>p-value</i> modèle)	18,07 ($p = 0,01.10^{-2}$)	12,79 ($p = 0,17.10^{-2}$)
$\exp(\hat{\delta}_{ML})$	0,6244	0,6084
$\chi_{L_{ML}}^2$ (<i>p-value</i> paramètre)	1,07 ($p = 30,10.10^{-2}$)	7,89 ($p = 0,50.10^{-2}$)
$\exp(\hat{\delta}_{TG})$	0,3986	1,1035
$\chi_{L_{TG}}^2$ (<i>p-value</i> paramètre)	17,51 ($p = 2,86.10^{-5}$)	4,38 ($p = 3,64.10^{-2}$)

Les résultats des paramètres sont cohérents avec les statistiques descriptives présentées dans le Tableau 30 et le Tableau 31. Pour le test de significativité des modèles, il ressort que la valeur du test empirique est supérieure à celle de la table au seuil de 5 % à 2 degrés de liberté (égale à 5,99 pour mémoire), que ce soit pour la population féminine ou masculine. Ainsi, les modèles sont significatifs au seuil de 5 %. Au niveau des variables toutefois, il apparaît que le paramètre du Mali pour les femmes n'est pas significatif au seuil de 5 % (pour ce paramètre individuel, la valeur du test empirique est inférieure à celle de la table au seuil de 5 % à 1 degré de liberté, égale à 3,84 pour mémoire). Les paramètres des autres variables sont en revanche significatifs au seuil de 5 %.

5.1.2. Résultats des ajustements de Lin et Ying

À l'image de l'illustration du modèle de Cox, on considère ici la population féminine et masculine de la population UEMOA, représentée ici par la Côte d'Ivoire, le Mali et le Togo (cf. Tableau 30 et Tableau 31).

Les résultats de l'estimation de $\hat{\gamma} = (\hat{\gamma}_{TG}; \hat{\gamma}_{ML})$ sont présentés dans le tableau suivant.

Tableau 34 – Modèle de Lin et Ying : ajustement (UEMOA – F. & H.)

Statistique	Femmes	Hommes
χ_W^2 (<i>p-value</i> modèle)	26,55 ($p = 1,71.10^{-6}$)	16,01 ($p = 0,03.10^{-2}$)
$\hat{\gamma}_{ML}$	-0,0693 %	-0,1429 %
$\chi_{W_{ML}}^2$ (<i>p-value</i> paramètre)	1,43 ($p = 23,25.10^{-2}$)	10,68 ($p = 0,11.10^{-2}$)
$\hat{\gamma}_{TG}$	-0,1129 %	0,0399 %
$\chi_{W_{TG}}^2$ (<i>p-value</i> paramètre)	26,14 ($p = 3,17.10^{-7}$)	4,07 ($p = 4,37.10^{-2}$)

Les résultats des paramètres sont cohérents avec les statistiques descriptives présentées dans le Tableau 30 (pour les femmes) et le Tableau 31 (pour les hommes) d'une part, et avec les coefficients d'abattement du modèle de Cox présentés dans le Tableau 33 (pour les femmes et les hommes) d'autre part. En termes de tests statistiques, à l'image des résultats du modèle de Cox, il apparaît que les modèles pour les femmes et pour les hommes sont significatifs au seuil de 5 %, même si le paramètre du Mali pour les femmes ne l'est pas.

En outre, concernant la comparaison entre le modèle multiplicatif (Cox) et le modèle additif (Lin et Ying) sur le volet opérationnel, on retrouve les principales conclusions des travaux de Cao [2005]. Ainsi, il apparaît que ces deux modèles multiplicatifs et additifs sont exploitables pour les données tronquées et/ou censurées, permettent de tenir compte de facteurs observables et présentent des *p-value* cohérentes pour les coefficients des variables du modèle.

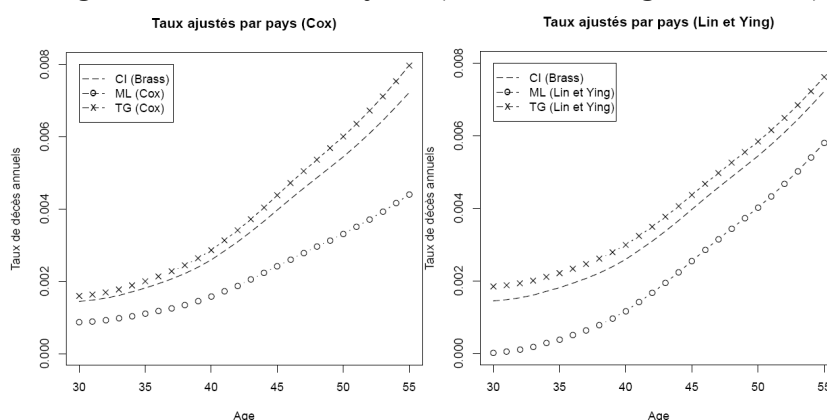
5.1.3. Illustration et comparaison des résultats des ajustements de Cox et de Lin et Ying

Les deux graphiques ci-dessous (Figure 36) présentent une comparaison entre les résultats du modèle de Cox et ceux du modèle de Lin et Ying, pour la population masculine uniquement. Le premier graphique présente les taux de décès ajustés de la Côte d'Ivoire, issus du modèle de Brass, ainsi que ceux du Mali et du Togo, déduits des paramètres $\hat{\delta} = (\hat{\delta}_{TG}; \hat{\delta}_{ML})$ du modèle de Cox. Sur le même principe, le second graphique présente les taux de décès ajustés de la Côte d'Ivoire issus du modèle de Brass, ainsi que ceux du Mali et du Togo, déduits des paramètres $\hat{\gamma} = (\hat{\gamma}_{TG}; \hat{\gamma}_{ML})$ du modèle de Lin et Ying (dans les deux graphiques, les courbes relatives à la Côte d'Ivoire sont donc identiques).

En termes statistiques, dans le cadre de l'ajustement des taux à partir du modèle de Brass pour la Côte d'Ivoire, il ressort que le modèle est significatif

au seuil de 5 % (test de Fisher) et que le R2 ajusté est égal à 83,8 %. On rappelle également que les paramètres retenus pour l'évaluation des taux ajustés à partir de Ling et Ying pour le Togo et le Mali sont significatifs au seuil de 5 %.

Figure 36 - Taux de décès ajustés (Cox et Lin et Ying, UEMOA - H)



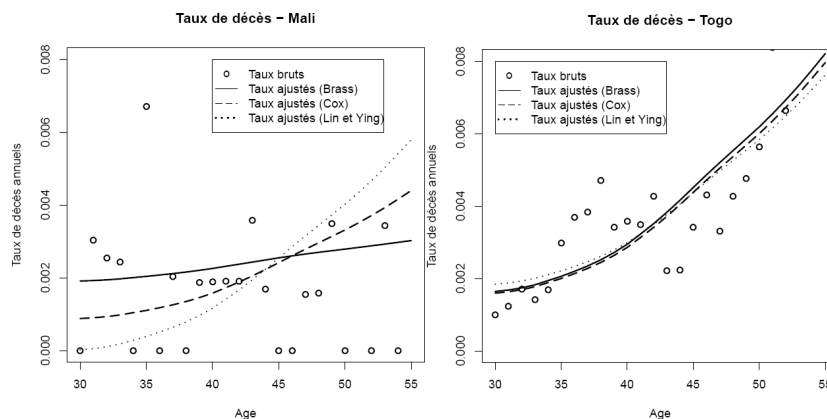
Dans la continuité des commentaires relatifs au choix entre ces deux modèles (*cf.* sous-section 2.2), il apparaît qu'avec le modèle d'hétérogénéité multiplicatif de Cox, les différences absolues entre les pays augmentent significativement lorsque l'âge augmente, contrairement au modèle d'hétérogénéité de Lin et Ying dans lequel les différences absolues sont constantes pour tous les âges.

5.2. Backtesting sur l'hétérogénéité (comparaison des modèles d'hétérogénéité)

L'objectif de l'étude est de comparer le risque d'estimation selon deux approches pour le traitement de l'hétérogénéité. Cette comparaison est présentée dans la section suivante (section 6), et on se concentre ici sur le traitement de l'hétérogénéité selon ces approches.

On considère une population composée de trois sous-populations : la Côte d'Ivoire (sous-population de base), le Mali et le Togo. Dans la première approche, l'hétérogénéité est prise en compte à partir de modèles indépendants pour chaque sous-population (en pratique, dans ce cas les taux de décès de chaque pays sont ajustés à partir d'un modèle de Brass). Dans la seconde approche, l'hétérogénéité est prise en compte à partir de modèles intégrant des facteurs observables (ici, les taux de décès sont ajustés à partir du modèle de Brass pour la Côte d'Ivoire et à partir du modèle de Cox ou de Lin et Ying pour le Mali et le Togo).

Les résultats des ajustements de taux de décès selon ces deux approches (qui comptent trois modèles au total) sont présentés sur le graphique ci-dessous, pour le Mali et le Togo (le graphique relatif à la Côte d'Ivoire n'est pas présenté car il est identique dans les deux approches).

Figure 37 – Présentation des taux de décès du Mali et Togo (H)

Dans le cas du Mali, il apparaît des taux de décès bruts nuls (*cf.* graphique ci-dessus), ce qui limite l'utilisation du logit des taux de décès bruts, pourtant indispensable à la modélisation de Brass. La convention suivante a donc été retenue pour l'ajustement (pour le modèle de Brass uniquement) : il a été considéré que les taux nuls sont égaux au plus faible taux de décès brut non nul de la plage d'âge retenue pour l'ajustement. Des approches alternatives auraient pu être par exemple de ne pas retenir les âges où l'on observe des taux bruts nuls ou de substituer les taux bruts nuls par les taux moyens observés, mais ces solutions majorent significativement les taux décès retenus pour l'ajustement.

Sur ces bases, pour le Mali, il apparaît que les taux de décès ajustés divergent et dépendent du modèle. Sur le plan statistique, le modèle de Brass pour le Mali (première approche) n'est pas significatif au seuil de 5 % (test de Fisher), et le R² ajusté est égal à 1,4 %. Par ailleurs, pour mémoire, les paramètres du Mali pour les modèles de Cox et de Lin et Ying (modèles de la deuxième approche) sont significatifs au seuil de 5 % (*cf.* Tableau 33 et Tableau 34). On retient donc ici que les modèles intégrant l'hétérogénéité permettent, contrairement au modèle de Brass, d'éviter de retenir des conventions en cas de taux bruts nuls d'une part et de présenter des résultats satisfaisants en termes de tests statistiques d'autre part.

Pour le Togo, il apparaît que les taux de décès ajustés sont proches, quel que soit le modèle considéré. On note ici que le modèle de Brass pour le Togo est significatif au seuil de 5 % (test de Fisher) et que le R² ajusté est égal à 69,2 %. Pour mémoire, les paramètres du Togo pour les modèles de Cox et de Lin et Ying sont significatifs au seuil de 5 % (*cf.* Tableau 33 et Tableau 34), ce qui est d'ailleurs cohérent avec les résultats des tests d'hétérogénéité effectués au chapitre précédent.

Afin d'apprécier la pertinence des taux de décès ajustés dans ces trois modèles, le tableau suivant présente des comparaisons entre les réalisations et les prédictions des décès (établies à partir des taux de décès ajustés et des expositions au risque par âge).

Tableau 35 – Comparaison réalisation / prédiction décès (UEMOA –H.)

Modèle global UEMOA - H (Brass global)				Modèles intégrant l'hétérogénéité (sans et avec fact. obs.)			
Pays	Décès observés	Décès prédits	Différence relative	Pays (modèle)	Décès observés	Décès prédits	Différence relative
Côte d'Ivoire	2 188	2 203	0,7%	Côte d'Ivoire (Brass)	2 188	2 144	-2,0%
Mali	27	44	63,8%	Mali (Brass ^(*))	27	29	8,4%
				Mali (Cox)		26	-3,2%
				Mali (Lin et Ying)		26	-4,9%
Togo	565	511	-9,6%	Togo (Brass)	565	565	-0,1%
				Togo (Cox)		548	-3,0%
				Togo (Lin et Ying)		550	-2,7%

(*) avec convention pour le traitement des taux de décès bruts nuls.

Le premier sous-tableau présente les écarts lorsque les taux de décès ajustés sont déterminés au global, sans prendre en compte l'hétérogénéité entre sous-populations. Le deuxième sous-tableau présente les écarts lorsque les taux de décès ajustés sont déterminés à partir de modèles intégrant l'hétérogénéité, soit à partir de modèles indépendants pour chaque sous-population (avec le modèle de Brass), soit à partir de modèles intégrant l'hétérogénéité à partir de facteurs observables (avec les modèles de Cox et de Lin et Ying). Une comparaison des deux sous-tableaux confirme la nécessité de tenir compte de l'hétérogénéité dans la population.

Dans le sous-tableau intégrant l'hétérogénéité il apparaît, hormis pour les taux ajustés du Mali avec le modèle de Brass, que les écarts entre les décès observés et les décès théoriques sont tous inférieurs à 5 %, ce qui confirme la prise en compte de l'hétérogénéité.

Dans le détail, pour ce deuxième sous-tableau on note dans un premier temps que le nombre de décès théoriques pour la Côte d'Ivoire sous-estime de 2 % le nombre de décès observés. Cet écart traduit le biais d'estimation introduit par la fonction logit du modèle de Brass. En effet, pour mémoire, le caractère concave de la fonction logit entre 0 et $\frac{1}{2}$ conduit, en utilisant l'inégalité de Jensen, à une sous-estimation des taux de décès (*cf.* Planchet et Théron [2006] pour une description plus complète de ce phénomène).

Les estimations du modèle de Brass pour la Côte d'Ivoire sont à la base des estimations issues des modèles de Cox et de Lin et Ying pour le Mali et le Togo. Aussi, les écarts observés pour le Mali et le Togo à partir de ces modèles incluent l'écart observé pour la Côte d'Ivoire à partir du modèle de Brass.

Enfin, malgré la convention retenue pour le traitement du Mali avec le modèle de Brass (substitution des taux nuls par le plus faible taux non nul observé), on note sur ce deuxième sous-tableau que les décès théoriques du Mali estimés avec le modèle de Brass majorent les décès observés. À titre de comparaison, avec la convention consistant à ne retenir que les âges où les taux de décès bruts sont non nuls, le nombre de décès prédits pour le Mali s'élève à 38 (soit un écart de 39,7 % avec les décès théoriques, contre 8,4 % avec la convention retenue). La convention retenue semble donc plus adaptée, bien que présentant des limites importantes (les écarts obtenus pour le modèle de Brass appliqué au Mali ne sont pas satisfaisants compte tenu des écarts observés pour les autres pays et les autres modèles).

Au final, il ressort de ces analyses que le choix d'un modèle intégrant l'hétérogénéité à partir de facteurs observables (par exemple le modèle de Cox ou le modèle de Lin et Ying), permet de conserver une bonne appréciation statistique du risque de décès pour les sous-populations hétérogènes au sein d'une population. En outre, ces modèles permettent de tenir en compte des taux annuels nuls pour des âges données (phénomène fréquent dans les petites populations), et permet ainsi d'éviter de retenir des hypothèses ou convention (contrairement au modèle de Brass).

6. Comparaison des risques d'estimation entre les modèles indépendants pour chaque sous-population et ceux intégrant l'hétérogénéité à partir de facteurs observables

Cette section présente le risque d'estimation dans les différents modèles retenus pour chacune des deux approches : modèle de Brass (modèle de la première approche), modèle de Cox et modèle de Lin et Ying (modèles de la seconde approche). Pour les trois modèles étudiés, les outils de mesure sont présentés à travers les taux de décès ajustés d'une part, et à travers les tables de mortalité d'autre part.

De manière générale, les approches présentées ici s'inscrivent dans la continuité de celles présentées en détail dans le chapitre précédent sur la mesure du risque d'estimation dans le cadre du modèle de Brass.

Par ailleurs, dans l'ensemble des applications numériques, on s'appuie sur $K = 1000$ simulations des taux bruts et on considère que pour chaque simulation k les réalisations de la loi normale centrée réduite $N(0;1)$, desquels on déduit les réalisations appropriées pour chaque sous-population, sont exactement les mêmes dans les trois modèles utilisés (pour une sous-population donnée). On note en outre que les taux bruts sont en pratique générés sous la contrainte $q_{x,h}^k > 0$ (lorsque cette condition n'est pas vérifiée pour une simulation k , celle-ci est annulée et reconduite pour le pays h).

Enfin, dans cette section, seules les illustrations numériques relatives au Togo seront présentées : les illustrations de la Côte d'Ivoire ne sont pas présentées car la modélisation est identique dans les deux approches, et celles du Mali ne le sont pas car elles font appel à des conventions arbitraires, ce qui peut être de nature à créer un biais dans l'exploitation et l'analyse des résultats. De surcroît, l'objectif de l'étude est de présenter les mesures de risque d'estimation selon le choix de modèle retenu pour tenir compte de l'hétérogénéité ; balayer le risque d'estimation des différents modèles en se limitant à la sous-population togolaise permet ainsi d'améliorer la clarté des illustrations.

6.1. Mesure du risque d'estimation sur les taux ajustés

On présente ici les outils de mesure du risque d'estimation sur les taux ajustés, ainsi qu'une illustration et interprétation des résultats pour les trois modèles étudiés.

6.1.1. Outils de mesure du risque d'estimation sur les taux

On dispose des taux de décès ajustés selon le modèle de Brass, le modèle de Cox ou le modèle de Lin et Ying. Pour simplifier les notations, on note $q_{x,h}(\hat{\theta})$ ces taux de décès ajustés pour un pays h (pour le Mali et le Togo, les taux de décès ajustés devraient théoriquement être notés $q_{x,h}(\hat{\theta}; \hat{\delta}_h)$ et $q_{x,h}(\hat{\theta}; \hat{\gamma}_h)$, respectivement pour les estimations à partir du modèle de Cox et de Lin et Ying). À partir de K simulations des taux bruts pour chaque pays h , on dispose également des taux de décès simulés selon le modèle de Brass, le modèle de Cox ou le modèle de Lin et Ying. Pour simplifier les notations, on note $q_{x,h}(\hat{\theta}^k)$ les taux de décès pour une simulation k et un pays h (pour le Mali et le Togo, les taux de décès simulés devraient théoriquement être notés $q_{x,h}(\hat{\theta}^k; \hat{\delta}_h^k)$ et $q_{x,h}(\hat{\theta}^k; \hat{\gamma}_h^k)$, respectivement pour les estimations à partir du modèle de Cox et de Lin et Ying).

Le risque d'estimation des taux de décès pour un pays h peut alors être mesuré à partir du coefficient :

$$c(\Psi_{x,h}) = \frac{\Psi_{x,h}}{q_{x,h}(\hat{\theta})},$$

où $\Psi_{x,h} = \sqrt{E \left[\left(q_{x,h}(\hat{\theta}^k) - q_{x,h}(\hat{\theta}) \right)^2 \right]}$. Ce coefficient constitue une mesure

de dispersion, en pourcentage et au titre du risque d'estimation, des taux de décès simulés autour du taux de décès ajusté.

6.1.2. Comparaison du risque d'estimation sur les taux

La comparaison du risque d'estimation sur les taux ajustés est réalisée à partir de la moyenne du coefficient $c(\Psi_{x,h})$, pour tous les âges $x \in [x_m, x_M]$.

Tableau 36 – Risque d'estimation sur les taux (moyenne) (Togo - H)

Population	Modèle de Brass (approche 1)	Modèle de Cox (approche 2)	Modèle de Lin et Ying (approche 2)
Togo $c(\psi_{TG})$	9,89 %	6,19 %	6,78 %

Il apparaît que l'utilisation du modèle de Brass (modèle indépendant pour chaque sous-population), conduit à un risque d'estimation supérieur à celui obtenu avec le modèle de Cox (modèle multiplicatif intégrant l'hétérogénéité à partir de facteurs observables) ou avec le modèle de Lin et Ying (modèle additif intégrant l'hétérogénéité à partir de facteurs observables). Pour le modèle de Lin et Ying, les comparaisons doivent toutefois être traitées avec prudence, compte tenu de l'hypothèse de constance des différences absolues des taux de décès instantanés.

Pour information, on note par ailleurs que le risque d'estimation est d'autant plus important que la population présente une faible exposition. Ainsi, à titre d'exemple le risque d'estimation avec le modèle de Brass pour la Côte d'Ivoire est égal à 4,73 %. De même, le risque d'estimation pour le Mali avec le modèle de Cox est égal à 22,27 %.

Les comparaisons présentées ci-dessus se limitent aux taux de décès simulés ; les comparaisons des tables de mortalité construites à partir des taux de décès simulés sont présentées dans la sous-section ci-dessous.

6.2. Mesure du risque d'estimation sur les tables

Outre la présentation et la comparaison des taux de décès, il est nécessaire de comparer les tables de mortalité associées.

6.2.1. Outil de mesure du risque d'estimation sur les tables

Pour la mesure du risque d'estimation sur les tables de mortalité, il convient d'utiliser une fonctionnelle propre à chaque table qui lui associe un nombre positif. L'espérance de vie est de ce point de vue une fonctionnelle naturelle, et dans le cadre de l'assurance, le montant de la provision en est une autre. Ces deux fonctionnelles sont ainsi utilisées dans les travaux ci-dessous.

Ainsi, dans un premier temps le risque d'estimation sur les tables est mesuré à partir de l'espérance de vie résiduelle à 30 ans, entre 30 et 55 ans. On utilise à cet effet la distribution des espérances de vie résiduelles établies à partir des taux simulés.

Dans un second temps, le risque d'estimation sur les tables est mesuré à partir des provisions. On considère plus précisément des provisions déterministes relatives à des engagements temporaires au décès (la durée de l'engagement est de d années, avec $d \geq 1$) et évaluées à partir des taux de décès simulés (dans ce cas, les taux de décès sont les seules sources d'aléas). En faisant l'hypothèse de décès en milieu d'année, on en déduit que le montant de la provision pour la réalisation k est :

$$L_{0,h}^k = \sum_{t=0}^{d-1} F_{x,h}^k(t) \times (1+r_{t+1})^{-t-1/2},$$

où $F_{x,h}^k(t)$ représente les flux probables des prestations à payer en t pour un individu d'âge x du pays h (pour la réalisation k des taux de décès simulés) et r_t représente le taux d'actualisation des flux de l'échéance t . L'impact du risque d'estimation sur les provisions peut alors être mesuré par le coefficient :

$$c(\Upsilon_h) = \frac{\Upsilon_h}{L_{0,h}},$$

où $L_{0,h}$ correspond à la provision calculée à partir des taux de décès ajustés

$q_{x,h}(\hat{\theta})$ et $\Upsilon_h = \sqrt{E \left[\left(L_{0,h}^k - L_{0,h} \right)^2 \right]}$. Ce coefficient permet d'avoir en

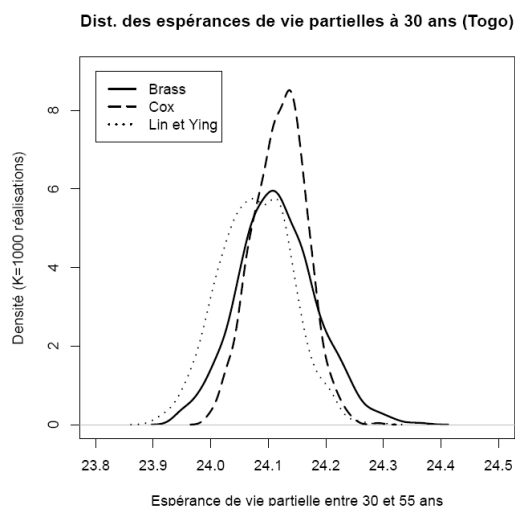
pourcentage une mesure de dispersion (au titre du risque d'estimation) autour de la provision calculée à partir des taux de décès ajustés.

On note enfin $\bar{L}_{0,h} = \frac{1}{K} \sum_{j=1}^K L_{0,h}^j$, la moyenne des provisions déterministes

calculées avec les taux de décès simulés.

6.2.2. Comparaison du risque d'estimation sur les tables

Dans un premier temps, on présente le risque d'estimation sur les tables à travers les espérances de vie partielles. Dans ce contexte, le graphique suivant présente, pour la sous-population du Togo, les estimations des fonctions de densité (par l'estimateur à noyau) des espérances de vie résiduelles (entre 30 et 55 ans) établies à partir des taux de décès simulés.

Figure 38 - Distribution des espérances de vie partielles du Togo (H)

Dans la continuité des observations et des conclusions ci-dessus, il apparaît sur ce graphique que la distribution de l'espérance de vie partielle issue des modèles de Cox présente une queue moins épaisse que celle issue du modèle de Brass. En d'autres termes, les tables de mortalité issues des taux simulés à partir du modèle multiplicatif intégrant l'hétérogénéité à l'aide de facteurs observables présentent une volatilité plus faible que les tables issues des taux simulés à partir du modèle de Brass utilisé indépendamment pour chaque sous-population. Dans le cas du modèle additif de Lin et Ying en revanche, il apparaît que les tables de mortalité présentent une volatilité comparable à celle issue du modèle de Brass.

Dans un second temps, il est question de quantifier ce risque d'estimation supplémentaire sur les tables dans un contexte d'assurance. À cet effet, on introduit la notion de provision, calculée avec la courbe d'actualisation de l'Institut des Actuaire au 31/12/2009⁴. En outre, on considère un assuré d'âge $x = 31$, un capital assuré égal à $C = 1$ et une durée de contrat égal à $d = 20$.

À ce titre, on présente dans le tableau ci-dessous les résultats synthétiques des estimations et des simulations de provisions déterministes, respectivement avec les taux ajustés et avec les taux simulés.

⁴ Courbe disponible sur le site de l'Institut des Actuaire : http://www.institutdesactuaire.com/gene/link.php?doc_link=../docs/2010003190313_I A20091231.xls.

Tableau 37 - Risque d'estimation sur les provisions (20 ans) (Togo - H)

Statistique	Brass		Cox		Lin et Ying	
	Prov. et taux ajustés	Prov. et taux simulés	Prov. et taux ajustés	Prov. et taux simulés	Prov. et taux ajustés	Prov. et taux simulés
Moyenne (L_0 ou \bar{L}_0)	$4,18.10^{-2}$	$4,03.10^{-2}$	$4,06.10^{-2}$	$4,01.10^{-2}$	$4,22.10^{-2}$	$4,17.10^{-2}$
Quantile à 0,5 %	Sans objet	$3,18.10^{-2}$	Sans objet	$3,52.10^{-2}$	Sans objet	$3,54.10^{-2}$
Quantile à 5 %	Sans objet	$3,52.10^{-2}$	Sans objet	$3,70.10^{-2}$	Sans objet	$3,74.10^{-2}$
Quantile à 95 %	Sans objet	$4,50.10^{-2}$	Sans objet	$4,34.10^{-2}$	Sans objet	$4,61.10^{-2}$
Quantile à 99,5 %	Sans objet	$4,74.10^{-2}$	Sans objet	$4,50.10^{-2}$	Sans objet	$4,83.10^{-2}$
Coefficient $c(r)$	Sans objet	7,91 %	Sans objet	5,15 %	Sans objet	6,41 %

Il ressort qu'avec les données masculines du Togo, la prise en compte du risque systématique diminue de 3,6 % la provision calculée lorsque l'on retient le modèle de Brass, alors que l'impact est plus faible avec les modèles de Cox de Lin et Ying (baisses de 1,2 % pour ces deux modèles). Concernant les quantiles, il ressort que les écarts avec les moyennes sont plus importants pour le modèle de Brass que pour les modèles de Cox et de Lin et Ying. Enfin, il apparaît que le coefficient $c(r)$ relatif à la modélisation à partir des modèles de Cox ou de Lin et Ying minore celui obtenu à partir du modèle de Brass (baisses respectives de 35 % et 20 %).

On retient donc que le choix d'un modèle intégrant l'hétérogénéité permet de réduire significativement le risque d'estimation associé à la construction de la loi d'expérience. Dans notre exemple, il apparaît cependant que le poids du risque d'estimation dans l'évaluation d'une provision (mesuré pour chaque modèle par la différence entre la provision calculée à partir des taux ajustées et la moyenne de celles calculées à partir des taux simulés) est comparable au poids du risque de modèle (mesuré par la différence de la provision calculée à partir des taux ajustées entre les trois modèles proposés). Une partie du risque d'estimation est ainsi transféré vers un risque de modèle.

7. Synthèse et conclusion

Dans un contexte d'hétérogénéité, cette étude a pour objectif d'apprécier l'évolution du risque d'estimation lors du passage d'une évaluation des taux de

décès à partir de modèles indépendants pour chaque sous-population (première approche), à un modèle intégrant directement l'hétérogénéité à partir de facteurs observables (seconde approche).

S'il est acquis que dans cette étude le choix du modèle pour la première approche est le modèle de Brass (dans la continuité des travaux réalisés dans le chapitre précédent), il convient de s'interroger sur le(s) choix de modèle(s) pour la seconde approche. L'analyse menée dans la section 2 a conduit au choix du modèle multiplicatif de Cox et du modèle additif de Lin et Ying.

Dès lors, l'ajustement des taux de décès pour chaque sous-population a été présenté et mis en œuvre pour ces deux modèles de la seconde approche. Une comparaison des résultats du calibrage et des tests statistiques associés à ces deux modèles illustre leur cohérence. Les modèles de Cox et de Lin et Ying présentent toutefois des différences, qui peuvent justifier le choix de l'un ou l'autre selon les besoins. Ainsi, alors que le modèle de Cox considère que les différences relatives des taux de décès instantanés sont stables (les différences absolues augmentent donc significativement avec l'âge), le modèle de Lin et Ying considère que les différences absolues sont stables (les différences relatives diminuent donc significativement avec l'âge). En conséquence, si ces deux modèles peuvent conduire à des résultats comparables sur les âges retenus pour l'ajustement, ils peuvent aboutir à des taux de décès sensiblement différents en dehors de cette plage d'âge (de surcroît, le modèle de Lin et Ying peut présenter d'importantes limites opérationnelles pour les âges faibles).

Après avoir vérifié que ces deux modèles prennent en compte le caractère hétérogène de la population, une mesure du risque d'estimation a été réalisée dans ce chapitre. À cet effet, plusieurs estimations des taux de décès issus du modèle de Brass (première approche) et des modèles de Cox et de Lin et Ying (deuxième approche) ont été simulées sur la base de simulations de taux de décès bruts (générés à partir de la loi des taux bruts initiaux, afin de refléter les fluctuations d'échantillonnage).

Les résultats obtenus illustrent une diminution de l'impact du risque d'estimation sur les provisions de plus de 50 % avec les modèles de la deuxième approche (modèles de Cox et de Lin et Ying), par rapport à l'impact observé avec le modèle de la première approche (modèle de Brass). En outre, avec les modèles de la seconde approche, le coefficient mesurant le risque d'estimation sur les provisions diminue de 20 % à 35 %. Néanmoins, dans notre exemple on note en parallèle que le poids du risque d'estimation dans les provisions (qui reflète la dispersion des estimations de provision due aux fluctuations d'échantillonnage des taux bruts, pour un modèle donné) est comparable au poids du risque de modèle (qui reflète la dispersion des estimations de provision due au choix du modèle retenu). Dans la recherche d'un modèle permettant de réduire le risque d'estimation, une attention particulière doit donc également être accordée aux conséquences en termes de risque de modèle.

Au final, le choix d'un modèle intégrant l'hétérogénéité à partir de facteurs observables (deuxième approche) présente plusieurs avantages, parmi lesquels on retient :

- la diminution de l'impact du risque d'estimation sur la provision estimée (baisse de plus de 50 % par rapport à l'impact observé avec le modèle de Brass utilisé indépendamment pour chaque sous-population) ;
- la capacité à modéliser les taux de décès lorsque les données sont sensiblement limitées (en particulier lorsqu'il existe des données avec des taux de décès bruts nuls, sachant que de telles données ne sont pas exploitables avec le modèle de Brass, sauf à retenir des conventions arbitraires).

Le choix de modèles de la seconde approche présente toutefois des inconvénients, parmi lesquels on note :

- l'hypothèse sur la constance des différences relatives (modèle de Cox) ou absolues (modèle de Lin et Ying) des taux de décès instantanés ;
- de manière générale, l'impact potentiel en termes de risque de modèle.

Bibliographie

- Barbi E., Caselli G., Vallin J. [2003] « Hétérogénéité des générations et âge extrême de la vie », *Institut Nationale d'Études Démographiques / Population*, Vol. 58, No. 2006-1.
- Cao H. [2005] *A comparison Between the Additive and Multiplicative Risk Models*, Mémoire d'actuariat, Université Laval.
- Cox D. R. [1972] « Regression Models and Life-Tables », *Journal of the Royal Society. Series B (Methodological)*, Vol. 34, No. 2.
- Delwarde A., Denuit M. [2006] « Construction de tables de mortalité périodiques et prospectives », *Economica*.
- Droesbeke J.-J., Fichet B., Tassi P. [1989] (éditeurs) « Analyse statistique des durées de vie », *Economica*.
- Hill C., Com-Nougé C., Kramar A., Moreau T., O'Quigley J., Senoussi R., Chastang C. [1990] « Analyse statistique des données de survie », *Inserm, Médecine – Sciences, Flammarion*.
- Kamega A., Planchet F. [2010], « Mesure du risque d'estimation associé à une table d'expérience », *Cahiers de recherche de l'ISFA, WP2136*.
- Kamega A., Planchet F. [2011], « Hétérogénéité : mesure du risque d'estimation dans le cas d'une modélisation intégrant des facteurs observables », *Bulletin Français d'Actuariat, Vol. 11, No. 21*.
- Klein J. P., Moeschberger M. L. [2005] « Survival Analysis – Techniques for Censored and Truncated Data », *Springer, 2nd edition*.
- Lin D. Y., Wei L.J., Ying Z. [1993] « Checking the Cox model with cumulative sums of martingale-based residuals », *Biometrika, n. 80*.
- Lin D. Y., Ying Z. [1994] « Semiparametric analysis of the additive risk model », *Biometrika, n. 81*.
- Lopez O. [2007] *Réduction de dimension en présence de données censurées*, Thèse de doctorat, Université de Rennes 1.
- Martinussen T., Scheike T. H. [2006] « Dynamic Regression Models for Survival Data », *Springer*.
- McCullagh P., Nelder J. A. [1989] « Generalized Linear Models », *Monographs on Statistics and Applied Probability 37, Chapman & Hall/CRC, 2nd edition*.
- Planchet F., Leroy G. [2009] « Quel niveau de segmentation pertinent ? », *La Tribune de l'Assurance, n. 142*.
- Planchet F., Théron P. [2006] « Modèles de Durée - Applications actuarielles », *Economica*.

-
- Therneau T. M., Grambsch P. M. [2000] « Modeling Survival Data – Extending the Cox Model », *Springer*.
- Vaupel J. W. [2002] « Life Expectancy at Current Rates vs. Current Conditions: A Reflexion Stimulated by Bongaarts and Feeney's "How Long Do We Live?" », *Demographic Research*, Vol. 7, Art. 8.
- Vaupel J. W., Manton K. G. Stallard E. [1979] « The impact of heterogeneity in individual frailty on the dynamics of mortality », *Demography*, Vol. 16, No. 3.
- Viallon V. [2006] *Processus empiriques, estimation non paramétrique et données censurées*, Thèse de doctorat, Université Paris 6.

PARTIE III

MORTALITÉ FUTURE : MODÉLISATION PROSPECTIVE DE LA MORTALITÉ ET MESURE DES RISQUES ASSOCIÉS DANS LA ZONE

Les travaux de la première et de la deuxième parties sont réalisés au titre de la mortalité actuelle uniquement. Cependant, parmi les résolutions des états généraux de 2007 il apparaît la volonté d'instaurer des régimes de retraite complémentaires gérés par les assureurs, et dans ce cadre la modélisation de la mortalité actuelle n'est pas suffisante. La modélisation de la mortalité future est donc nécessaire, et cette troisième partie de l'étude s'inscrit dans ce cadre. À cet effet toutefois, deux scénarios distincts sont considérés.

Dans le premier scénario, on considère une situation dans laquelle aucune donnée sur les tendances passées de la mortalité n'est disponible. Dans ce cas, une modélisation de la mortalité future à partir d'une extrapolation des tendances passées n'est pas possible. Aussi en pratique, dans ce scénario, qui correspond à la situation réelle aujourd'hui dans la région, la mortalité future est modélisée à partir d'avis d'experts sur l'évolution possible de la mortalité d'expérience dans la région. Une analyse du risque systématique lié au choix des avis d'expert y est également présentée.

Dans le second scénario, des données d'expérience sur les tendances passées de la mortalité sont supposées disponibles (si ce n'est pas le cas dans la région aujourd'hui, de telles données devraient devenir disponibles dans les prochaines années dans la région). La modélisation de la mortalité future est alors classiquement réalisée à partir des tendances passées et d'avis d'expert. L'analyse menée dans ce scénario consiste alors à fournir des outils pour mesurer les risques systématiques associés à cette modélisation (au titre des fluctuations d'échantillonnage – risque d'estimation – et au titre du choix d'avis d'expert).

Chapitre 5

Modélisation prospective en l'absence de données sur les tendances passées et mesure des risques associés¹

Sommaire

CHAPITRE 5 MODÉLISATION PROSPECTIVE EN L'ABSENCE DE DONNÉES SUR LES TENDANCES PASSÉES ET MESURE DES RISQUES ASSOCIÉS	161
1. Introduction.....	162
2. Choix du modèle pour projeter la mortalité.....	164
2.1. Quelques réflexions et quelques outils pour modéliser la mortalité d'expérience future	165
2.2. Modèle de Lee-Carter : limites, extension et alternatives	170
2.3. Approche de Bongaarts : justification du choix	174
3. Ajustement du modèle de Bongaarts.....	176
3.1. Ajustement statistique.....	176
3.2. Ajustement par avis d'expert	179
4. Application : anticipation de la mortalité d'expérience future	183
4.1. Résultats de l'ajustement du modèle.....	183
4.2. Sensibilités et mesure du risque d'avis d'expert	188
5. Synthèse et conclusion	191
Bibliographie.....	193

Ce chapitre présente une méthode simple pour modéliser la mortalité d'expérience future en l'absence de données d'historique suffisantes. Ainsi, avec un ajustement paramétrique, une hypothèse sur la forme de l'évolution du niveau de mortalité (évolution linéaire ou exponentielle) et un avis d'expert sur l'espérance de vie générationnelle à un âge donné, ce travail présente la

¹ Ce chapitre s'appuie principalement sur l'article Planchet et Kamega [2011b], relatif à la modélisation de la mortalité prospective en présence de petits échantillons, à partir d'informations externes en utilisant l'approche de Bongaarts.

construction d'une table de mortalité prospective complète. Un exemple de mise en œuvre est présenté avec des données de la population assurée d'Afrique subsaharienne francophone.

1. Introduction

En 1971, Omran [2005] propose une théorie pour rendre compte de l'évolution de la mortalité qui accompagne la transition démographique en Europe d'abord (à partir du XVIII^{ème} siècle), et dans le reste du monde ensuite (un siècle ou deux plus tard). Cette théorie, intitulée transition épidémiologique, reconnaît trois phases dans l'évolution de la mortalité : « l'âge de la peste et de la famine » pendant lequel la mortalité est forte et fluctuante, « l'âge du recul des pandémies » pendant lequel l'espérance de vie augmente fortement et « l'âge des maladies de dégénérescence et de société » pendant lequel la mortalité se stabilise à un niveau bas et l'espérance de vie plafonne. Dans les dernières décennies du XX^{ème} siècle toutefois, la mortalité a repris sa tendance à la baisse dans la plupart des pays occidentaux, compte tenu du recul des maladies cardio-vasculaires et du contrôle des maladies de société (*cf.* par exemple Kamega [2006] pour une présentation des mesures prises en France pour lutter contre les maladies de société à partir des années 1970). Ces avancées ont conduit plusieurs auteurs à étendre la théorie de transition épidémiologique à une quatrième phase, et c'est ainsi que s'est développée la théorie de la transition sanitaire, ajoutant à la donne épidémiologique les différentes réponses de la société aux questions de santé (*cf.* par exemple Meslé et Vallin [2007]).

Meslé et Vallin [2007] avancent que ces transitions revêtent un caractère universel, bien que certains pays rencontrent de grandes difficultés à franchir certaines étapes. En effet, à titre d'exemple les pays d'Afrique subsaharienne n'ont pas franchi la seconde phase de la transition à cause du VIH/Sida et du paludisme, et les pays d'Europe centrale et de l'ex-URSS ne sont pas parvenus à contrôler les maladies de société ni à lutter efficacement contre les maladies cardio-vasculaires. En avançant le caractère universel de ces transitions, les auteurs anticipent donc une hausse générale de l'espérance de vie, mais considèrent qu'elle s'observera de manière plus ou moins prononcée selon les pays.

Si pour des raisons évidentes cette hausse attendue de l'espérance de vie peut être bien accueillie par les populations, il convient de rappeler qu'elle engendre des contraintes en termes de ressources en période de retraite. Il est donc nécessaire que la population dispose d'un financement lui permettant de se couvrir contre ce risque viager. À cet effet, la rente viagère apparaît être une solution adaptée, comme le rappellent Brouhns et Denuit [2002] par exemple.

Avec les rentes viagères, le risque viager est transféré de l'assuré à l'assureur, et ce dernier doit alors évaluer son engagement d'assurance avec des tables de mortalité appropriées (c'est d'ailleurs dans ce contexte que depuis 1993, les tables réglementaires en France pour le provisionnement des rentes viagères sont des tables de mortalité prospectives, prenant en compte l'évolution de l'espérance de vie).

En complément, avec les référentiels Solvabilité 2 (en Europe) et IFRS, les assureurs sont tenus d'évaluer leurs engagements d'assurance à partir d'hypothèses réalistes et fines. En pratique, pour les contrats de rentes viagères, les assureurs sont ainsi appelés à construire des tables de mortalité prenant en compte le niveau actuel et l'évolution attendue de la mortalité de leur population assurée (il s'agit de tables d'expérience prospectives spécifiques à chaque assureur). Cet exercice peut toutefois être complexe car (*cf.* Kamega [2011]) :

- les théories et les tendances que l'on peut observer et anticiper au sein de la population générale ne sont pas directement vérifiées au sein de la population assurée, compte tenu de l'existence d'une sélection à l'entrée qui conduit à disposer d'une population assurée différente de la population générale (et en général présentant des taux de décès inférieurs à ceux de la population générale à tous les âges) ;
- l'assureur ne peut pas directement inclure l'évolution attendue des différentes causes de décès dans ses analyses et ses anticipations car l'accès à l'information sur les causes de décès est généralement limité pour les populations assurées ;
- la construction d'une table d'expérience prospective conduit à déterminer une dérive temporelle, et nécessite à ce titre un recul historique et un volume de données que les assureurs n'ont pas dans leur portefeuille.

Généralement, deux approches sont envisagées pour construire des tables prospectives à partir de données d'expérience : une démarche endogène utilisant principalement les données du portefeuille ou une démarche exogène utilisant également des références externes de mortalité (*cf.* Planchet et Lelieur [2007]).

Aujourd'hui, la seconde approche est souvent retenue pour pallier les difficultés associées à des échantillons de taille réduite, comme souvent en ce qui concerne les portefeuilles d'expérience, même à l'échelle nationale pour un pays comme la France. En effet, à titre d'exemple pour la construction des tables d'expérience TGH/TGF05, réglementaires en France depuis 2007, les données d'expérience par génération ont été positionnées par rapport aux tables de la population générale française de 1962 à 2000 (sachant que les données d'expérience utilisées pour la construction de ces tables réglementaires comptent 700 000 rentiers et 2 000 000 de lignes et couvrent la période 1993-2005, *cf.* Planchet [2006]).

De manière générale, plus le périmètre des données d'expérience est restreint (faible recul, faible exposition, faible nombre de rentiers, etc.), plus les difficultés pour construire une loi d'expérience prospective sont nombreuses et le recours à des informations externes (positionnement par rapport à des tables de mortalité externes, calibrage de paramètres par avis d'expert, etc.) devient nécessaire.

Dans ce contexte, ce travail s'attache à proposer une démarche pour construire, sur la base d'informations externes, des lois d'expérience prospectives à partir d'un échantillon restreint (faible recul). En pratique, la

solution proposée consiste à intégrer des contraintes liées aux avis d'expert dans la modélisation des taux de décès prospectifs.

Plan

L'étude présente dans un premier temps (*cf.* section 2) les étapes du processus du choix du modèle retenu : de la présentation d'outils et de méthodes classiques pour projeter la mortalité au choix de l'approche de Bongaarts. Dans un second temps (*cf.* section 3), l'étude présente l'ajustement du modèle de Bongaarts, à partir de données statistiques d'une part et à partir de références externes (avis d'expert) d'autre part. La troisième étape (*cf.* section 4) est alors consacrée à la mise en œuvre du modèle pour anticiper la mortalité future de la population d'expérience d'Afrique subsaharienne francophone. Des tests de sensibilité y sont également réalisés dans un contexte d'assurance.

L'ensemble des travaux est conduit sur la base d'une population en « groupe fermé », arrêtée à une date d'inventaire, compte tenu du contexte de l'étude (évaluation de l'engagement de l'assureur en termes de rentes viagères).

2. Choix du modèle pour projeter la mortalité

Dans un contexte où le volume des données d'expérience est limité, les problématiques liées à l'estimation des paramètres et de la tendance future de la mortalité revêtent une importance particulière dans le choix du modèle.

En effet, d'une part, lors de l'ajustement des paramètres aux données il existe un risque que la valeur estimée diffère de la valeur correspondant au risque sous-jacent. Ce risque est directement lié aux fluctuations d'échantillonnage (*cf.* Kamega et Planchet [2010] et Planchet et Kamega [2011a] pour une présentation du risque d'estimation en présence de fluctuations d'échantillonnage) et est d'autant plus important que la taille de l'échantillon est restreinte. En conséquence, dans la présente étude, une attention particulière doit être accordée au nombre de paramètres dans le modèle.

D'autre part, la constitution de la tendance future pour la mortalité doit idéalement être le fruit de la combinaison d'un ajustement statistique aux données passées, et d'un ajustement par avis d'expert pour tenir compte du contexte futur. En effet, les projections de mortalité ne doivent pas se limiter à des approches descriptives, consistant à reproduire les tendances passées, mais à des approches prédictives, consistant à anticiper le futur. À cet effet, la prise en compte d'une dimension quantitative et qualitative est fortement souhaitable pour les projections de mortalité. En pratique toutefois, en présence d'un faible volume de données, la dimension quantitative pour déterminer les tendances futures de la mortalité ne peut pas être intégrée car on ne dispose pas des tendances passées. Les tendances futures de mortalité sont alors complètement déterminées à partir d'une dimension qualitative (à partir d'informations et de contraintes externes). Aussi, dans la présente étude, une attention particulière doit également être accordée à la prise en compte d'une dimension qualitative dans le modèle.

Dans cette section, on présente les étapes du processus du choix du modèle compte tenu, entre autres, de ces contraintes sur le nombre de paramètres et sur la prise en compte d'une dimension qualitative. Ainsi, dans un premier temps on présente les principales problématiques rencontrées dans le cadre d'une projection de la mortalité (*cf.* sous-section 2.1), puis on présente le modèle classique de Lee-Carter ainsi que ses principales limites, extensions et alternatives (*cf.* 2.2), et enfin on présente et justifie le choix du modèle de Bongaarts plutôt que celui de Lee-Carter ou l'une de ses extensions (*cf.* 2.3).

2.1. Quelques réflexions et quelques outils pour modéliser la mortalité d'expérience future

Comme indiqué en introduction, l'existence d'une sélection à l'entrée conduit à disposer d'une population assurée différente de la population générale. Ainsi, en pratique l'allongement de la durée de vie au sein de la population générale d'un pays ne se traduit pas automatiquement par un allongement analogue au sein de la population assurée d'un organisme assureur.

Une autre différence significative entre ces deux populations porte sur la profondeur des données disponibles (*cf.* Barrieu et al. [2010]) : généralement de 5 à 15 ans d'historique au sein d'une population assurée, contre plusieurs dizaines d'années au sein d'une population générale. C'est notamment pour cette raison que les actuaires modélisent généralement la tendance de la mortalité de la population assurée à partir de la tendance de la mortalité de la population générale, malgré les différences existant entre ces deux populations. Cette approche n'est toutefois pas satisfaisante et c'est dans ce contexte qu'une réflexion s'impose sur la procédure à suivre pour projeter la mortalité d'expérience.

Ainsi, cette sous-section présente quelques outils et méthodes propres à la modélisation de la mortalité future de la population assurée. En pratique, cette sous-section présente les méthodes de projection par cause (*cf.* paragraphe 2.1.1), par extrapolation (*cf.* paragraphe 2.1.2) et par avis d'expert (*cf.* paragraphe 2.1.3). Le dernier paragraphe de cette sous-section (*cf.* 2.1.4) est consacré aux méthodes de fermeture des tables.

Les risques liés à l'incertitude des projections de mortalité ne sont toutefois pas repris ici (pour information, une revue de ces risques est présentée dans Booth et Tickle [2008] et Planchet et Kamega [2011a]).

2.1.1. Hétérogénéité et projection de la mortalité par cause

Ce paragraphe présente le traitement de l'hétérogénéité et l'analyse de la mortalité par cause dans le cadre de projections.

Prise en compte des facteurs d'hétérogénéité : une démarche utile et nécessaire mais présentant des risques

Au-delà des différences entre population générale et population assurée évoquées ci-dessus, il existe des différences au sein de la population assurée : entre portefeuilles assurés d'une part, compte tenu des différences dans la

sélection, et au sein d'un même portefeuille assuré d'autre part, compte tenu des facteurs d'hétérogénéité.

L'identification et la prise en compte de ces facteurs d'hétérogénéité sont utiles et nécessaires dans le cadre d'une projection de la mortalité.

En effet, d'une part il apparaît que la modélisation du risque de décès fournit des résultats différents au niveau agrégé (c'est-à-dire en considérant la population au global) et désagrégé (c'est-à-dire en segmentant la population en sous-populations), ce qui traduit un biais d'hétérogénéité (*cf.* par exemple Dreesbeke et al. [1989]). Comme indiqué dans Kamega et Planchet [2011b], ce résultat est souvent expliqué par le phénomène « mobile-stable », selon lequel les individus du segment avec un taux de sortie élevé sortent en premier et augmentent ainsi la proportion d'individus du segment avec un taux de sortie faible lorsque le temps s'écoule. Ainsi, on comprend que dès que la répartition de la population évolue, comme souvent lors de l'évaluation d'une provision d'assurance, une prise en compte de l'hétérogénéité est indispensable, afin de disposer d'une loi d'expérience robuste.

D'autre part, la prise en compte de l'hétérogénéité peut permettre d'améliorer l'intégration d'une dimension qualitative dans les projections de mortalité. À titre d'exemple, pour affiner les hypothèses de projection pour la mortalité, Van Den Berg Jeths et al. [2001] structurent les différents facteurs d'hétérogénéité en distinguant deux catégories : celle des facteurs exogènes (consommation de tabac, consommation d'alcool, comportements alimentaires, activité physique, etc.) et celle des facteurs endogènes (prédispositions héréditaires, sexe, âge, indice de masse corporelle, etc.). Il apparaît alors que dans la catégorie des facteurs exogènes, les facteurs de risque sont très sensibles aux politiques sanitaires et de santé publique et les projections de mortalité associées peuvent ainsi naturellement intégrer une dimension qualitative (par avis d'expert). Dans la catégorie des facteurs endogènes, l'impact des politiques sur la mortalité future est plus modeste et la dimension qualitative dans les projections de mortalité est en conséquence plus limitée.

Au final toutefois, la prise en compte de l'hétérogénéité doit être appréciée en tenant compte notamment des problématiques de choix de segmentation optimale (*cf.* Planchet et Leroy [2009]) et de risque d'estimation (*cf.* chapitres sur l'hétérogénéité), qui peuvent conduire le modèle à devenir inadapté.

Prise en compte des causes de mortalité : problème de données, de cohérence et de qualité statistique

Comme indiqué en annexe de Kamega et Planchet [2011a], l'analyse des causes de mortalité pourrait être utile pour justifier les différences de mortalité entre sous-populations, même si le lien entre l'hétérogénéité et la mortalité par cause n'est pas simple : il s'agit de domaines distincts dont les relations sont complexes. Les analyses de l'hétérogénéité et de la mortalité par cause apparaissent toutefois complémentaires. En effet, à titre d'exemple, la prise en compte de l'hétérogénéité entre les fumeurs et les non fumeurs doit se traduire dans l'analyse de la mortalité par cause par un poids significatif des décès dus au cancer des poumons dans la sous-population des fumeurs.

L'analyse et la projection de la mortalité par cause présente des avantages et des inconvénients. Parmi les avantages, on note notamment qu'il est plus facile avec cette approche de tenir compte de la connaissance sur la science médicale, sur le comportement des individus ou sur les changements environnementaux (à titre d'exemple il est plus simple de prendre en compte la connaissance épidémiologique pour des causes de décès spécifiques que pour des décès toutes causes). Sur le plan qualitatif, cette démarche permet donc de disposer de projections plus structurées.

En pratique toutefois l'analyse de la mortalité par cause est peu utilisée compte tenu de l'insuffisance des données et des informations disponibles. En effet, dans un premier temps la cause de décès renseignée peut être relativement arbitraire. Par ailleurs, il est peu fréquent qu'un décès résulte d'une cause unique, et on est donc généralement confronté à des situations de multi-causalité. Sur ce point, on peut par exemple citer les études de l'Australian Bureau of Statistics (*cf.* ABS [2009]) qui illustrent qu'en 2007 3,1 causes de décès sont recensés en moyenne par décès dans le pays et que seul 20 % des décès enregistrés relève d'une cause de décès unique. Enfin, une troisième limite dans les données est la forte interaction entre les risques de décès, ce qui rend difficile l'identification de la (ou des) cause(s) exacte(s) du décès (le risque de décéder d'une grippe est différent selon que l'on vive ou non avec le VIH/Sida). Ce problème d'interdépendance des causes de décès est notamment à l'origine du problème de cohérence observée entre les projections de mortalité par cause et les projections de mortalité toutes causes (*cf.* par exemple Wilmoth [1995], qui constate que les projections de mortalité par cause conduisent à des taux de mortalité supérieurs à ceux issus de projections de mortalité globales).

Un autre inconvénient significatif de l'analyse de la mortalité par cause est la dégradation de la qualité statistique des estimations (compte tenu de la baisse du volume d'information et de la multiplication des modèles, *cf.* Alho [2007]).

C'est dans ce contexte que Tabeau et al. [2001a] indiquent que si les analyses par causes sont souhaitables pour les projections de court ou moyen terme, elles s'avèrent particulièrement risquées pour les projections de long terme qui nous intéressent ici.

2.1.2. Projection par extrapolation des taux de décès

Ce paragraphe présente les avantages et inconvénients d'une projection de la mortalité par extrapolation (selon qu'il s'agisse d'une extrapolation par période ou par génération), ainsi que les particularités de cette approche en termes d'interdépendance.

Extrapolation des taux de décès : projection par période ou par génération ?

Comme évoqué ci-dessus, les projections de mortalité toutes causes sont plus robustes sur le plan statistique que les projections par cause, mais en contrepartie elles sont moins intéressantes sur le plan qualitatif. Parmi les projections toutes causes, on distingue généralement les projections par période (dans ce cas on considère une dérive des taux de décès au fil du temps) ou par génération (dans ce cas on considère une dérive au fil des générations).

Les projections toutes causes par génération présentent l'avantage de pouvoir facilement prendre en compte l'évolution des facteurs d'hétérogénéité exogènes sur la mortalité, car dans ce cas on peut aisément intégrer la période latente entre le changement de comportement et son impact sur les taux de décès. Cette approche présente toutefois une limite importante, qui porte sur les censures généralement observées dans les données pour les personnes nées au cours des 100 dernières années. En effet, avec ces censures on ne dispose pas d'observations complètes pour certaines générations et on est usuellement amené à retenir des hypothèses qui biaisent les projections (*cf.* Tabeau et al. [2001a]).

Ce problème n'apparaît pas dans les projections toutes causes par période, ce qui est un avantage pour cette méthode alternative. Un autre argument justifiant l'intérêt de la méthode de projection par période (aux dépens de la méthode par génération) est que l'effet période sur la mortalité est généralement plus fort que l'effet génération, comme le rappellent Booth et Tickle [2008]. L'approche par période présente toutefois un inconvénient significatif par rapport à l'approche par génération : la prise en compte d'une dimension qualitative dans les projections y est plus limitée.

Au final, les démarches par période ou par génération présentées ici apparaissent donc complémentaires, chacune d'entre elles présentant des avantages et des inconvénients. C'est d'ailleurs dans ce contexte que ce sont développés les modèles APC (Age-Période-Cohorte), selon lesquels les taux de décès évoluent selon l'âge, la date (période) et la génération (*cf.* un exemple de modèle APC au paragraphe 2.2.2 avec le modèle de Renshaw et Haberman [2005]).

Traitement de l'interdépendance dans les extrapolations : interdépendance entre les paramètres, entre les âges et entre les périodes

On note que dans le cadre d'extrapolation à partir de modèle paramétrique incluant plusieurs paramètres, une attention particulière doit être accordée au traitement de l'interdépendance entre les paramètres, qu'il s'agisse d'une extrapolation par période ou par génération (sur ce point, *cf.* Booth et Tickle [2008] ou Delwarde et Denuit [2006]).

Par ailleurs, au-delà de l'interdépendance entre les paramètres évoquée ci-dessus, une réflexion s'impose également sur l'interdépendance entre les âges et entre les périodes en cas d'extrapolation.

Aussi, dans leur article sur la dépendance cachée au sein du risque de longévité, Loisel et Serant [2007] apportent un éclairage sur la mesure et l'illustration de l'interdépendance entre les âges et entre les périodes (leurs illustrations concernent les 60-90 ans observés sur la période 1969-1999 en France à partir de données issues de l'INSEE). Les auteurs apprécient dans un premier temps l'interdépendance entre les âges pour une période t . Ils considèrent à cet effet la variation des logits des taux de décès entre t et $t+1$ (variation intitulée « logit's Deltas ») et constatent alors une corrélation entre les âges. Dans un second temps, ils étudient l'interdépendance entre les périodes à partir de la représentation de la moyenne sur tous les âges des

« logit's Deltas » sur la période 1969-1999. Il y apparaît que la série constituée correspond à un processus autorégressif d'ordre 1 (dont le coefficient est significativement négatif). Il existe donc bien une interdépendance entre les périodes.

2.1.3. Projection par avis d'expert

Au-delà des approches qui consistent à projeter les causes de mortalité (cf. 2.1.1) ou à directement extrapoler la mortalité globale (par période ou par génération, cf. 2.1.2), la projection de la mortalité peut être réalisée par avis d'expert.

Généralement, la méthode retenue à cet effet consiste à poser une contrainte sur le niveau de l'espérance de vie à une date future (à l'image de ce qui est proposé dans Planchet [2007] sur la base du modèle de Lee-Carter). Booth et Tickle [2008] présentent toutefois plusieurs démarches alternatives : avis d'expert sur les tendances des taux de décès par causes (et par âge), avis d'expert sur la vitesse de réduction des taux de mortalité par âge, etc.

L'approche par avis d'expert permet d'intégrer, au moins sur un plan qualitatif, la connaissance épidémiologique, démographique, environnementale, etc. dans les projections. En outre, elle présente également un intérêt particulier lorsque les données d'expérience disponibles sont limitées (comme dans le cadre de la présente étude). L'inconvénient de cette approche est sa subjectivité et son biais potentiel. Ainsi, Booth et Tickle [2008] rappellent que le conservatisme et les divergences dans les avis d'expert ne permettent pas d'envisager des projections de long terme avec cette méthode. En outre, les résultats présentés par les auteurs illustrent que les projections par avis d'expert aux États-Unis, au Royaume-Uni et en Australie ont eu tendance à sous-estimer à terme l'amélioration de la mortalité, que ce soit par rapport aux projections par extrapolation ou par rapport aux observations *ex post*.

L'approche consistant à projeter la mortalité par avis d'expert apparaît ainsi nécessaire en cas de d'échantillons restreints, mais présente des limites importantes pour des projections de long terme. Cette approche est par ailleurs complémentaire aux deux autres.

2.1.4. Fermeture de la table

Quelle que soit l'approche retenue au final, la question du traitement de la mortalité aux grands âges se pose. En effet, au-delà de 90 ans, on dispose rarement de données de bonne qualité (du moins dans un contexte d'assurance). Il convient alors de trouver des solutions pour extrapoler la mortalité aux grands âges, avant ou après l'ajustement : on parle de « fermeture de table de mortalité ».

Ces questions sur la fermeture de la table soulèvent le problème du comportement de la mortalité aux grands âges : existence et niveau d'un éventuel âge limite (biologiquement maximal), décroissance de la croissance des taux de décès au-delà d'un certain âge, etc. (cf. Delwarde et Denuit [2006]). De nombreux modèles ont ainsi été développés par des démographes et des actuaires pour tenir compte de ces réalités. Quashie et Denuit [2005] présentent et comparent plusieurs de ces modèles.

Comme le rappellent Planchet et Thérond [2006], la question de la fermeture de la table est importante dans le cas de la construction d'une table pour des provisionnements de rentes viagères, mais cette importance doit être relativisée si les rentiers d'âge très élevé sont en proportion modeste dans le portefeuille. En effet, à titre d'exemple, les auteurs illustrent qu'en figeant les taux de décès de la table TH00-02 à 95 ans (en provoquant alors une sortie brutale à 120 ans), l'impact sur la provision à 75 ans est inférieur à 1 % (impact inférieur à 3 % sur une provision à 85 ans). De même, Quashie et Denuit [2005] illustrent que l'écart entre deux méthodes de fermeture en termes de provisionnement n'est véritablement significatif qu'à des âges très élevés.

2.2. Modèle de Lee-Carter : limites, extension et alternatives

La sous-section précédente présente les avantages et inconvénients de quelques outils et méthodes pour modéliser la mortalité future, dans le contexte de l'étude. Cette nouvelle sous-section présente désormais les caractéristiques de l'approche classique de Lee-Carter pour modéliser la mortalité d'expérience future. En particulier, cette sous-section s'attache à mettre en perspective les problématiques formulées dans l'analyse des outils et méthodes existants avec les propriétés de l'approche de Lee-Carter.

2.2.1. Présentation du modèle de Lee-Carter

Le modèle de Lee-Carter, aujourd'hui un standard, a déjà fait ses preuves de maintes reprises en démographie. Il a notamment prouvé son efficacité sur diverses populations (USA, Canada, Chili, Japon, Belgique, Suède, etc., cf. références associées dans Delwarde et Denuit [2006]).

En termes de méthodes, la démarche de Lee-Carter consiste à estimer les taux de mortalité futurs à partir d'une extrapolation des tendances passées, après avoir décomposé la mortalité en deux composantes, l'une propre à l'âge et l'autre tendancielle. Il s'agit ainsi d'une projection à partir d'une extrapolation par période. Le modèle s'écrit (cf. Lee [2000]) :

$$\ln(\mu_{x,t}) = \alpha_x + \beta_x k_t + \varepsilon_{x,t},$$

où α_x s'interprète comme la valeur moyenne des $\ln(\mu_{x,t})$ au cours du temps, k_t représente l'évolution générale de la mortalité au cours du temps, β_x traduit la sensibilité (à l'échelle logarithmique) de la mortalité instantanée à l'âge x par rapport à k_t et $\varepsilon_{x,t} \sim N(0; \sigma_\varepsilon)$ représente l'erreur du modèle (sous l'hypothèse d'homoscédasticité). Afin de rendre le modèle identifiable, il convient par ailleurs d'ajouter deux contraintes sur les paramètres ; on retient généralement :

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

Les paramètres du modèle sont estimés par une méthode de moindres carrés.

Il convient de noter que le modèle de Lee-Carter s'apparente donc un modèle semi-paramétrique dans lequel la forme de la mortalité n'est pas contrainte.

2.2.2. Limites, extensions et alternatives au modèle de Lee-Carter

Ce paragraphe présente quelques limites du modèle de Lee-Carter, ainsi que des extensions ou des alternatives envisageables pour y remédier.

Traitement de l'hétéroscédasticité : l'approche de Brouhns et al. [2002] (log-Poisson)

Une première limite du modèle de Lee-Carter est l'hypothèse d'homoscédasticité des erreurs requise pour l'estimation par moindres carrés. En effet, cette condition est rarement vérifiée en pratique puisque la variance du logarithme du taux instantané de mortalité croît aux âges élevés, du fait notamment de la baisse des effectifs survivants à ces âges.

Pour pallier à cette limite, Brouhns et al. [2002] ont proposé une adaptation consistant à modéliser le nombre de décès à l'âge x durant l'année t à l'aide d'une loi de Poisson. Ils modélisent ainsi $D_{x,t}$, qui suit une loi de Poisson de paramètre $R_{x,t}\mu_{x,t}$, où $R_{x,t}$ représente l'exposition au risque à l'âge x durant l'année t et $\mu_{x,t} = \exp(\alpha_x + \beta_x k_t)$ (l'expression du taux de décès instantané est donc identique à celle du modèle de Lee-Carter, avec les mêmes interprétations des paramètres et les mêmes contraintes). Comme le rappellent Delwarde et Denuit [2006], cette nouvelle spécification, qui recourt au principe du maximum de vraisemblance pour estimer les paramètres, permet de lever l'hypothèse d'homoscédasticité et reconnaît la variabilité plus grande des taux bruts aux âges élevés.

Extension de la dimension temporelle : l'approche de Renshaw et Haberman [2003]

L'approche de Lee-Carter classique consiste à extrapoler dans le futur des tendances constatées dans le passé, en ne faisant intervenir qu'un seul paramètre temporel. Ainsi, l'approche de Lee-Carter dispose d'une dimension qualitative très limitée et peut en conséquence se révéler trop rigide dans certaines situations. En effet, Renshaw et Haberman [2003] constatent par exemple que le modèle classique de Lee-Carter ne permet pas de rendre compte de l'impact des maladies de société (suicide notamment), ni de l'effet du VIH/Sida dans la mortalité des 20-39 en Angleterre et au Pays de Galles au cours des dernières décennies du XX^{ème} siècle. Il s'agit d'une seconde limite pour l'approche de Lee-Carter.

Pour pallier cette limite, Renshaw et Haberman [2003] propose de rajouter une dimension temporelle au modèle de Lee-Carter classique pour projeter la mortalité. Le modèle s'écrit alors :

$$\ln(\mu_{x,t}) = \alpha_x + \beta_x^1 k_t^1 + \beta_x^2 k_t^2 + \varepsilon_{x,t},$$

avec des notations analogues à celle du modèle de Lee-Carter classique. Ce nouveau dispose alors d'une dimension qualitative plus importante.

Prise en compte de la génération : l'approche de Renshaw et Haberman [2005]

Comme évoqué ci-dessus, l'approche de Lee-Carter classique présente une faible dimension qualitative et la prise en compte d'un avis d'expert dans l'évolution présumée de la mortalité (au titre de l'accès aux soins, des progrès de la médecine, de l'apparition ou de la suppression de maladies, de l'évolution du style de vie, etc.) est donc limitée².

L'alternative de Renshaw et Haberman [2003], qui consiste à améliorer la dimension qualitative du modèle en augmentant la dimension temporelle, est une première approche pour limiter cette insuffisance (*cf.* ci-dessus). En complément, toujours afin d'améliorer la dimension qualitative, Renshaw et Haberman [2005] ont étendu le modèle de Lee-Carter classique afin de tenir compte d'un effet de génération. À cet effet, ils considèrent que l'expression des taux de décès instantanés s'écrit :

$$\mu_{x,t} = \exp\left(\alpha_x + \beta_x^0 \gamma_{t-x} + \beta_x^1 k_t\right),$$

où l'effet génération est représenté par $\beta_x^0 \gamma_{t-x}$ (la génération étant symbolisée par $z = t - x$). Il s'agit d'un modèle APC (Age-Période-Cohorte), dans lequel la mortalité est sensible à l'âge, à la période et à la génération.

On remarque que si $\beta_x^0 = 0$ on retrouve l'expression classique du modèle de Lee-Carter (prenant en compte uniquement les effets d'âge et de période), et d'un autre côté si $\beta_x^1 = 0$ l'on dispose d'un modèle ne prenant en compte que les effets d'âge et de génération.

Calibrage en présence de données limitées : quelques alternatives à Lee-Carter et à ses extensions

Le modèle de Lee-Carter et ses extensions s'appuient, entre autres, sur le calibrage de la tendance pour ensuite l'extrapoler pour les années futures. En particulier, ces modèles sont construits pour un contexte de gros volume de données dans lequel on suppose qu'il n'y a pas de fluctuations d'échantillonnage significative à un âge fixé. Pour toutes ces raisons, un historique de données important est nécessaire. Si à l'échelle d'une population nationale le volume de données nécessaire est généralement disponible, à l'échelle d'une population assurée le volume de données se révèlent généralement insuffisant pour structurer correctement la table complète. En outre, le nombre élevé de paramètres, notamment pour les extensions au modèle de Lee-Carter, et le faible volume de données peut conduire à des

² Dans d'autres contextes, la faible dimension qualitative du modèle de Lee-Carter peut toutefois être présenté comme un avantage (comme le font par exemple Delwarde et Denuit [2006]), notamment lorsque les avis d'experts divergent significativement.

irrégularités importantes au titre des fluctuations d'échantillonnage (ce phénomène est illustré dans Planchet et Lelieur [2007]).

Pour faire face à ces limites, il peut être utile de se tourner vers des modèles alternatifs moins paramétrés, mettant en jeu des expressions analytiques portant sur les âges et/ou les années. À titre d'exemple, on peut s'appuyer sur des régressions linéaires simples des logits des taux bruts (qui présentent l'avantage de varier entre $-\infty$ et $+\infty$; en outre, on constate empiriquement que le logit des taux de décès est approximativement affine au-delà de 25 ans). Dans ce cas, deux paramètres sont estimés.

Ainsi, en régressant sur l'année calendaire notée t (pour un âge x fixé), on peut considérer le modèle alternatif suivant :

$$\ln\left(q_{x,t}/(1-q_{x,t})\right) = \alpha_x + \beta_x \times t + \varepsilon_{x,t}.$$

Par ailleurs, en régressant sur les logits des taux de la table de référence notés $\ln\left(q_{x,t}^{ref}/(1-q_{x,t}^{ref})\right)$, on peut également considérer le modèle alternatif suivant :

$$\ln\left(q_{x,t}/(1-q_{x,t})\right) = \alpha + \beta \times \ln\left(q_{x,t}^{ref}/(1-q_{x,t}^{ref})\right) + \varepsilon_{x,t}.$$

Cette deuxième démarche permet au surplus une extrapolation aisée des logits des taux d'expérience dans les plages d'âge pour lesquelles les données d'expérience seraient insuffisantes.

Pour information, ces deux exemples sont détaillés dans Planchet et Thérond [2006].

Sensibilité de la mortalité au cours du temps : quelques alternatives à Lee-Carter

Une autre limite importante du modèle de Lee-Carter est l'hypothèse que la sensibilité de la mortalité instantanée à l'âge x par rapport à k_t , notée β_x (avec $d \ln(\mu_{x,t})/dt = \beta_x \times dk_t/dt$, soit $\beta_x = d \ln(\mu_{x,t})/dk_t$), est constante au cours du temps (*cf.* Planchet et Thérond [2006]). Cette hypothèse revient à considérer qu'à un âge donné, l'amélioration de la mortalité au cours du temps est la même, quelles que soient les dates de la période considérée. Cette contrainte n'est pas vérifiée en pratique. En effet, à titre d'exemple Bongaarts [2004] illustre sur la base de la population féminine suédoise que l'amélioration de la mortalité à 75 ans sur la période 1950-2000 est plus forte que celle observée sur les périodes 1850-1900 et 1900-1950. De manière générale, l'auteur rappelle que l'amélioration de la mortalité au cours du temps a tendance à diminuer dans le temps pour les plus jeunes, et qu'elle a tendance à augmenter dans le temps pour les plus âgés (ce qui est cohérent avec les théories de transitions évoquées en introduction).

L'approche de Bongaarts [2004], présentée ci-après en 2.3, permet, entre autres, de relâcher cette contrainte (*cf.* notamment en 2.3.2).

2.3. Approche de Bongaarts : justification du choix

Le modèle de Lee-Cater présente l'avantage d'avoir fait ses preuves à maintes reprises dans la modélisation de la mortalité future et de disposer de bonnes propriétés, ce qui en fait un standard.

Cette approche présente toutefois quelques limites. Certaines sont « universelles » (problème d'hétéroscédasticité, constance de la sensibilité de la mortalité au cours du temps), et d'autres sont principalement liées au contexte de l'étude sur de petits échantillons (faible dimension qualitative, grand nombre de paramètres – notamment pour les extensions du modèle –).

Il convient alors de retenir au final un modèle qui ne reprenne pas ces limites, tout en y gardant les principaux avantages (simplicité opérationnelle, facilité d'interprétation, peu de paramètres – dans l'approche classique –, etc.). Aussi, cette sous-section présente et justifie le choix de l'approche de Bongaarts.

2.3.1. Présentation de l'approche de Bongaarts

On considère ici le modèle logistique décalé de Bongaarts [2004] défini par :

$$\mu_{x,t} = \frac{\alpha_t e^{\beta x}}{1 + \alpha_t e^{\beta x}} + \gamma_t,$$

où les paramètres α_t et γ_t sont dépendants du temps, et le paramètre β est indépendant du temps (Bongaarts [2004] constatant empiriquement que le paramètre β dépend peu du temps).

Dans cette spécification des taux de décès instantanés, on reconnaît facilement une adaptation du modèle paramétrique proposé par Thatcher [1999] dans laquelle les taux de décès ne sont pas constants au cours du temps³. On retrouve ainsi la dimension explicative du modèle, dans laquelle Bongaarts [2004] considère que la composante liée au taux instantané de décès environnemental (notée $(\mu_{x,t})_e = \gamma_t$) et celle liée au vieillissement (notée

$(\mu_{x,t})_v = \frac{\alpha_t e^{\beta x}}{1 + \alpha_t e^{\beta x}}$) dépendent du temps. Pour mémoire, le paramètre γ_t

représente généralement l'ajustement de la mortalité pour les jeunes adultes, le paramètre α_t représente le niveau de mortalité et le paramètre e^{β} représente l'augmentation du risque de mortalité avec l'âge.

Concernant la composante des taux de décès liée au vieillissement, Bongaarts relève en particulier que l'évolution à la hausse ou à la baisse de la mortalité au cours du temps peut être interprétée comme un décalage d'âge au

³ Pour mémoire, le modèle de Thatcher correspond lui-même à une adaptation du modèle classique de Makeham, ajusté pour corriger la surestimation des taux de décès conditionnels aux âges élevés observée en pratique.

cours du temps. Cette interprétation est possible car il montre que la composante vieillesse de la mortalité à l'âge x pour l'année t est identique à la valeur de la composante à l'âge $x - S_{t_0,t}$ pour l'année t_0 , soit :

$$(\mu_{x,t})_v = \frac{\alpha_{t_0} e^{\beta(x-S_{t_0,t})}}{1 + \alpha_{t_0} e^{\beta(x-S_{t_0,t})}},$$

avec $S_{t_0,t} = -\frac{\ln(\alpha_t/\alpha_{t_0})}{\beta}$. Cette propriété, qui n'est toutefois pas valable pour les âges faibles et proches de 0, est à l'origine de l'expression « modèle logistique décalé ».

Sur la base de ce modèle, Bongaarts [2004] propose une procédure en quatre étapes pour anticiper la mortalité future :

- 1) Calibrer le modèle de Thatcher $\mu_x = \frac{\alpha e^{\beta x}}{1 + \alpha e^{\beta x}} + \gamma$ pour différentes périodes ou années passées t (ajustement statistique). Ce calibrage permet de disposer de séries pour les paramètres α_t , β_t et γ_t .
- 2) Fixer la valeur du paramètre β en considérant la moyenne de β_t , et déterminer à nouveau les séries pour les paramètres α_t et γ_t (les nouvelles séries ne devant différer que légèrement de celles issus de la première étape).
- 3) Extrapoler les paramètres α_t et γ_t obtenus dans l'étape 2 sur la durée de projection souhaitée.
- 4) En déduire une projection des taux de mortalité à partir du modèle de Bongaarts, sur la base du paramètre β et des paramètres extrapolés α_t et γ_t .

2.3.2. Justification de l'approche de Bongaarts

L'approche de Bongaarts [2004] consiste donc à ajuster les observations de chaque année à l'aide d'un modèle paramétrique (en l'occurrence ici le modèle de Thatcher), puis à modéliser l'évolution au cours du temps de ces paramètres afin de fournir des projections de la mortalité dans le futur. Cette démarche présente de nombreux atouts au regard des principes et caractéristiques évoquées ci-dessus. Ainsi, l'approche de Bongaarts :

- permet de ne pas majorer le nombre de paramètres du modèle de Lee-Carter classique : l'information annuelle est condensée dans un petit nombre de paramètres ;
- est moins contrainte dans la dimension temporelle que le modèle de Lee-Carter classique, étant donné qu'elle fait intervenir deux paramètres temporels (contre un seul dans Lee-Carter) ;

- considère que l'amélioration de la mortalité au cours du temps évolue dans le temps (contrairement au modèle de Lee-Carter). En effet, Bongaarts [2004] montre que dans son modèle l'amélioration de la mortalité au cours du temps évolue au fil des années (il considère notamment que cette amélioration est égal au produit du taux d'évolution de la mortalité selon l'âge et de la sensibilité au temps de l'espérance de vie, ces deux composantes évoluant dans le temps) ;
- fournit une dimension qualitative à la modélisation, d'une part compte tenu de la forme paramétrique du modèle sous-jacent (modèle de Thatcher, avec un paramètre au titre de la mortalité environnementale, un paramètre au titre du niveau de mortalité et un paramètre au titre l'augmentation du risque de décès avec l'âge), et d'autre part compte tenu de la propriété de « décalage » du modèle de Bongaarts (qui permet d'interpréter l'évolution de la mortalité au cours du temps comme un décalage d'âge) ;
- accorde une attention particulière à la modélisation de la mortalité aux grands âges (compte tenu de son lien étroit avec le modèle de Thatcher).

De surcroît, Bongaarts [2004] précise que son approche est adaptée à des populations pour lesquelles on ne dispose qu'une d'une (ou de très peu) d'année(s) d'historique, ce qui est souvent le cas dans les pays non industrialisés (les données dans ces pays étant souvent limitées). Dans ce cas, Bongaarts [2004] indique que la projection des paramètres α_t et γ_t se fait alors uniquement à partir d'hypothèses et que son approche devient de fait principalement qualitative.

3. Ajustement du modèle de Bongaarts

Après avoir présenté et justifié le choix de l'approche de Bongaarts dans la section précédente, il convient ici d'en effectuer un ajustement. L'ajustement est réalisé à partir des données individuelles d'une part (ajustement statistique, *cf.* 3.1) et à partir d'informations externes d'autre part (ajustement par avis d'expert, *cf.* 3.2).

3.1. Ajustement statistique

Cette sous-section présente l'ajustement statistique des paramètres, réalisé dans le cadre de l'étape 1 de l'approche de Bongaarts [2004]. Dans un premier temps, on présente la méthode statistique retenue à cet effet (*cf.* 3.1.1), puis on présente dans un second temps le test statistique retenu (*cf.* 3.1.2).

3.1.1. Méthode : ajustement par maximum de vraisemblance (et moindres carrés)

La méthode retenue est l'ajustement par (pseudo) maximum de vraisemblance discret (méthode de De Vylder, *cf.* Delwarde et Denuit [2006] ou Planchet et Thérond [2006]).

De la méthode du maximum de vraisemblance à un problème de moindres carrés

Soit d_x le nombre de décès observé à l'âge x , parmi R_x individus soumis au risque de décès à cet âge. On considère alors le nombre de décès observé d_x comme une réalisation de variables aléatoires D_x de la loi binomiale de paramètres $(R_x; q_x)$, où q_x correspond à la probabilité annuelle de décès théorique à l'âge x . La vraisemblance associée à la réalisation d'un nombre d_x de décès s'écrit alors :

$$\binom{R_x}{d_x} q_x^{d_x} (1-q_x)^{R_x-d_x},$$

et on cherche alors à maximiser :

$$\sum_x d_x \ln(q_x) + (R_x - d_x) \ln(1-q_x).$$

En pratique toutefois, on réalise généralement l'approximation de la loi de \hat{q}_x (estimation des taux de décès basée sur l'échantillon) par une loi normale, et en remplaçant la variance théorique par la variance estimée le problème de maximisation ci-dessus peut être valablement approximé par un problème de minimisation de (cf. Delwarde et Denuit [2006] ou Planchet et Thérond [2006]) :

$$\sum_x \frac{R_x}{\hat{q}_x (1-\hat{q}_x)} (q_x - \hat{q}_x)^2.$$

Dans ce cas, la méthode de vraisemblance est ramenée, pour chaque année (ou période), à un problème de moindres carrés pondérés dans le cas non linéaire.

Évaluation du taux de décès théorique

Les quantités R_x et \hat{q}_x étant observées (en particulier, \hat{q}_x est déterminé à partir de l'estimation des taux bruts de Hoem, cf. une description de la méthode dans Planchet et Thérond [2006]), il reste à déterminer l'expression de q_x (taux de décès théorique) pour résoudre le problème de moindres carrés pondérés. On considère à cet effet :

$$q_x = 1 - \exp\left(-\int_x^{x+1} \mu_u du\right) = 1 - \exp\left(-\int_x^{x+1} \left(\frac{\alpha e^{\beta u}}{1 + \alpha e^{\beta u}} + \gamma\right) du\right).$$

En posant $v_u = 1 + \alpha e^{\beta u}$ et en intégrant par changement de variable (à l'image de ce qui est proposé dans Planchet [2010]), on obtient :

$$q_x = 1 - \exp\left(-\int_x^{x+1} \frac{v'_u}{\beta v_u} du\right) \exp(-\gamma) = 1 - \exp\left(-\frac{1}{\beta} \int_x^{x+1} \frac{1}{v} dv\right) \exp(-\gamma),$$

et on en déduit :

$$q_x = 1 - \exp(-\gamma) \left(\frac{v_x}{v_{x+1}} \right)^{1/\beta},$$

avec $v_u = 1 + \alpha e^{\beta u}$. C'est à l'aide de cette expression que le problème d'optimisation est résolu.

Détermination des valeurs initiales des paramètres

L'algorithme utilisé pour l'estimation des paramètres ne converge vers la vraie valeur du paramètre qu'à condition de partir d'une valeur initiale assez proche de cette valeur. Une attention particulière doit ainsi être accordée à la détermination de valeurs initiales acceptables.

À cet effet, on s'appuie sur une propriété présentée dans Planchet et Thérond [2006], sous l'hypothèse que les taux suivent une loi de Makeham. Cette hypothèse est acceptable ici étant donné qu'en pratique le calibrage est réalisé sur les âges compris entre 30 et 55 ans (*cf.* section 4 ci-après), et que pour ces âges le modèle de Thatcher est assimilable au modèle de Makeham, étant donné que $\alpha e^{\beta x}$ est généralement petit (il s'agit de la première des deux propriétés importantes soulevées par Thatcher [1999] dans son article).

Avec le modèle de Makeham, on a :

$$q_x = 1 - \exp\left(-\int_x^{x+1} (\alpha e^{\beta u} + \gamma) du\right),$$

soit au final :

$$q_x = 1 - e^{-\gamma} \left[\exp\left(-\frac{\alpha}{\beta}\right) \right]^{e^{\beta x} (e^{\beta} - 1)}.$$

On observe alors que l'on a :

$$\ln(1 - q_x) = -\gamma + e^{\beta x} (e^{\beta} - 1) \times \left(-\frac{\alpha}{\beta}\right).$$

Pour les q_x proches de zéro, on peut faire l'approximation $\ln(1 - q_x) \approx -q_x$ et

on a donc $-q_x \approx -\gamma - \frac{\alpha}{\beta} e^{\beta x} (e^{\beta} - 1)$. Il en résulte que

$q_{x+1} - q_x \approx \frac{\alpha}{\beta} e^{\beta x} (e^{\beta} - 1)^2$, ce qui conduit à remarquer en prenant le logarithme

que :

$$\ln(q_{x+1} - q_x) \approx \ln(e^{\beta}) \times x + \ln\left(\frac{\alpha}{\beta} (e^{\beta} - 1)^2\right).$$

Sous l'hypothèse que sur les âges retenus pour l'ajustement, les taux de mortalité suivent une loi de Makeham, les points $(x, y = \ln(q_{x+1} - q_x))$ sont donc alignés sur une droite de pente β . Il s'agit ainsi de réaliser une régression linéaire classique, par moindres carrés, pour en déduire des valeurs initiales pour les paramètres du modèle de Thatcher : la pente de la droite et l'ordonnée à l'origine permettent de déterminer une estimation de β et α ; et la relation

$\gamma = -\ln(1 - q_x) - \frac{\alpha}{\beta} e^{\beta x} (e^\beta - 1)$ permet de déterminer une estimation de γ (en vérifiant que $-\ln(1 - q_x) - \frac{\alpha}{\beta} e^{\beta x} (e^\beta - 1)$ dépend faiblement de l'âge x ⁴).

3.1.2. Test statistique

Dans le cadre de cette étude, on se limite pour les illustrations au test de significativité globale des paramètres. On suppose ici que ce test est réalisé à partir de la statistique du rapport des vraisemblances qui suit une distribution du Khi-deux sous l'hypothèse nulle H_0 (cette statistique compte parmi les plus utilisées dans ce contexte, avec la statistique de Wald, cf. par exemple Klein et Moeschberger [2005] pour une présentation illustrée).

On considère donc $H_0 : \delta = 0$, soit :

$$\chi_L^2(p) = 2 \left[L(\hat{\delta}) - L(0) \right],$$

où δ représente le vecteur des paramètres (α , β et γ ici) et p représente la dimension δ .

3.2. Ajustement par avis d'expert

L'étape 1 de l'approche de Bongaarts préconise de réaliser l'ajustement statistique pour différentes années ou périodes passées. Toutefois, dans le cadre de cette étude on considère des données limitées et insuffisantes pour réaliser l'ajustement statistique pour différentes années (ou périodes). Aussi, on considère ici, dans le cadre de l'étape 1 de l'approche de Bongaarts, un ajustement statistique pour une année (ou période) de référence uniquement. L'étape 2 de l'approche de Bongaarts relative au paramètre β_t est alors immédiate (puisque l'on a $\beta_t = \beta$). La projection des paramètres γ_t et α_t pour l'étape 3 est ensuite réalisée par avis d'expert. À cet effet, on distingue la projection du paramètre γ_t et celle du paramètre α_t .

Ajustement de la mortalité environnementale

Pour le paramètre γ_t , représentant l'ajustement au titre de la mortalité environnementale (indépendant de l'âge mais concernant principalement les

⁴ En pratique, on pourra considérer la moyenne pour tous les âges x de l'ajustement.

jeunes adultes), il sera supposé qu'il est constant tout au long de la durée de projection. On a ainsi $\gamma_t = \gamma$.

Ajustement du niveau de la mortalité : première approche

Pour le paramètre α_t , représentant le niveau de mortalité, deux approches sont envisagées pour introduire l'avis d'expert.

La première consiste à utiliser les caractéristiques du modèle logistique décalé. Dans ce cas, pour chaque t futur, on déduit les α_t à partir des décalages d'âge $S_{t_0,t}$ fixés par avis d'expert. Cette approche apparaît toutefois faiblement appréhensive, ce qui limite une prise en compte optimale de l'avis d'expert.

Ajustement du niveau de la mortalité : seconde approche (présentation)

On se tourne alors vers une seconde approche, qui consiste à utiliser l'espérance de vie générationnelle à une date t future et à un âge pivot x donné, à l'image de ce qui est proposé dans Planchet [2007]. Cette notion, notée $e_{x,t}$, est intuitivement plus aisée à intégrer et se calcule par :

$$e_{x,t} = \sum_{h>0} \prod_{k=0}^{h-1} (1 - q_{x+k,t+k}),$$

où h et k sont des entiers. En reprenant l'expression des q_x du modèle retenu, on obtient :

$$e_{x,t} = \sum_{h>0} \exp(-\gamma h) \prod_{k=0}^{h-1} \left(\frac{v_{x+k,t+k}}{v_{x+1+k,t+k}} \right)^{1/\beta},$$

avec $v_{u,t} = 1 + \alpha_t e^{\beta u}$. On remarque ici que l'on suppose que la répartition des décès entre x et $x+1$ s'observe à une même date t : on considère ainsi une évolution en escalier des taux de décès au fil du temps.

Dans le cadre de cette seconde approche, on pose ensuite une expression simple pour projeter α_t , en fonction de t et de deux paramètres a_α et b_α :

- soit $\alpha_t = a_\alpha \times t + b_\alpha$, si l'on anticipe une évolution linéaire du niveau de mortalité ;
- ou $\alpha_t = \exp(a_\alpha \times t + b_\alpha)$, si l'on anticipe une évolution exponentielle du niveau de mortalité.

Quelques indications pour faciliter le choix de l'expert entre ces deux expressions pour la projection du niveau de mortalité sont présentées dans le sous-paragraphe suivant.

On en déduit alors facilement une expression de la forme $e_{x,t} = \varphi_{x,t}(a_\alpha, b_\alpha)$ pour l'espérance de vie générationnelle. On se ramène au final à chercher $\theta = (a_\alpha, b_\alpha)$ minimisant l'écart quadratique entre la valeur à dire d'expert pour l'espérance de vie générationnelle (notée $e_{x,t}(\text{expert})$ à un âge pivot x et à une date t) et la prédiction de cette valeur selon l'expression de $e_{x,t}$ ci-dessus. On cherche ainsi à résoudre :

$$\underset{(a_\alpha, b_\alpha)}{\text{Min}} \left[\left(e_{x,t}(\text{expert}) - \varphi_{x,t}(a_\alpha, b_\alpha) \right)^2 \right],$$

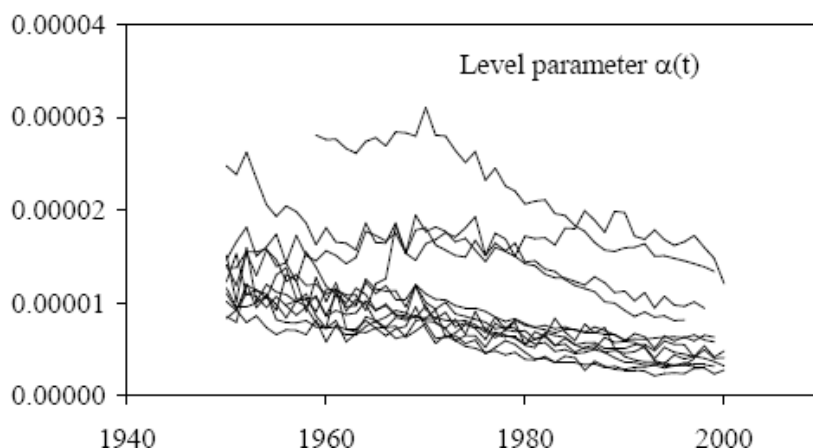
avec $\varphi_{x,t}(a_\alpha, b_\alpha) = e_{x,t}$.

On note qu'ici l'avis de l'expert s'observe à deux niveaux : d'une part pour le choix de l'expression pour la projection du niveau de mortalité α_t (cf. détails dans le sous-paragraphe suivant), et d'autre part pour le choix de l'espérance de vie générationnelle $e_{x,t}(\text{expert})$.

Ajustement du niveau de la mortalité : seconde approche (choix de l'expression pour la projection du niveau de mortalité)

En pratique, l'expression de la projection du niveau de mortalité α_t dépend de la forme attendue pour la courbe d'évolution du paramètre α_t et de la durée de projection, sachant que dans tous les cas on doit raisonnablement s'attendre à une décroissance du paramètre compte tenu du contexte de hausse attendue de l'espérance de vie (cf. introduction de ce chapitre).

Pour information, dans le cadre de l'étape 1, Bongaarts [2004] obtient les allures suivantes pour l'évolution du paramètre α_t de 1950 à 2000 au titre de la population générale féminine de 14 pays industrialisés. Il apparaît bien, une tendance à la baisse du niveau de mortalité.

Figure 39 – Évolution du niveau de mortalité (extrait de Bongaarts [2004])Figure 3a: Estimates of level parameter α in the logistic model for 14 countries, females, 1950-2000

Selon la durée de projection et l'allure attendue pour la prédiction du paramètre α_t , l'expression de la projection du niveau de mortalité peut s'assimiler à une décroissance linéaire ou à une décroissance exponentielle.

D'un côté, si l'on anticipe que la baisse absolue du niveau de mortalité sera constante dans les prochaines années, il convient de retenir une décroissance linéaire. Dans ce cas α_t s'exprime simplement par :

$$\alpha_t = a_\alpha \times t + b_\alpha,$$

et on dispose alors d'une suite arithmétique de valeur initiale $\alpha_0 = b_\alpha$ et de raison a_α (avec $a_\alpha < 0$). La limite de cette approche est que pour des horizons de projection importants, α_t peut mathématiquement devenir négatif (ce qui n'est pas souhaitable en pratique). Cette approche peut ainsi être envisagée lorsque l'on considère que la baisse du niveau de mortalité s'effectuera à un rythme constant (en valeur absolue) dans le futur et que la projection est réalisée sur un horizon mesuré.

D'un autre côté, si l'on anticipe que c'est plutôt la baisse relative du niveau de mortalité qui sera constante dans les prochaines années, il convient de retenir une décroissance exponentielle. En effet, dans ce cas on considère une diminution de α_t proportionnelle à son niveau, et on pose donc $\frac{d\alpha_t}{dt} = a_\alpha \times \alpha_t$ (avec $a_\alpha < 0$), soit :

$$\alpha_t = \alpha_0 \times \exp(a_\alpha \times t).$$

Au final, on dispose ainsi d'une suite géométrique de valeur initiale α_0 et de raison $\exp(a_\alpha)$, et en posant $\alpha_0 = \exp(b_\alpha)$, on a :

$$\alpha_t = \exp(a_\alpha \times t + b_\alpha).$$

Cette expression présente l'avantage de rendre mathématiquement impossible l'obtention de niveaux de mortalité négatifs. Cette approche peut ainsi être envisagée lorsque l'on considère que la baisse du niveau de mortalité diminuera (en valeur absolue) dans le futur et est parfaitement adaptée aux horizons de projection importants.

4. Application : anticipation de la mortalité d'expérience future

Cette section s'attache à mettre en application la démarche proposée ci-dessus pour modéliser la mortalité prospective en présence de données limitées.

En particulier, les applications s'inscrivent dans le cadre de la modélisation de la mortalité prospective en Afrique subsaharienne francophone et s'appuient à ce titre sur les données d'expérience utilisées pour la construction des tables de mortalité réglementaires dans la zone CIMA/FANAF pour les contrats d'assurance en cas de vie et en cas de décès. Ces données ont été collectées en 2009, sont observées sur les années 2003 à 2006 et comptent des troncatures à gauche (relatives aux entrées après le 01/01/2003) et des censures à droite (relatives aux sorties avant le 31/12/2006 pour un motif autre que le décès). Ces données, qui sont exploitables uniquement sur les 30-55 ans, sont présentées en détail dans le deuxième chapitre.

Dans le cadre de cette étude, on limite l'exploitation de ces données à la population masculine de la Côte d'Ivoire, ce qui permet de répondre à la volonté de tenir compte de l'hétérogénéité au sein de la zone étudiée (zone qui pour mémoire comprends plusieurs pays, *cf.* chapitres précédents sur l'hétérogénéité).

4.1. Résultats de l'ajustement du modèle

Cette sous-section présente les résultats des ajustements statistiques (*cf.* 4.1.1) et par avis d'expert (*cf.* 4.1.2), ainsi qu'une illustration des projections des taux de mortalité (*cf.* 4.1.3).

4.1.1. Résultats de l'ajustement statistique

L'ajustement statistique est réalisé au titre des points 1 et 2 de l'approche de Bongaarts. Les statistiques descriptives relatives aux données utilisées sont présentées dans le tableau suivant, pour chaque année d'observation de la population assurée.

Tableau 38 – Statistiques ventilées par année (Côte d’Ivoire - Homme)

Homme Côte d’Ivoire (pop. ass.)	Exposition	Âge moyen	Taux décès moyen	Taux décès moyen (borne inf. à 95%)	Taux décès moyen (borne sup. à 95%)
2003	163 755	43,7 ans	0,36%	0,33%	0,38%
2004	161 056	44,0 ans	0,40%	0,37%	0,43%
2005	125 797	44,4 ans	0,43%	0,40%	0,47%
2006	99 048	43,4 ans	0,42%	0,38%	0,46%
2003-2006	549 656	43,9 ans	0,40%	0,38%	0,41%

La faible profondeur des données (quatre année d’historique uniquement) et les fluctuations d’échantillonnage associées à la taille de l’échantillon ne permettent pas d’identifier de tendance dans l’évolution de la mortalité au fil des années. Aussi, dans le cadre de cette étude, on considère une seule période d’observation t (comme le permet Bongaarts [2004]), correspondant à la période 2003-2006.

Pour revenir sur l’approche de Bongaarts, dans une première étape, on réalise l’ajustement des trois paramètres α_t , β_t et γ_t à partir de la méthode présentée ci-dessus en 3.1.1 (moindres carrés pondérés).

À cet effet, on commence par déterminer des valeurs initiales acceptables, et on réalise donc une régression linéaire simple de $y = \ln(q_{x+1} - q_x)$ sur x , telle que $y = a_y \times x + b_y + \varepsilon_y$. On obtient les estimations suivantes : $\hat{a}_y = 0,06451$ et $\hat{b}_y = -10,68118$ (avec des probabilités de rejeter à tort l’hypothèse H_0 de nullité du paramètre selon le test de Student respectivement égales à 1,8 % et à $9,27.10^{-10}$). À partir de la démarche présentée en 3.1.1, on en déduit les valeurs initiales suivantes : $\hat{\alpha}_i = 1,89.10^{-7}$, $\hat{\beta}_i = 6,45.10^{-2}$ et $\hat{\gamma}_i = 3,58.10^{-3}$.

À partir de ces valeurs initiales, on procède à la mise en œuvre de l’optimisation pour estimer les paramètres α_t , β_t et γ_t . Les résultats sont présentés dans le tableau suivant. À titre d’information, les résultats des estimations de Bongaarts pour la population générale masculine (moyenne des estimations entre 1950 et 2000, cf. Bongaarts [2004]) sont également présentés pour la France et pour la moyenne de 14 pays industrialisés.

Tableau 39 – Ajustement statistique (étapes 1 et 2 de Bongaarts)

Période (t)	Alpha (α_t)	Beta (β_t)	Gamma (γ_t)
Valeur initiale (régression lin.)	1,89E-07	6,45E-02	3,58E-03
Côte d'Ivoire, exp. 2003-2006	2,05E-04	6,45E-02	-3,07E-05
Bongaarts (France) (*)	4,20E-05	1,01E-01	9,80E-04
Bongaarts (14 pays indus.) (*)	3,12E-05	1,05E-01	7,50E-04

(*) Moyenne des valeurs moyennes obtenues entre 1950 et 2000 pour la population générale masculine (respectivement pour la France et pour 14 pays industrialisés, cf. Bongaarts [2004]).

Il ressort de ces résultats que le niveau de mortalité moyen de la population assurée ivoirienne entre 2003 et 2006 est supérieur au niveau de mortalité moyen de la population générale française et des pays industrialisés entre 1950 et 2000. En revanche, dans le cas de la population assurée de la Côte d'Ivoire, l'évolution de la mortalité avec l'âge est plus modérée. Enfin, on note que dans le cas de la population assurée ivoirienne, l'ajustement de la mortalité pour les jeunes adultes correspond à un ajustement à la baisse (à l'image de ce qu'obtient Thatcher [1999] pour les hommes d'Angleterre et du Pays de Galles âgés de plus de 30 ans entre 1980 et 1982).

Sur le plan statistique, il apparaît que le modèle est significatif au seuil de 5 % (la valeur du test empirique est supérieure à celle de la table au seuil de 5 % à 3 degrés de liberté).

Pour illustrer l'adéquation de l'ajustement aux taux bruts, le graphique ci-dessous compare les taux bruts et les taux ajustés déduits de l'estimation des paramètres présentés dans le Tableau 39, et le tableau suivant compare les décès observés et les décès prédits (décès théoriques déduits des taux ajustés).

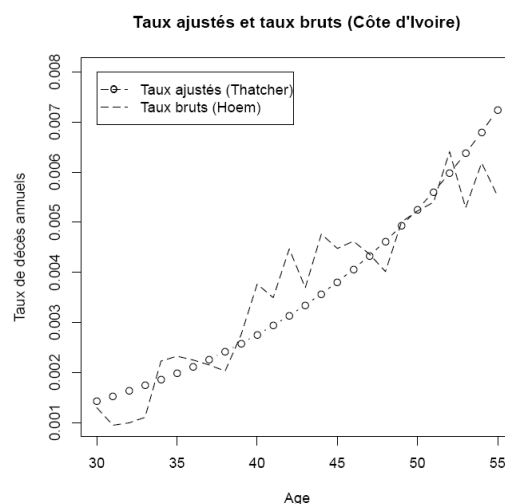
Figure 40 – Taux de décès ajustés et taux de décès bruts (Côte d'Ivoire - H)

Tableau 40 – Comparaison des réalisations / prédictions de décès

Homme Côte d'Ivoire (pop. ass.)	Exposition	Décès observés	Décès prédits	Différences relatives	Décès prédits (borne inf. à 95%)	Décès prédits (borne sup. à 95%)	Taux décès moyen (décès observés)	Taux décès moyen (décès prédits)
30-39 ans	144 931	276	294	6,4%	260	327	0,19%	0,20%
40-44 ans	126 517	514	401	-21,9%	362	441	0,41%	0,32%
45-49 ans	149 578	672	650	-3,3%	600	699	0,45%	0,43%
50-55 ans	128 630	726	779	7,4%	725	834	0,56%	0,61%
Total	549 656	2 188	2 124	-2,9%	2 034	2 214	0,40%	0,39%

Au global, il apparaît un écart inférieur à 3 % entre les décès théoriques (prédits à partir des taux ajustés) et les décès observés. Par tranches d'âge toutefois, il apparaît des écarts plus importants : les décès prédits pour les 40-44 ans minorent ainsi de plus de 20 % les décès observés.

4.1.2. Résultats de l'ajustement par avis d'expert

L'ajustement par avis d'expert est relatif à l'étape 3 de l'approche de Bongaarts, dans laquelle les paramètres α_t et γ_t sont projetés pour chaque t futur. En pratique, cette projection est réalisée dans notre étude par avis d'expert.

Concernant la mortalité environnementale γ_t , on considère qu'elle est constante tout au long de la durée de projection. On a ainsi $\hat{\gamma}_t = \hat{\gamma} = -3,07.10^{-5}$.

Concernant le niveau de mortalité α_t , on considère qu'il évolue dans le temps selon l'expression $\alpha_t = \exp(a_\alpha \times t + b_\alpha)$; on considère ainsi une décroissance exponentielle du niveau de mortalité au fil des années. Pour mémoire, les paramètres $\theta = (a_\alpha, b_\alpha)$ sont obtenus en minimisant l'écart quadratique entre la valeur d'expert de l'espérance de vie générationnelle (notée $e_{x,t}(\text{expert})$ à un âge x et à une date t) et la prédiction de cette valeur (prédiction à partir de l'expression de $e_{x,t} = \varphi_{x,t}(a_\alpha, b_\alpha)$, cf. sous-section 3.2).

La projection par avis d'expert du paramètre $\alpha_t = \exp(a_\alpha \times t + b_\alpha)$ est ainsi directement déduite de la valeur d'expert retenue pour l'espérance de vie générationnelle à un âge pivot x et à une date t . Pour faciliter le choix de l'expert, on présente dans le tableau ci-après trois références pour l'espérance de vie générationnelle, à l'âge $x = 30$ et à la date $t = t_0 = 0$ (en considérant que la date initiale $t = t_0 = 0$ correspond à l'année 2006) :

- La première est issue de la table d'expérience ajustée en supposant que sur toute la durée de projection le niveau de mortalité α_t est constant. Pour le calcul de l'espérance de vie, on considère par ailleurs que la table issue des taux ajustés se ferme brutalement à 105 ans (pour

mémoire, la méthode de fermeture n'a un impact significatif sur le provisionnement que si les individus considérés pour l'évaluation ont un âge très élevé à la date de calcul, cf. 2.1.4).

- La deuxième est issue des tables de la population assurée masculine française (tables prospectives réglementaires TGH 05). Pour mémoire, cette table se ferme à 120 ans.
- La troisième est issue des tables de la population générale masculine française (tables prospectives INED H). Pour information, cette table se ferme à 105 ans.

Tableau 41 – Références pour l'espérance de vie à 30 ans (en 2006)

Espérance de vie générationnelle (Age = 30, Date/Année = 2006)	
Côte d'Ivoire (Table du moment sans évolution de l'espérance de vie, population assurée, années 2003-2006) (*)	51,4
TGH05 (Tables prospectives réglementaires en France, population assurée, générations 1900-2005)	60,6
INED H (Tables prospectives en France, population générale, générations 1897-1996)	54,0

(*) Dans ce cas il ne s'agit pas d'une espérance de vie générationnelle, mais d'une espérance de vie simple, ne prenant donc pas en compte l'évolution de la mortalité.

Ainsi par exemple, en considérant par avis d'expert que l'espérance de vie générationnelle de la population assurée ivoirienne est raisonnablement inférieure à celle de la population générale française, on peut supposer qu'elle est égale à 53 ans (contre 54 ans pour la population générale française).

On cherche alors les valeurs de $\theta = (a_\alpha, b_\alpha)$ permettant de minimiser :

$$\left(53 - \sum_{h>0} \exp(-\hat{\gamma}h) \prod_{k=0}^{h-1} \left(\frac{1 + \alpha_{t+k} e^{\hat{\beta}(x+k)}}{1 + \alpha_{t+k} e^{\hat{\beta}(x+1+k)}} \right)^{1/\hat{\beta}} \right)^2,$$

avec $\alpha_t = \exp(a_\alpha \times t + b_\alpha)$. À cet effet, on considère que les valeurs initiales des paramètres $\theta = (a_\alpha, b_\alpha)$ correspondent aux valeurs obtenues à la date initiale $t = t_0 = 0$, soit $(a_\alpha)_i = 0$ et $(b_\alpha)_i = \ln(\hat{\alpha}_0) = -8,49$. On obtient alors les estimations suivantes : $\hat{a}_\alpha = -3,24.10^{-3}$ et $\hat{b}_\alpha = -8,49$. L'estimation par avis d'expert du niveau de mortalité à la date t est ainsi $\hat{\alpha}_t = \exp(-3,24.10^{-3} \times t - 8,49)$ (pour $t = 0$ en 2006, $t = 1$ en 2007, etc.). Dans ce cas, l'estimation de l'espérance de vie théorique est $\hat{e}_{x,t} = 53,0$ (on retrouve ainsi le « 53 ans » fixé par avis d'expert).

4.1.3. Illustrations des projections

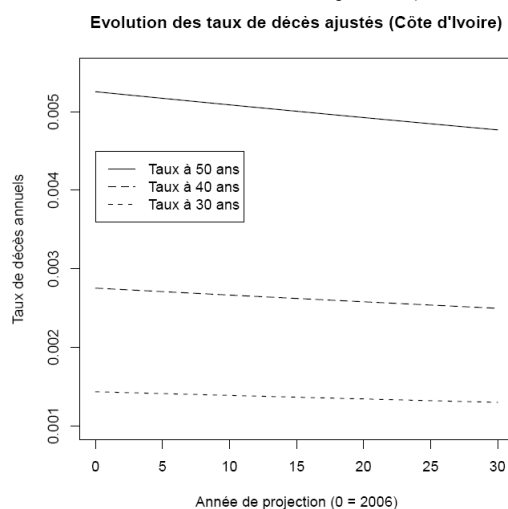
Au final, les taux de décès ajustés à un âge x et à une date t sont notés :

$$q_{x,t} = 1 - \exp(-\hat{\gamma}) \left(\frac{1 + \exp(\hat{a}_\alpha \times t + \hat{b}_\alpha) e^{\hat{\beta}x}}{1 + \exp(\hat{a}_\alpha \times t + \hat{b}_\alpha) e^{\hat{\beta}(x+1)}} \right)^{1/\hat{\beta}},$$

et en considérant que $t=0$ en 2006, on a ainsi $\hat{\gamma} = -3,07.10^{-5}$, $\hat{a}_\alpha = -3,24.10^{-3}$, $\hat{b}_\alpha = -8,49$ et $\hat{\beta} = 6,45.10^{-2}$. Cette expression permet notamment de projeter les taux de décès dans le cadre de l'étape 4 de l'approche de Bongaarts.

Sur ces bases, le graphique suivant présente l'évolution dans le temps des taux de décès à différents âges (en l'occurrence à 30 ans, 40 ans et 50 ans).

Figure 41 – Évolution des taux de décès ajustés (Côte d'Ivoire - H)



Au fil du temps, il ressort une baisse des taux de décès ajustés cohérente avec l'amélioration anticipée de la mortalité (hausse de l'espérance de vie).

4.2. Sensibilités et mesure du risque d'avis d'expert

Pour mémoire, les résultats des projections présentés dans la sous-section précédente sont issus d'un avis d'expert à plusieurs niveaux : forme de la décroissance du paramètre représentant le niveau de la mortalité α_t (choix d'une décroissance exponentielle), niveau de l'espérance de vie générationnelle à 30 ans (âge pivot) en 2006 (choix de 53 ans), âge de fermeture de la table (choix de 105 ans).

Dans cette sous-section, on apprécie la sensibilité des projections à ces différentes hypothèses. En particulier, le premier paragraphe (*cf.* 4.2.1) est consacré à l'analyse de la sensibilité des projections de α_t aux choix de l'âge

pivot et de l'avis d'expert associé sur l'espérance de vie générationnelle. Dans le second paragraphe (cf. 4.2.2), on se place dans un contexte d'assurance et on mesure la sensibilité d'une provision d'assurance *best estimate* au risque d'avis d'expert (tant sur le niveau de l'espérance de vie, que sur la forme de la décroissance du niveau de mortalité α_t et sur l'âge de fermeture de la table).

4.2.1. Cohérence des avis d'expert (sensibilité au choix de l'âge pivot)

Pour apprécier la cohérence des avis d'expert, on sollicite un avis d'expert sur l'espérance de vie à différents âges pivots, et on étudie la cohérence des projections du niveau de mortalité α_t qui en découlent. Dans le cadre de cette analyse, on considère que le paramètre représentant le niveau de mortalité suit une décroissance exponentielle.

Tableau 42 – Cohérence des avis d'expert (Côte d'Ivoire - H)

Avis d'expert sur l'esp. de vie gén. 2006 (esp. de vie du moment issue de la table d'expérience)	a	b	alpha (décroissance exponentielle)			
			t=0	t=10	t=20	t=30
A 30 ans : 53 ans (51,4 ans)	-3,24E-03	-8,49E+00	2,05E-04	1,98E-04	1,92E-04	1,86E-04
A 40 ans : 43 ans (42,3 ans)	-1,88E-03	-8,49E+00	2,05E-04	2,01E-04	1,97E-04	1,94E-04
A 40 ans : 44 ans (42,3 ans)	-4,47E-03	-8,49E+00	2,05E-04	1,96E-04	1,87E-04	1,79E-04
A 50 ans : 34 ans (33,7 ans)	-1,07E-03	-8,49E+00	2,05E-04	2,03E-04	2,00E-04	1,98E-04
A 50 ans : 35 ans (33,7 ans)	-4,69E-03	-8,49E+00	2,05E-04	1,95E-04	1,86E-04	1,78E-04

Il apparaît logiquement que dans l'expression du niveau de mortalité, la sensibilité à l'avis d'expert sur l'espérance de vie s'illustre principalement sur la raison de la suite géométrique, notée $\exp(a_\alpha)$ pour mémoire. La valeur initiale de la suite géométrique, notée $\alpha_0 = \exp(b_\alpha)$, semble quant à elle inchangée dans les différents scénarios.

La cohérence et l'ajustement des différentes opinions d'experts peuvent ainsi être évalués à partir du seul paramètre a_α .

4.2.2. Risque d'avis d'expert sur les provisions

Pour mesurer le risque d'avis d'expert sur les provisions, on considère un régime de rentes en cours de service et une population de rentiers composée d'un seul individu de sexe masculin et âgé de $x_0 = 55$ ans au 31/12/2009 (date initiale de calcul notée i_0 , avec $i_0 = t_0 + 3$ en considérant que $t_0 = 0$ correspond à l'année 2006). La rente annuelle individuelle est supposée égale à 1 € à la date initiale de calcul (i_0), et est supposée être revalorisée de 2 % chaque année. Le régime prévoit le versement de cette rente jusqu'au décès.

En pratique, il s'agit ici de présenter la provision *best estimate* dans plusieurs situations. À cet effet, on considère un scénario central dans lequel l'espérance de vie générationnelle à 30 ans est égale à 53 ans, le niveau de

mortalité connaît une décroissance exponentielle et la table d'expérience prospective construite se ferme à 105 ans (hypothèses d'avis d'expert retenues jusqu'alors dans ce travail). Dans les scénarios alternatifs testés, on considère une évolution de ces hypothèses d'avis d'expert.

Dans tous les cas, on évalue des provisions déterministes : on suppose ainsi qu'il n'y a pas d'aléa sur la durée de vie restante des assurés.

La provision *best estimate* est calculée à la date initiale à partir de la relation suivante :

$$L_{i_0} = \sum_{i=1}^{\omega} F(i) \times (1+r_i)^{-i},$$

où r_i représente le taux d'actualisation des flux de l'échéance i (issu de la courbe des taux à la date initiale), ω représente la dernière échéance de versement des flux et $F(i) = (1+2\%)^{i-1} \prod_{u=0}^{i-1} (1 - q_{x_0+u, i_0+u})$ représente les flux probables des prestations à payer en i (à terme échu), avec x_0 l'âge de l'individu à la date initiale i_0 et q_{x_0+u, i_0+u} le taux de décès ajusté issu de la table construite à partir des avis d'expert. La courbe d'actualisation retenue est la courbe de l'Institut des Actuaire⁵ au 31/12/2009.

Une illustration des résultats des calculs de provisions dans les différents scénarios est présentée dans le tableau suivant.

⁵ Courbe disponible sur le site de l'Institut des Actuaire : http://www.institutdesactuaire.com/gene/link.php?doc_link=../docs/2010003190313_I A20091231.xls.

Tableau 43 – Sensibilité de la provision aux avis d’expert (Côte d’Ivoire - H)

Avis d'expert sur l'esp. de vie gén. à 30 ans en 2006 / forme décroissance niveau mortalité / âge fermeture (estimation esp. de vie gén. théorique)	Provision <i>best estimate</i> au 31/12/2009 (homme âgé de 55 ans en 2009)	
	Provision <i>best estimate</i>	Ecart relatif avec (1)
52 ans / décroiss. expo. / fermeture 105 ans (52,00 ans)	20,40	-1,1%
53 ans / décroiss. expo. / fermeture 105 ans (53,00 ans) (1)	20,62	0,0%
53 ans / décroiss. linéaire / fermeture 105 ans (53,14 ans)	20,64	0,1%
53 ans / décroiss. expo. / fermeture 120 ans (53,00 ans)	20,61	0,0%
54 ans / décroiss. expo. / fermeture 105 ans (54,00 ans)	20,85	1,1%
57 ans / décroiss. expo. / fermeture 105 ans (57,00 ans)	21,53	4,4%

À la lecture de ce tableau, il apparaît plusieurs enseignements. D’une part, on note qu’une évolution d’un an de l’avis d’expert sur l’espérance de vie générationnelle à 30 ans engendre, toutes choses égales par ailleurs, une évolution d’environ 1 % de la provision d’assurance. De même, une évolution quatre ans (soit de plus de 7 %) de l’avis d’expert sur l’espérance de vie générationnelle à 30 ans engendre une évolution de 4 à 5 % de la provision d’assurance (en pratique, l’impact sur la provision est minoré par l’effet actualisation).

D’autre part, il ressort que pour l’évolution du niveau de mortalité, le choix d’une décroissance linéaire aux dépens d’une décroissance exponentielle a un impact limité sur le montant de la provision. En revanche, le choix de la forme de la décroissance semble influencer sur la performance du processus de minimisation permettant de rapprocher l’espérance de vie théorique à celle issue de l’avis d’expert (avec la décroissance exponentielle on converge plus précisément vers l’avis d’expert pour l’espérance de vie).

Enfin, il apparaît que le passage d’une fermeture de 105 ans à 120 ans a, toutes choses égales par ailleurs, également un impact limité sur la provision (à l’image de ce qui était annoncé au paragraphe 2.1.4).

5. Synthèse et conclusion

Dans le cadre de régimes de rentes, l’assureur est tenu d’apprécier le niveau actuel et l’évolution attendue de la mortalité spécifique de sa population assurée, compte tenu du contexte de hausse générale de l’espérance de vie (y compris dans les pays en développement, *cf.* Meslé et Vallin [2007]) et d’introduction de nouveaux référentiels justifiant l’évaluation réaliste et fine des risques (par opposition à l’approche prudente souvent retenue). Il doit alors construire des tables d’expérience prospectives.

À cet effet, lorsque le volume de données est restreint, en particulier lorsque l’on dispose de peu d’années d’historique (comme c’est le cas dans les pays d’Afrique subsaharienne francophone), l’utilisation du modèle classique de

Lee-Carter n'est pas souhaitable et le recours à des informations externes est nécessaire. La dimension qualitative (avis d'expert) dans les projections prend alors une place importante. Autre conséquence : la présence de données limitées engendre un risque significatif d'estimation des paramètres compte tenu des fluctuations d'échantillonnage (*cf.* Kamega et Planchet [2010] et Planchet et Kamega [2011a]). Dans ce cas, une attention particulière doit également être accordée au nombre de paramètres dans le modèle.

Dans ce contexte, l'approche proposée par Bongaarts [2004], qui consiste à utiliser le modèle de Thatcher [1999] et à faire évoluer les paramètres au cours du temps, se révèle utile pour répondre à ces contraintes. Aussi, dans ce chapitre le modèle de Bongaarts est utilisé pour projeter la mortalité d'expérience. À cet effet, dans un premier temps, les trois paramètres du modèle sont estimés à la date initiale par ajustement statistique. Dans un second temps, en termes de projection, deux des trois paramètres (ceux relatifs à la mortalité environnementale et à l'augmentation de la mortalité avec l'âge) sont supposés constants au cours du temps, et les valeurs futures du troisième paramètre (celui relatif au niveau de la mortalité) sont déduites de l'avis d'expert retenu pour l'espérance de vie générationnelle.

La solution proposée dans ce travail permet alors de disposer d'un outil simple pour projeter la mortalité d'expérience en l'absence de données d'historique suffisantes (quatre ans dans l'exemple numérique présenté ici). Seul un avis d'expert sur le niveau de l'espérance de vie générationnelle (et dans une moindre sur la forme de la décroissance du niveau de mortalité et sur l'âge de fermeture de la table) est nécessaire.

En complément, une appréciation du risque lié au choix de l'avis d'expert est réalisée dans ce travail. Pour un engagement viager à 55 ans, il apparaît alors qu'un écart d'un an dans le choix de l'espérance de vie conduit à un écart d'environ 1 % sur la provision, et qu'un écart de 4 ans conduit à un écart de 4 à 5 %. Les impacts du choix de la forme de l'évolution de la mortalité (forme linéaire ou exponentielle) ou du choix de l'âge de fermeture de la table apparaissent secondaires.

Il convient toutefois de noter que certaines problématiques susceptibles de faire l'objet de développements importants n'ont pas été traitées dans le présent travail. Ainsi, aucune projection de la mortalité par cause n'a été réalisée alors qu'une telle analyse pourrait être utile (*cf.* paragraphe 2.1.1), notamment dans le contexte des données africaines utilisées ici compte tenu de l'impact potentiel du VIH/Sida dans la mortalité. Par ailleurs, ni l'interdépendance entre les paramètres, ni celle entre les âges, ni celle entre les périodes (*cf.* paragraphe 2.1.2) n'ont été étudiées, alors qu'elles sont confirmées par plusieurs études (*cf.* par exemple Booth et Tickle [2008] et Loisel et Serant [2007]). Enfin, hormis le risque systématique d'avis d'expert, les mesures d'incertitudes (risque systématique d'estimation, risque systématique de modèle, risque mutualisable) n'ont pas été développés ici, alors qu'elles présentent un intérêt particulier dans le contexte de données limitées (*cf.* par exemple Booth et Tickle [2008] et Planchet et Kamega [2011a]).

Bibliographie

- ABS [2009] « Causes of death », *Australian Bureau of Statistics*.
- Alho J. M. [2007] « Méthodes utilisées pour établir les projections relatives à la mortalité – Distributions prédictives de la mortalité future », *ISSA, Quinzième Conférence internationale des actuaires et statisticiens de la sécurité sociale*.
- Barrieu P., Bensusan H., El Karoui N., Hillairet C., Loisel S., Ravanelli C., Salhi Y. [2010] « Understanding, Modelling and Managing Longevity Risk: Key Issues and Main Challenges », *Scandinavian Actuarial Journal*.
- Bongaarts J. [2004] « Long-Range Trends in Adult Mortality: Models and Projections Methods », *Population Council, WP192*.
- Booth H., Tickle L. [2008] « Mortality modeling and forecasting: A review of methods », *The Australian Demographic and Social Research Institute, WP3*.
- Brouhns N., Denuit M. [2002] « Risque de longévité et rentes viagères – I. Évolution de la mortalité en Belgique de 1880 à nos jours », *Belgian Actuarial Bulletin, Vol. 2, No. 1*.
- Brouhns N., Denuit M., Vermunt J. K. [2002] « A Poisson log-bilinear regression approach to the construction of projected lifetables », *Insurance: Mathematics and Economics, No. 31*.
- Delwarde A., Denuit M. [2006] « Construction de tables de mortalité périodiques et prospectives », *Economica*.
- Droesbeke J.-J., Fichet B., Tassi P. [1989] (éditeurs) « Analyse statistique des durées de vie », *Economica*.
- Kamega A. [2006] *Impact de la sursuicidité en Bretagne sur les garanties temporaires décès*, Mémoire d'actuariat, EURIA.
- Kamega A. [2011] *Quelle anticipation de l'espérance de vie dans les portefeuilles d'assurance ?*, Chaire transitions démographiques, transitions économiques, Intervention du 27/01/2011.
- Kamega A., Planchet F. [2010], « Mesure du risque d'estimation associé à une table d'expérience », *Cahiers de recherche de l'ISFA, WP2136*.
- Kamega A., Planchet F. [2011a], « Analyse et comparaison des populations générale et assurée en Afrique subsaharienne francophone pour anticiper la mortalité future », *Cahiers de recherche de l'ISFA, WP2138*.
- Kamega A., Planchet F. [2011b], « Hétérogénéité : mesure du risque d'estimation dans le cas d'une modélisation intégrant des facteurs observables », *Bulletin Français d'Actuariat, Vol. 11, No. 21*.
- Klein J. P., Moeschberger M. L. [2005] « Survival Analysis – Techniques for Censored and Truncated Data », *Springer, 2nd edition*.

- Lee R. [2000], « The Lee-Carter Method for Forecasting Mortality, with Various Extensions and Applications », *North American Actuarial Journal*, Vol. 4, No. 1.
- Loisel S., Serant D. [2007], « In the core of longevity risk: hidden dependence in stochastic mortality models and cut-offs in prices of longevity swaps », *Cahiers de recherche de l'ISFA*, WP2044.
- Omran A. R. [2005] « The Epidemiologic Transition: A Theory of the Epidemiology of Population Change », *The Milbank Quarterly*, Vol. 83, No. 4, Reprinted.
- Meslé F., Vallin J. [2007] « De la transition épidémiologique à la transition sanitaire », *Médecine Tropicale*, No. 67.
- Planchet F. [2006], « Tables de mortalité d'expérience pour les portefeuilles de rentiers – (Tables TGH 05 et TGF 05) – Notice de présentation », *Institut des Actuaire*.
- Planchet F. [2007], « Prospective models of mortality with forced drift – Application to the longevity risk for life annuities », *Proceedings of the 11th IME Congress*.
- Planchet F. [2010] *Modèles de durée – Tables de mortalité*, Support de cours 2010-2011, Version 2.11 (novembre 2010).
- Planchet F., Kamega A. [2011a], « Construction d'une table de mortalité prospective pour un régime de rentes : prise en compte du risque d'estimation », *Cahiers de recherche de l'ISFA*, WP2140.
- Planchet F., Kamega A. [2011b], « Mortalité prospective en cas de petits échantillons : modélisation à partir d'informations externes en utilisant l'approche de Bongaarts », *Cahiers de recherche de l'ISFA*, WP2141.
- Planchet F., Lelieur V. [2007] « Utilisation des méthodes de Lee-Carter et log-Poisson pour l'ajustement de tables de mortalité dans le cas de petits échantillon », *Bulletin Français d'Actuariat*, Vol. 7, No. 14.
- Planchet F., Leroy G. [2009] « Quel niveau de segmentation pertinent ? », *La Tribune de l'Assurance*, n. 142.
- Planchet F., Théron P. [2006] « Modèles de Durée – Applications actuarielles », *Economica*.
- Quashie A., Denuit M. [2005] « Modèles d'extrapolation de la mortalité aux grands âges », *Institut des Sciences Actuarielles et Institut de Statistique, Université Catholique de Louvain*, WP.
- Renshaw A., Haberman S. [2003] « Lee-Carter mortality forecasting with age-specific enhancement », *Insurance: Mathematics and Economics*, Vol. 33.
- Renshaw A., Haberman S. [2005] « Mortality Reduction Factors Incorporating Cohort Effects », *Actuarial Research Paper (Cass Business School)*, No. 160.

- Tableau E., Ekamper P., Huisman C., Bosch A. [2001a] « Predictiong Mortality from Period, Cohort or Cause-Specific Trends: a Study of Four European Countries », *extrait de* Tableau E., Van Den Berg Jeths A., Heathcote C. (edited by) [2001b] « Forecasting Mortality in Developed Countries – Insights from a Statistical, Demographic and Epidemiological Perspective », *Kluwer*.
- Thatcher A. R. [1999] « The long-term pattern of adult mortality and the highest attained age », *Journal of the Royal Statistical Society*, 162.
- Van Den Berg Jeths A., Hoogenveen R., De Hollander G., Tableau E. [2001] « A Review of Epidemiological Approaches to Forecasting Mortality and Morbidity », *extrait de* Tableau E., Van Den Berg Jeths A., Heathcote C. (edited by) [2001b] « Forecasting Mortality in Developed Countries – Insights from a Statistical, Demographic and Epidemiological Perspective », *Kluwer*.
- Wilmoth J. R. [1995] « Are mortality projections always more pessimistic when disaggregated by cause of death? », *Mathematical Population Studies*, Vol. 5(4).

Chapitre 6

Modélisation prospective en présence de données sur les tendances passées et mesure des risques associés¹

Sommaire

CHAPITRE 6 MODÉLISATION PROSPECTIVE EN PRÉSENCE DE DONNÉES SUR LES TENDANCES PASSÉES ET MESURE DES RISQUES ASSOCIÉS	197
1. Introduction.....	198
2. Construction de tables de mortalité prospectives.....	200
2.1.Présentation des données disponibles	200
2.2.Présentation de la méthode retenue.....	201
2.3.Validation de la table obtenue et présentation des risques.....	202
3. Encadrement des taux de décès bruts au titre des fluctuations d'échantillonnage.....	204
3.1.Intervalles de confiance pour les taux bruts.....	204
3.2.Bandes de confiance pour les taux bruts.....	205
4. Encadrement des tables de mortalité au titre des fluctuations d'échantillonnage (risque d'estimation)	207
4.1.Ré-échantillonnage des taux bruts au titre des fluctuations d'échantillonnage et nouvelles constructions de tables.....	207
4.2.Construction d'intervalles de confiance sur les tables à partir des espérances de vie conditionnelles partielles.....	209
5. Impact des fluctuations d'échantillonnage dans les évaluations de l'assureur (risque d'estimation)	213
6. Comparaison du risque d'estimation et du risque d'avis d'expert	215
7. Synthèse et conclusion	217
Bibliographie.....	219

¹ Ce chapitre s'appuie principalement sur l'article Planchet et Kamega [2011], relatif à la prise en compte du risque d'estimation dans le cadre de la construction d'une table de mortalité prospective pour un régime de rentes.

De manière générale, la nécessité de déterminer des provisions *best estimate* conduit les assureurs de régimes de rentes à devoir construire des tables de mortalité prospectives rendant compte des spécificités du portefeuille considéré. Comme évoqué dans le chapitre précédent, cette construction se heurte classiquement à plusieurs difficultés pratiques, compte tenu notamment des limites dans le volume et la profondeur des données d'expérience.

En l'absence de données d'expérience sur les tendances passées (c'est-à-dire par exemple lorsque la profondeur des données est inférieure ou égale à 5 ans), une démarche pour construire une table de mortalité prospective a été proposée dans le chapitre précédent (démarche qui s'appuie principalement sur des avis d'expert).

En présence de données d'expérience sur les tendances passées (ce que l'on suppose dans ce chapitre, même si ce n'est pas le cas aujourd'hui dans la région étudiée), il est souvent nécessaire d'utiliser des méthodes de positionnement par rapport à une table de référence externe, généralement dans le cadre d'une régression des logits d'expérience sur les logits de référence. La table ainsi obtenue est alors entachée, entre autres, d'un risque d'estimation (due aux fluctuations d'échantillonnage dans les données ventilées par année pour tenir compte de la tendance) et d'un risque d'avis d'expert (dû au choix de la table de référence), qui ont pour conséquence une incertitude sur le niveau de la provision *best estimate* à constituer. On présente alors dans ce chapitre une méthode de quantification de cette incertitude.

1. Introduction

Comme évoqué dans le chapitre précédent, dans un régime de rentes le risque viager est transféré de l'assuré à l'assureur. Ce dernier doit alors évaluer son engagement avec des tables de mortalité appropriées, prenant en compte l'augmentation de l'espérance de vie de sa population assurée.

Dans le cadre de la construction de telles tables, et plus généralement dans le cadre de projections de la mortalité, il est nécessaire de décrire les risques en fonction de leur caractère mutualisable ou systématique, compte tenu des différences existantes entre un aléa frappant différemment les diverses têtes de la population (risque mutualisable) et un aléa frappant de manière systématique l'ensemble de la population (risque systématique, non mutualisable et donc potentiellement plus dangereux pour l'assureur ou le régime).

En reprenant la classification d'Alho, on distingue plus précisément les quatre sources de risques techniques suivantes (*cf.* détails dans Booth et Tickle [2008] par exemple) :

- le risque systématique de modèle, dû à une mauvaise spécification du modèle ;
- le risque systématique d'estimation des paramètres, dû à une mauvaise estimation des paramètres du modèle, compte tenu par exemple des fluctuations d'échantillonnage ;
- le risque systématique d'avis d'expert, dû à un mauvais jugement de l'expert lors de la prise en compte d'informations externes ;

- le risque mutualisable, dû aux fluctuations de l'espérance empirique autour de l'espérance mathématique, compte tenu des fluctuations d'échantillonnage.

Le risque mutualisable, relatif au caractère aléatoire des décès individuels (que ce soit pour les observations passées ou pour les prévisions futures, comme le rappelle Pitacco [2004]), n'est pas traité ici et on se concentre dans cette étude sur les risques systématiques. Comme indiqué ci-dessus, les risques systématiques peuvent être répartis en trois classes : le risque de modèle, le risque d'estimation et le risque d'avis d'expert.

En pratique, on considère dans ce chapitre un groupe restreint observé sur un historique relativement court d'une dizaine d'années pour construire des tables et projeter la mortalité, et en termes de risques on se limite aux risques d'estimation et d'avis d'expert. L'attention accordée au risque systématique d'estimation, lié aux fluctuations d'échantillonnage, est due à l'importance des fluctuations d'échantillonnage lorsque le volume de données est restreint (comme le rappelle Olivieri [2001] dans son papier sur l'incertitude dans les projections de mortalité). Par ailleurs, l'attention accordée, dans une moindre mesure, au risque systématique d'avis d'expert, lié au choix de la table de référence externe, est justifié par le fait que l'utilisation d'un groupe restreint pour construire la table prospective conduit à utiliser une table de référence externe (*cf.* Planchet et Lelieur [2007]). Le troisième risque systématique, le risque de modèle, n'est pas traité dans cette étude, mais on peut toutefois rappeler ici qu'il s'agit d'un sujet utile qui peut se révéler complexe et pour lequel il existe un espace important pour la recherche future (*cf.* Booth et Tickle [2008]).

En résumé, on se concentre donc ici sur le risque d'estimation des paramètres lié aux fluctuations d'échantillonnage et, dans une moindre mesure, sur le risque d'avis d'expert lié au choix de la référence externe retenue.

Ces deux risques sont des risques systématiques et ne se mutualisent donc pas. Il s'agit ainsi de risques potentiellement dangereux, et susceptibles d'avoir un impact non négligeable dans les évaluations de l'assureur (provision *best estimate*, marge de risque, besoin de capital, coût des risques non couvrables, etc.). Une réflexion s'impose alors sur la mesure de ces risques dans le cadre de la construction de tables prospectives d'expérience.

Dans ce contexte, ce chapitre évalue et compare ces deux sources de risques, en accordant notamment une attention particulière au risque d'estimation lié aux fluctuations d'échantillonnage. Ainsi, après avoir présenté dans la section 2 les données et la méthode retenue pour construire et valider la table, on présente dans la section 3 une démarche pour encadrer les taux de décès bruts au titre des fluctuations d'échantillonnage. La section 4 s'attache ensuite à proposer une méthode pour mesurer le risque d'estimation sur les tables au titre de ces fluctuations d'échantillonnage. Cette méthode permet notamment d'obtenir, à partir de simulations, une table de mortalité pour un niveau de quantile donné. Avec cette méthode, il est alors possible de construire des intervalles de confiance pour les provisions sans réaliser de nouvelles simulations (comme illustré dans la section 5). Enfin, la section 6 est consacrée à la comparaison du risque d'estimation dû aux fluctuations

d'échantillonnage au risque d'avis d'expert exprimé au travers du choix de l'information externe retenue.

Le problème de l'interdépendance des sources de risques entre les périodes ou entre les âges (*cf.* Booth et Tickle [2008] et Loisel et Serant [2007]) n'est pas traité dans ce travail.

2. Construction de tables de mortalité prospectives

Cette section est consacrée à la présentation de la méthode de construction de tables de mortalité prospectives d'expérience par sexe pour l'évaluation des engagements d'un régime de retraite.

2.1. Présentation des données disponibles

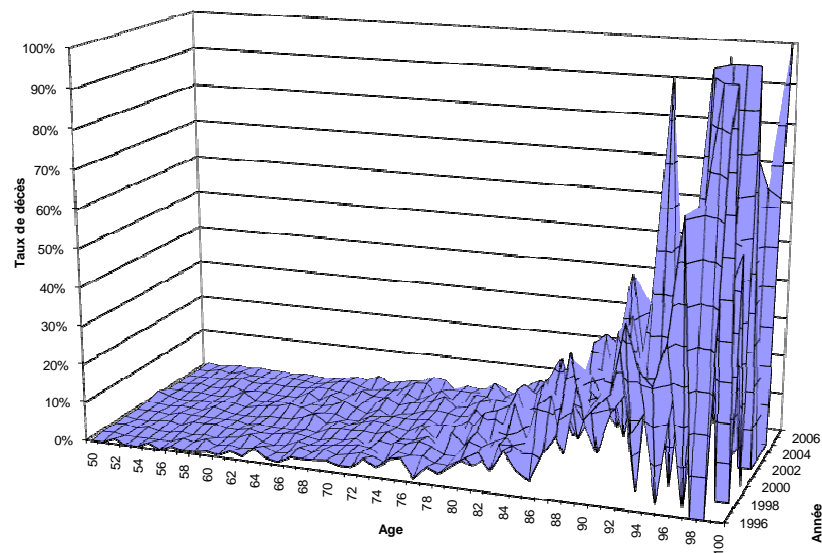
On dispose de données relatives à un régime de retraite français de 1996 à 2007 avec une exposition au risque concentrée sur la plage d'âge 50-100 ans. L'exposition annuelle globale est d'environ 20 000 années avec un *sex-ratio* de 40 %.

Les taux bruts, notés \hat{q}_{xt} pour $x = 50, \dots, 100$ et $t = 1996, \dots, 2007$, sont calculés en rapportant le nombre de décès d_{xt} à l'exposition au risque R_{xt} (méthode de Hoem²), soit :

$$\hat{q}_{xt} = d_{xt} / R_{xt} .$$

Ils ont l'allure suivante, pour la population masculine :

Figure 42 - Taux de décès bruts (hommes)



² Ce qui s'avère très proche de l'estimateur de Kaplan-Meier dans le cas présent.

Cette figure présente des irrégularités dans les taux bruts de décès, directement liées aux fluctuations d'échantillonnage.

2.2. Présentation de la méthode retenue

Le volume des données utilisables est modeste : avec une exposition au risque annuelle d'environ 8 000 pour les hommes et 12 000 pour les femmes sur les 12 années de la période d'observation, il n'est pas envisageable de proposer une construction endogène des tables d'expérience *via* un modèle classique de type Lee-Carter (*cf.* par exemple Lee [2000]) ou l'une de ses variantes (*cf.* notamment Brouhns et al. [2002], Renshaw et Haberman [2003] et Renshaw et Haberman [2005]).

Aussi, il a été décidé de s'appuyer sur une référence externe et de construire les tables d'expérience H/F en positionnant la mortalité d'expérience par rapport à cette référence. Les tables réglementaires françaises TGH/F 05 ont été choisies comme référence, compte tenu du fait qu'elles décrivent la survie des rentiers de portefeuilles d'assureurs français (*cf.* Planchet [2006]). L'ajustement est effectué sur les tables du moment reconstituées de 1996 à 2007. Le modèle utilisé repose sur :

$$\ln\left(\frac{\hat{q}_{xt}}{1-\hat{q}_{xt}}\right) = a \times \ln\left(\frac{q_{xt}^{réf}}{1-q_{xt}^{réf}}\right) + b + \varepsilon_{xt},$$

où ε_{xt} désigne le terme d'erreur et $q_{xt}^{réf}$ représente le taux de décès, à l'âge x et l'année calendaire t , pour la table de référence (issue de la table TGH 05 pour les hommes et de la table TGF 05 pour les femmes).

Ce modèle s'apparente ainsi au modèle relationnel de Brass ; il présente notamment l'avantage d'éviter l'hypothèse de proportionnalité sous-jacente au modèle relationnel de Cox et d'être moins paramétré que le modèle relationnel de Hannerz (ces trois modèles relationnels sont présentés et comparés dans Delwarde et al. [2004]).

Par ailleurs, l'utilisation des logits dans ce modèle relationnel est justifiée par le fait qu'empiriquement on constate que $\ln(q_{xt}/(1-q_{xt}))$ est approximativement affine au-delà de 25 ans (cette observation est à l'origine du modèle paramétrique proposé dans Thatcher [1999]). Dès lors, il est naturel de proposer une relation affine entre les logits de différentes tables de mortalité. Au surplus, l'utilisation des logits dans un contexte de rentes permet d'introduire une certaine prudence dans l'ajustement des q_{xt} qui en résulte : en effet, par concavité de la fonction logistique pour des taux de décès inférieurs à $\frac{1}{2}$, avec l'inégalité de Jensen on montre que les logits sont biaisées négativement si les taux de décès sont estimés sans biais (*cf.* Planchet et Théron [2006]).

Les taux ajustés du jeu de tables finalement utilisés sont obtenus en inversant la fonction logistique :

$$q_{xt} = \frac{\exp(y_{xt})}{1 + \exp(y_{xt})},$$

avec $y_{xt} = \hat{a} \ln \left(\frac{q_{xt}^{réf}}{1 - q_{xt}^{réf}} \right) + \hat{b}$.

Les paramètres $\theta = (a, b)$ correspondant sont estimés en minimisant la « distance » sur la base du critère de moindres carrés pondérés suivant :

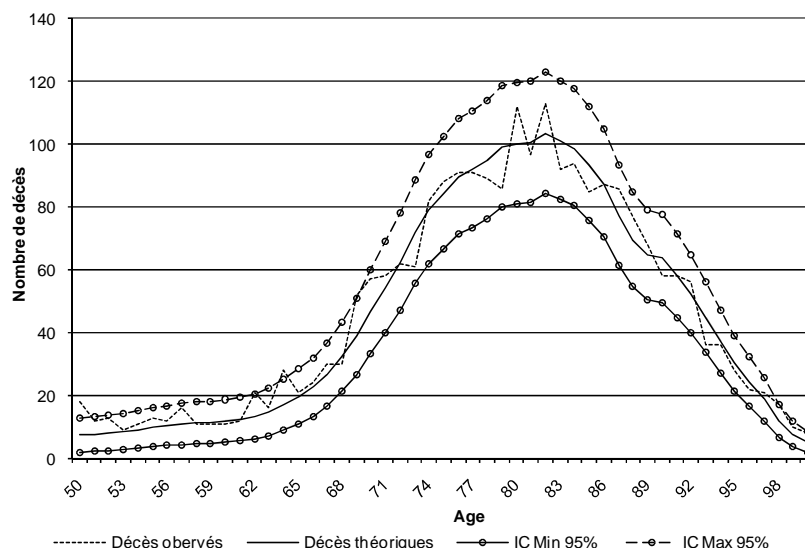
$$D = \sum_{x,t} R_{xt} (\hat{q}_{xt} - q_{xt})^2,$$

où R_{xt} représente l'exposition au risque à l'âge x l'année calendaire t . Ce critère est proche de celui obtenu par la méthode de De Vylder (*cf.* Delwarde et Denuit [2006], Planchet et Thérond [2006] ou chapitre précédent), lorsque l'on ramène la méthode de maximum de vraisemblance à un problème de moindres carrés en utilisant l'approximation de la loi de \hat{q}_{xt} par une loi normale et en remplaçant la variance théorique par la variance estimée (pour mémoire, dans ce cas le critère s'écrit $D^{(*)} = \sum_{x,t} \frac{R_{xt}}{\hat{q}_{xt}(1-\hat{q}_{xt})} (\hat{q}_{xt} - q_{xt})^2$). Le critère D retenu ici présente notamment l'avantage d'inclure les observations où les \hat{q}_{xt} sont nuls, ce qui est fréquent lorsque le groupe de la population assuré est restreint.

L'optimisation est réalisée sur la plage 50-100 ans et sur la période 1996-2007 pour les hommes, puis pour les femmes.

2.3. Validation de la table obtenue et présentation des risques

La table ainsi obtenue est classiquement validée en comparant les décès prédits et observés âge par âge sur l'ensemble de la période d'observation. En effectuant l'approximation usuelle d'une loi binomiale par une loi normale, on obtient pour les hommes les résultats ci-après :

Figure 43 – Décès prédits et des décès observés 1996-2007 (hommes)

Les courbes avec des marqueurs représentent les bornes inférieure et supérieure de l'intervalle de confiance ponctuel (IC) à 95 % pour les taux de décès ajustés. On peut formuler deux observations à la lecture de ce graphique :

- les décès observés se situent à l'intérieur des intervalles de confiance théoriques à 95 % sur l'essentiel des âges de la plage d'âge considérée, signe d'une représentation correcte de la réalité par la table d'expérience ;
- l'imprécision est importante, l'erreur relative³ sur le nombre de décès étant supérieure à 18 % à tous les âges (l'erreur relative globale est d'environ 25 % sur l'ensemble des âges).

On obtient pour les femmes des conclusions analogues avec une erreur relative supérieure à 14 % à tous les âges (l'erreur relative globale est d'environ 20 % sur l'ensemble des âges).

Le niveau élevé des erreurs relatives sur les nombres de décès impose une attention particulière sur l'analyse des risques. En pratique, l'erreur relative entre les décès théoriques et les décès observés est due à plusieurs sources de risques de nature différente :

- pour les décès théoriques, il existe un risque systématique lié à l'incertitude dans la spécification du modèle, dans l'estimation des paramètres et dans le choix de la table de référence externe ;
- pour les décès observés, il existe un risque mutualisable, conséquence directe de la taille finie du groupe entraînant une fluctuation de l'espérance empirique autour de l'espérance mathématique (même si l'on considère que les trois sources de risque systématique sont

³ Mesurée par la demi-longueur relative de l'intervalle de confiance à 95 %.

négligeables, on doit s'attendre à un écart entre les décès théoriques et les décès observés du fait des fluctuations d'échantillonnage).

Dans le cadre de cette étude, on s'attache principalement à apprécier le risque d'estimation des paramètres du modèle (*cf.* sections 4 et 5), c'est-à-dire le risque systématique associé à l'utilisation du calibrage d'une loi de durée dont on sait qu'il n'est pas exact du fait des erreurs d'échantillonnage affectant sa construction. Dans un premier temps toutefois, on apprécie le niveau des fluctuations d'échantillonnage sur les taux de décès observés (*cf.* section 3). L'ensemble des applications numériques présentées ci-après est réalisé pour la population masculine uniquement.

3. Encadrement des taux de décès bruts au titre des fluctuations d'échantillonnage

Les taux de décès bruts, c'est-à-dire les taux de décès observés, présentent des irrégularités directement liées aux fluctuations d'échantillonnage. Dans cette section on cherche à construire un encadrement des taux de décès bruts au titre de ces fluctuations.

En pratique, on cherche notamment à encadrer les taux de décès bruts simultanément sur une tranche d'âge (bandes de confiance, *cf.* sous-section 3.2), et non uniquement ponctuellement à un âge donné (intervalles de confiance ponctuels, *cf.* sous-section 3.1), puisque les tables de mortalité que l'on construit pour le provisionnement sont appliquées à plusieurs âges, et non uniquement à un âge ponctuel donné.

3.1. Intervalles de confiance pour les taux bruts

On cherche ici dans un premier temps un intervalle de confiance pour q_{xt} (taux de décès théorique) à partir de \hat{q}_{xt} (taux de décès observé), pour un âge x et une année t . Pour mémoire on retient ici l'estimateur de taux bruts de Hoem, qui considère entre autres que les décès suivent une loi binomiale (*cf.* Planchet et Théron [2006]). On peut alors réaliser l'approximation de la loi de \hat{q}_{xt} par une loi normale, et en remplaçant la variance théorique par la variance

estimée, on a la loi $N\left(\hat{q}_{xt}; \sqrt{\frac{\hat{q}_{xt}(1-\hat{q}_{xt})}{R_{xt}}}\right)$. L'expression approchée des bornes

de l'intervalle de confiance ponctuel de q_{xt} à l'âge $x = x_0$ est alors (pour une période t fixée) :

$$\hat{q}_{xt} \pm u_{\alpha/2} \sqrt{\frac{\hat{q}_{xt}(1-\hat{q}_{xt})}{R_{xt}}},$$

où $u_{\alpha/2}$ désigne le quantile d'ordre $\alpha/2$ de la loi normale centrée réduite et $1-\alpha$ désigne le niveau de confiance (avec $\alpha \in]0,1[$). Cette approximation de

l'intervalle de confiance permet d'encadrer les taux de décès tel que (pour une période t fixée) :

$$P\left(q_{xt} \in \hat{q}_{xt} \pm u_{\alpha/2} \sqrt{\frac{\hat{q}_{xt}(1-\hat{q}_{xt})}{R_{xt}}}, x = x_0\right) = 1 - \alpha.$$

3.2. Bandes de confiance pour les taux bruts

On souhaite désormais encadrer les taux de décès simultanément sur tous les âges x d'une plage d'âge $[x_0, x_0 + n]$ (où n est un nombre entier positif) pour une année t . L'encadrement des taux de décès correspond donc désormais à une bande de confiance, et non plus à un intervalle de confiance ponctuel.

Dans les analyses de la mortalité, les approches les plus classiques pour construire des bandes de confiance concernent les estimateurs de fonctions de survie. Les méthodes les plus répandues sont notamment les bandes de confiance de Nair [1984] et celles de Hall et Wellner [1980] (cf. également l'ouvrage de Klein et Moeschberger [2005] sur ce point), qui s'appuient entre autres sur les propriétés de convergence de l'estimateur de Kaplan-Meier pour la fonction de survie.

On souhaite ici construire des bandes de confiance pour les taux de décès, et non pour des fonctions de survie. En pratique, on cherche ainsi $f(\hat{q}_{xt})$ tel que :

$$P\left(q_{xt} \in \hat{q}_{xt} \pm f(\hat{q}_{xt}), \forall x \in [x_0, x_0 + n]\right) = 1 - \alpha,$$

pour une année t fixée. À cet effet, on s'appuie sur la méthode d'estimation de Sidak, qui repose sur le principe d'inflation du seuil du test lorsque le nombre de tests augmente (cf. par exemple Abdi [2007]).

Pour mémoire, une bande de confiance au niveau de confiance $1 - \alpha$ sur la plage d'âges $[x_0, x_0 + n]$ peut être présentée comme une collection d'intervalles de confiance pour les différents âges $x \in [x_0, x_0 + n]$ construits de manière à avoir un intervalle simultané de probabilité égal à $1 - \alpha$. Soit donc :

$$P\left(q_{xt} \in \hat{q}_{xt} \pm f(\hat{q}_{xt}), x = x_0\right) = 1 - \beta,$$

l'intervalle de probabilité de niveau $1 - \beta$ (avec $\beta \in]0, 1[$) pour q_{xt} à l'âge $x = x_0$ et pour une période t fixée. La probabilité simultanée d'encadrer les taux de décès q_{xt} aux deux âges $x = x_0$ et $x = x_0 + 1$ est alors $(1 - \beta)^2$, en supposant l'encadrement indépendant sur ces deux âges. En répétant l'opération de manière à inclure tous les âges de $[x_0, x_0 + n]$, il apparaît alors, toujours sous l'hypothèse d'indépendance, que la probabilité simultanée

d'encadrer les taux de décès q_{xt} pour les différents âges $x \in [x_0, x_0 + n]$ est $(1-\beta)^{n+1}$.

Sur ces bases, on peut ainsi construire une bande de confiance au seuil α sur la tranche d'âge $[x_0, x_0 + n]$, en constituant des intervalles de confiance ponctuels pour chaque âge $x \in [x_0, x_0 + n]$ au seuil :

$$\beta = 1 - (1 - \alpha)^{1/(n+1)},$$

puisque dans ce cas on a bien $(1-\alpha) = (1-\beta)^{n+1}$. On rappelle ici que la méthode de Sidak ci-dessus est une amélioration de la méthode de Bonferroni (cf. par exemple Bland et Altman [1995]), qui considère que $\beta = \alpha/(n+1)$.

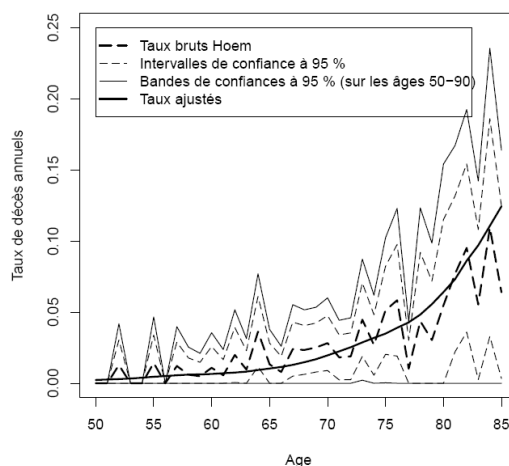
Au final, une approximation de la bande de confiance permettant d'encadrer simultanément les taux de décès sur tous les âges $[x_0, x_0 + n]$ à partir de la méthode de Sidak est (pour une période t fixée) :

$$P \left(q_{xt} \in \hat{q}_{xt} \pm u_{\beta/2} \sqrt{\frac{\hat{q}_{xt}(1-\hat{q}_{xt})}{R_{xt}}}, \forall x \in [x_0, x_0 + n] \right) = 1 - \alpha,$$

avec $\beta = 1 - (1 - \alpha)^{1/(n+1)}$.

Les intervalles et bandes de confiance ci-dessus permettent d'encadrer les taux de décès bruts au titre des fluctuations d'échantillonnage, respectivement pour un âge donné ou sur une tranche d'âge.

La figure ci-dessous illustre les taux bruts, les intervalles de confiance, les bandes de confiance (sur la tranche d'âge $[x_0 = 50, x_0 + n = 90]$) et les taux ajustés des hommes pour l'année 1996. Pour plus de lisibilité, l'illustration graphique est réalisée entre 50 et 85 ans uniquement.

Figure 44 – Intervalles et bandes de confiance pour l'année 1996 (hommes)

Il apparaît logiquement que les bandes de confiance sont plus larges que les intervalles de confiance. En outre, l'intérêt des bandes de confiance est ici illustré par l'allure des taux de décès théoriques, qui pour mémoire sont construits en prenant en compte les informations sur les décès sur tous les âges retenues pour l'ajustement (en l'occurrence ici les 50-100 ans). En effet, il apparaît qu'entre 50 et 85 ans, les taux de décès théoriques se situent tous dans les bandes de confiance des taux bruts, alors qu'à 77 ans le taux de décès théorique se situe en dehors de l'intervalle de confiance ponctuel. Cette observation confirme donc ici que les bandes de confiance sont plus adaptées pour encadrer les taux de décès bruts sur un ensemble d'âge.

4. Encadrement des tables de mortalité au titre des fluctuations d'échantillonnage (risque d'estimation)

Les irrégularités des taux bruts liées aux fluctuations d'échantillonnage créées une incertitude dans l'estimation des paramètres du modèle retenu pour ajuster les taux de décès (risque d'estimation). La description des taux de décès ajustés par les seules valeurs estimées est donc insuffisante et il est nécessaire de les décrire par des indicateurs numériques complémentaires. À cet effet, on présente ici, entre autres, un encadrement des taux de décès ajustés (et donc du jeu de tables de mortalité prospectives) associé à ces fluctuations d'échantillonnage.

Dans la continuité de ce qui a été présenté dans la section précédente, on cherche ici à encadrer directement les taux de décès ajustés sur un ensemble d'âge (c'est-à-dire la table de mortalité), et non uniquement pour un âge ponctuel donné.

4.1. Ré-échantillonnage des taux bruts au titre des fluctuations d'échantillonnage et nouvelles constructions de tables

On cherche désormais à mesurer l'impact des fluctuations d'échantillonnage sur l'estimation des paramètres du modèle retenu pour ajuster les taux de décès.

La démarche retenue à cet effet consiste, dans un premier temps, à ré-échantillonner les taux de décès bruts à partir de la méthode de simulation directe des taux présentée dans Kamega et Planchet [2010].

Pour mémoire, les taux bruts issus des données d'expérience sont estimés selon l'approche de Hoem, qui considère entre autres que le nombre de décès D_{xt} suit une loi binomiale $B(R_{xt}; q_{xt})$. Dans ce contexte, pour simuler les taux bruts \hat{q}_{xt}^k on peut s'appuyer sur la simulation du nombre de décès d_{xt}^k selon $B(R_{xt}; \hat{q}_{xt}^k)$, en retenant la méthode de simulation de Monte-Carlo. On en déduit alors pour chaque simulation k ($k \in [1, K]$, K étant le nombre total de simulations) :

$$\hat{q}_{xt}^k = d_{xt}^k / R_{xt}.$$

À partir de ces différents taux bruts ré-échantillonnés, on détermine une nouvelle estimation des paramètres $\theta = (a, b)$, notés $\theta^k = (a^k, b^k)$ au titre de la simulation k , en minimisant la « distance » $D^k = \sum_{x,t} R_{xt} (\hat{q}_{xt}^k - q_{xt}^k)^2$. On peut alors facilement obtenir une estimation des taux ajustés pour la simulation k , à partir de :

$$q_{xt}^k = \frac{\exp(y_{xt}^k)}{1 + \exp(y_{xt}^k)},$$

avec $y_{xt}^k = \hat{a}^k \ln \left(\frac{q_{xt}^{réf}}{1 - q_{xt}^{réf}} \right) + \hat{b}^k$. Ces différents taux ajustés issus des

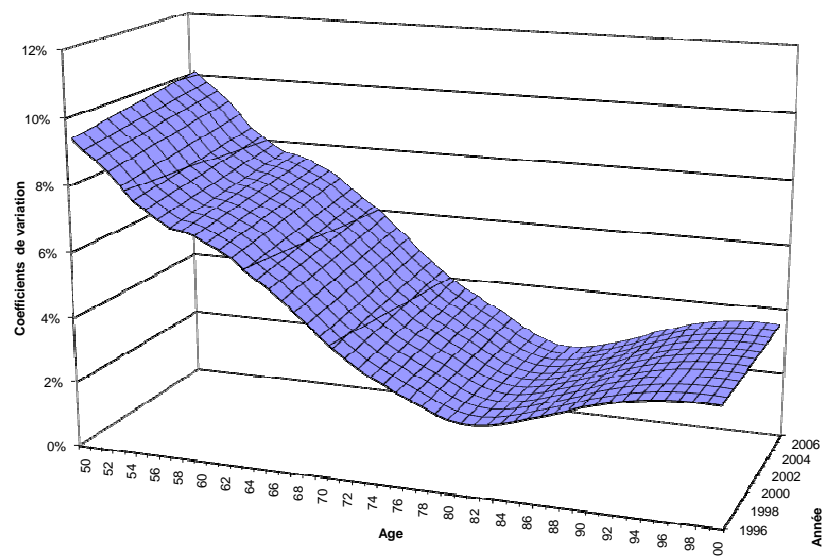
simulations k représentent la dispersion des taux de décès ajustés au titre du risque systématique sur l'estimation des paramètres compte tenu des fluctuations d'échantillonnage (risque d'estimation).

Pour quantifier cette dispersion des taux de décès ajustés au titre des fluctuations d'échantillonnage, on calcule une approximation des coefficients de variation des taux de décès, notés cv_{xt} pour un âge x et une année calendaire t . Ces coefficients sont calculés empiriquement par la relation suivante :

$$cv_{xt} = \frac{\sqrt{\frac{1}{K-1} \sum_k \left(q_{xt}^k - \frac{1}{K} \sum_k q_{xt}^k \right)^2}}{\frac{1}{K} \sum_k q_{xt}^k}.$$

En considérant $K = 5000$ simulations, on obtient les résultats des coefficients de variation présentés dans la figure suivante.

Figure 45 – Coefficients de variation des taux ajustés (hommes)



Il apparaît notamment que les coefficients de variation ont tendance à augmenter au fil des années (pour un âge donné), ce qui est cohérent avec la baisse des taux de décès à un âge donné au fil des années (compte tenu de l'amélioration de l'espérance de vie).

En outre, malgré l'augmentation logique des fluctuations d'échantillonnage lorsque l'âge augmente (comme illustrée sur la Figure 42 relative aux taux de décès bruts des hommes), il apparaît une diminution de la valeur des coefficients jusqu'à 80 ans environ. Cette baisse est liée à la hausse des taux de décès au fil des âges. À partir de 80 ans toutefois, les coefficients repartent à la hausse, ce qui signifie que l'effet lié à la hausse des taux de décès avec l'âge devient plus faible que l'effet lié à la hausse des fluctuations d'échantillonnage avec l'âge du fait de la baisse de l'exposition.

4.2. Construction d'intervalles de confiance sur les tables à partir des espérances de vie conditionnelles partielles

À l'image de ce qui a été réalisé en 3.2 avec les bandes de confiance sur les taux bruts, on cherche ici à encadrer les taux de décès ajustés au titre du risque d'estimation sur une tranche d'âge dans son ensemble et non uniquement à un âge ponctuel donné. En d'autres termes, on souhaite ici un encadrement d'une table de mortalité au titre du risque d'estimation. L'approche retenue à cet effet consiste à considérer une fonctionnelle propre à chaque table de mortalité qui lui associe un nombre positif, puis à lui appliquer un encadrement. L'espérance de vie générationnelle est de ce point de vue une fonctionnelle naturelle.

Construction d'un encadrement pour les tables de mortalité prospectives

On utilise plus précisément l'espérance de vie générationnelle partielle, entre les âges x_0 et $x_0 + n$ uniquement (avec $x_0 + n$ inférieur à l'âge limite de la table). En pratique, on ne prend donc pas en compte les âges très élevés (au-delà de $x_0 + n = 95$ ans par exemple), compte tenu de la faible exposition des effectifs à ces âges (à titre d'exemple, les plus de 95 ans représentent moins de 1 % de l'effectif sous risque chez les hommes dans notre étude) et de leur faible impact dans les analyses (*cf.* par exemple Planchet et Thérond [2006] et Quashie et Denuit [2005] qui illustrent que lorsque les rentiers âgés sont en proportion modeste dans le portefeuille, la fermeture de la table a une importance limitée).

À l'âge $x = x_0$ et à une date t , l'espérance de vie générationnelle partielle est notée $EV_{xt(x:x+n)}$. On a alors avec les taux ajustés issus des données d'expérience :

$$EV_{xt(x:x+n)} = \sum_{h=1}^n \prod_{u=0}^{h-1} (1 - q_{x+u,t+u}).$$

Ainsi, en considérant les taux ajustés issus des simulations, on a donc pour la simulation k :

$$EV_{xt(x:x+n)}^k = \sum_{h=1}^n \prod_{u=0}^{h-1} (1 - q_{x+u,t+u}^k),$$

avec les mêmes notations. Dans la suite, l'espérance de vie générationnelle partielle $EV_{xt(x:x+n)}$ entre les âges x et $x+n$ est notée simplement EV_{xt} (à l'âge $x = x_0$ et à une date t) pour faciliter la lecture.

Les espérances de vie générationnelles partielles représentent les tables de mortalité prospectives sur une tranche d'âge. Pour encadrer ces tables au titre du risque d'estimation avec une probabilité de $1 - \alpha$, on utilise alors les intervalles de confiance autour de l'espérance de vie générationnelle partielle entre x et $x+n$ ans, à l'âge $x = x_0$ et à une date t fixée. L'encadrement prend alors la forme suivante :

$$P\left(EV_{xt}^i < EV_{xt}^k \leq EV_{xt}^s, x = x_0\right) = 1 - \alpha,$$

où EV_{xt}^i et EV_{xt}^s représentent respectivement la borne inférieure et la borne supérieure de l'espérance de vie générationnelle partielle, et $1 - \alpha$ (avec $1 - \alpha \in]0, 1[$) représente la probabilité que l'espérance de vie générationnelle partielle issue de la simulation k se trouve entre ces bornes.

Compte tenu de l'approche retenue pour illustrer l'impact des fluctuations d'échantillonnage, l'encadrement des tables à partir des espérances de vie est

réalisé sur la base des quantiles. On dispose à cet effet à la date $t = t_0$ de K espérances de vie générationnelles partielles simulées, déduits des K simulations des taux de décès (chaque simulation k de l'espérance de vie générationnelle partielle est calculée à partir de la simulation k des taux de décès ajustés à partir de la simulation k des taux bruts aux âges 50 à 100 ans et aux périodes 1996 et 2007).

En considérant un seuil fixé à α , on retient les quantiles à $\alpha/2$ et $1-\alpha/2$ des K simulations d'espérance de vie générationnelle partielle. On peut alors estimer à une date $t = t_0$ les bornes EV_{xt}^i et EV_{xt}^s par :

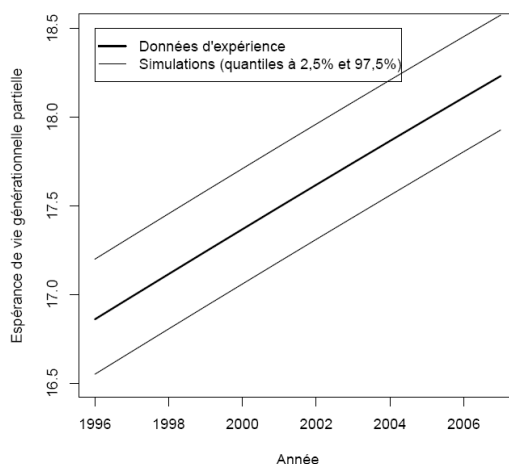
$$EV_{xt}^i = \inf \left\{ EV_{xt}^g \in [EV_{xt}^1, \dots, EV_{xt}^K] \mid P(EV_{xt}^k \leq EV_{xt}^g) \geq \alpha/2, x = x_0 \right\}$$

$$EV_{xt}^s = \inf \left\{ EV_{xt}^g \in [EV_{xt}^1, \dots, EV_{xt}^K] \mid P(EV_{xt}^k \leq EV_{xt}^g) \geq 1 - \alpha/2, x = x_0 \right\},$$

avec $g \in [1, K]$, pour un âge $x = x_0$ et une date $t = t_0$ fixés. Ces bornes permettent notamment d'en déduire les numéros des simulations i et s correspondantes et de sélectionner alors les taux de décès ajustés q_{xt}^i et q_{xt}^s associés, pour différents âges $x \in [x_0, x_0(\text{fin_table})]$ et différentes dates $t \in [t_0, t_0(\text{fin_table})]$. Au final, on dispose ainsi d'un encadrement de la table de mortalité prospective d'expérience avec un intervalle de probabilité de $1 - \alpha$.

Illustration de l'encadrement des tables de mortalité prospectives

Cet encadrement des tables de mortalité est illustré ci-dessous pour les hommes à partir de l'encadrement des espérances de vie générationnelles partielles entre les âges $x = x_0 = 67$ et $x + n = 95$, à l'âge $x = x_0 = 67$ ans (âge moyen de la population masculine exposée au risque) et aux périodes $t = t_0 = 1996$ à $t_0 + m = 2007$. En d'autres termes, sur la figure ci-dessous, les espérances de vie générationnelles partielles EV_{67t} sont ainsi encadrées par les bornes EV_{67t}^i et EV_{67t}^s (pour $t \in [t_0, t_0 + m]$). Le niveau de confiance retenu est $1 - \alpha = 95\%$ et on considère toujours $K = 5000$ simulations.

Figure 46 – Intervalles de confiance à 95 % sur les tables (esp. de vie, h.)

Il apparaît logiquement que les espérances de vie générationnelles partielles augmentent au fil des années.

Par ailleurs, pour l'année $t = t_0 = 1996$ par exemple, il apparaît au titre des fluctuations d'échantillonnage des écarts d'environ 2 % (que ce soit à la hausse ou à la baisse) autour des espérances de vie générationnelles partielles issues des taux ajustés à partir des données.

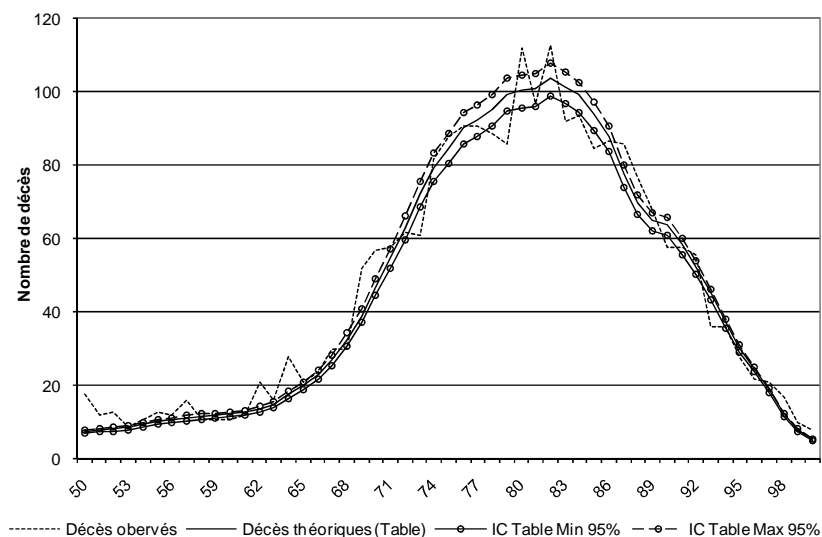
Ces écarts d'environ 2 % sont observés dans différents scénarios pour les âges de début et de fin retenus pour le calcul de l'espérance de vie générationnelle partielle, comme l'illustre le tableau suivant (toujours pour l'année $t = t_0 = 1996$).

Tableau 44 – Écarts entre les quantiles des esp. de vie et les esp. de vie issues des taux ajustés (h.)

Esp. de vie gén, partielle (pour $t=1996$)	Données d'expérience (1)	Simulations à partir des données d'expérience			
		Quantile à 2,5% (2)	Ecart relatif (2)/(1)-1	Quantile à 97,5% (3)	Ecart relatif (3)/(1)-1
Age: 50 / Age fin: 90	32,7	32,2	-1,6%	33,3	1,7%
Age: 60 / Age fin: 90	22,7	22,3	-1,7%	23,1	1,9%
Age: 67 / Age fin: 90	16,1	15,8	-1,7%	16,4	1,9%
Age: 70 / Age fin: 90	13,5	13,2	-1,7%	13,7	1,8%
Age: 50 / Age fin: 95	33,9	33,3	-1,7%	34,5	1,9%
Age: 60 / Age fin: 95	23,6	23,2	-1,8%	24,0	2,0%
Age: 67 / Age fin: 95	16,9	16,5	-1,8%	17,2	2,0%
Age: 70 / Age fin: 95	14,2	13,9	-1,9%	14,5	2,0%

Sur ce tableau, il apparaît en particulier que les écarts des quantiles ne sont pas symétriques, compte tenu notamment du caractère non linéaire de la fonction logit retenue pour l'ajustement des taux de décès.

Pour compléter l'encadrement des tables de mortalité, on présente également ci-dessous une comparaison des décès théoriques issus des taux de décès ajustés sur la base des données d'expérience à ceux issus des quantiles à 2,5 % et 97,5 % des taux de décès ajustés sur la base des simulations (simulations au titre des fluctuations d'échantillonnage pour mémoire).

Figure 47 – Décès prédits / données et quantiles (1996-2007, h.)

Les intervalles de confiance représentent ainsi les quantiles à 2,5 % et 97,5 % (simulations i et s) des tables de mortalité au titre du risque systématique d'estimation des paramètres du modèle retenu.

En particulier, il apparaît que l'erreur relative globale sur l'ensemble des âges au titre du seul risque systématique d'estimation lié aux fluctuations d'échantillonnage atteint dans nos illustrations environ 4 %, à la hausse comme à la baisse (pour mémoire l'erreur relative globale au titre de l'ensemble des risques identifiés est égale à environ 25 %, cf. Figure 43).

5. Impact des fluctuations d'échantillonnage dans les évaluations de l'assureur (risque d'estimation)

Cette section s'attache à présenter les conséquences du risque d'estimation, au titre des fluctuations d'échantillonnage, sur la provision *best estimate*.

À cet effet, on considère un régime de rentes en cours de service et une population de rentiers de sexe masculin. La rente annuelle individuelle est supposée égale à 1 € à la date initiale de calcul (0), et est supposée être revalorisée de 2 % chaque année. Le régime prévoit le versement de cette rente jusqu'à 95 ans.

Dans cette section, l'impact du risque d'estimation est donc mesuré sur la provision *best estimate* de l'assureur. En pratique, il s'agit ici d'une part de présenter la provision *best estimate* calculée avec la table construite à partir des données, puis d'autre part de présenter un encadrement à 95 % de la provision *best estimate* au titre du risque d'estimation lié aux fluctuations d'échantillonnage. Dans les deux cas, on évalue des provisions déterministes : pour un jeu de tables donné, on considère ainsi qu'il n'y a pas d'aléa sur la

durée de vie restante des assurés (on ne tient donc pas compte du risque mutualisable).

La provision *best estimate* obtenue avec la table construite à partir des données est calculée à la date initiale à partir de la relation suivante :

$$L_0 = \sum_{t=1}^{\omega} F(t) \times (1+r_t)^{-t},$$

où r_t représente le taux d'actualisation des flux de l'échéance t (issu de la courbe des taux à la date initiale), ω représente la dernière échéance de versement des flux et $F(t) = \sum_{j \in J} (1+2\%)^{t-1} \prod_{u=0}^{t-1} (1 - q_{x_d(j)+u, t_d+u})$ représente les flux probables des prestations à payer en t (à terme échu), avec J l'ensemble des j individus du portefeuille, $x_d(j)$ l'âge de l'individu j à la date initiale t_d et $q_{x_d(j)+u, t_d+u}$ le taux de décès ajusté issu de la table construite à partir des données d'expérience.

Pour l'encadrement à 95 %, la provision obtenue avec la table correspondant au quantile à 2,5 % (donc issue de la simulation i) est notée L_0^i à la date initiale et est calculée par :

$$L_0^i = \sum_{t=1}^{\omega} F^i(t) \times (1+r_t)^{-t},$$

où $F^i(t) = \sum_{j \in J} (1+2\%)^{t-1} \prod_{u=0}^{t-1} (1 - q_{x_d(j)+u, t_d+u}^i)$. De même, la provision obtenue avec la table correspondant au quantile à 97,5 % (donc issue de la simulation s) est notée L_0^s à la date initiale et est calculée par :

$$L_0^s = \sum_{t=1}^{\omega} F^s(t) \times (1+r_t)^{-t},$$

où $F^s(t) = \sum_{j \in J} (1+2\%)^{t-1} \prod_{u=0}^{t-1} (1 - q_{x_d(j)+u, t_d+u}^s)$.

Une illustration de l'encadrement de ces provisions est présentée dans le tableau suivant. On considère à cet effet un échantillon de la population d'expérience utilisée pour la construction de la table. Cet échantillon est composé de 301 têtes masculines, âgés de 59 à 89 ans au 31/12/2009 (date initiale retenue pour l'application numérique) et dont l'âge moyen est 66 ans. Pour mémoire le régime prévoit le versement d'une rente jusqu'à 95 ans,

revalorisée annuellement de 2 %. La courbe d'actualisation retenue est la courbe de l'Institut des Actuaire⁴ au 31/12/2009.

Tableau 45 – Encadrement de la provision avec les quantiles des tables (h.)

Provision	Données d'expérience (1)	Simulations à partir des données d'expérience			
		Quantile à 2,5% (2)	Ecart relatif (2)/(1)-1	Quantile à 97,5%	Ecart relatif (3)/(1)-1
	4 752	4 683	-1,4%	4 827	1,6%

Il apparaît ainsi que les bornes de l'intervalle de confiance à 95 % au titre des fluctuations d'échantillonnage encadrent la provision *best estimate* avec un écart d'environ 1,5 %, à la hausse ou la baisse. Cet écart est proche de celui évalué ci-dessus pour l'espérance de vie générationnelle partielle (*cf.* sous-section 4.2, Tableau 44), mais demeure plus faible compte tenu de l'effet lié à l'actualisation des flux.

6. Comparaison du risque d'estimation et du risque d'avis d'expert

Les sections précédentes présentent le risque d'estimation associé aux fluctuations d'échantillonnage dans les données. Dans cette section on présente également le risque d'avis d'expert au titre du choix du jeu de tables de référence pour l'évaluation de l'évolution tendancielle de la mortalité. En d'autres termes, dans le modèle relationnel de Brass, on cherche désormais à apprécier l'impact du choix du jeu de tables de référence dans l'estimation des

$$\text{taux de décès ajustés } q_{xt} = \frac{\exp(y_{xt})}{1 + \exp(y_{xt})}, \text{ avec } y_{xt} = \hat{a} \ln \left(\frac{q_{xt}^{réf}}{1 - q_{xt}^{réf}} \right) + \hat{b}.$$

Les tables de référence retenues pour l'ajustement dans les sections précédentes sont les tables réglementaires françaises TGH/F 05. Pour mémoire, ces tables sont construites à partir de données d'expérience et ont elles-mêmes été obtenues à partir d'un positionnement par rapport aux tables de la population générale française de 1962 à 2000 (pour mémoire, les données d'expérience utilisées pour la construction de ces tables réglementaires comptent 700 000 rentiers et 2 000 000 de lignes, et couvrent la période 1993-2005, *cf.* Planchet [2006]).

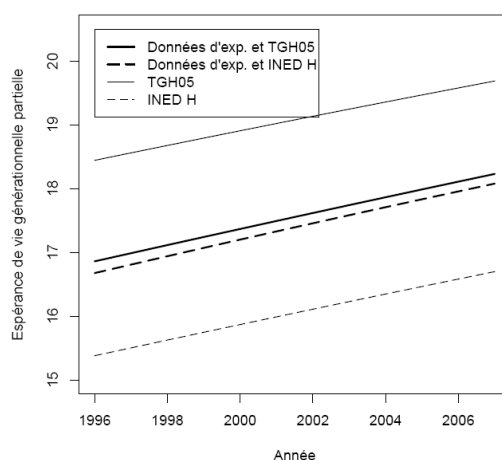
Dans cette section, on compare l'ajustement du modèle à partir de deux jeux de tables de référence différents : d'une part à partir des tables de la population assurée (tables prospectives réglementaires TGH/F 05), et d'autre part à partir des tables de la population générale française (tables prospectives de l'INED⁵). Dans les deux scénarios, la méthode d'ajustement des paramètres $\theta = (a, b)$ correspond exactement à celle présentée dans la sous-section 2.2 ; seule l'hypothèse relative aux tables de référence évolue dans le deuxième scénario.

⁴ Courbe disponible sur le site de l'Institut des Actuaire : http://www.institutdesactuaire.com/gene/link.php?doc_link=../docs/2010003190313_IA20091231.xls.

⁵ Institut National d'Études Démographiques.

Pour les illustrations, à l'image des sections précédentes, on se limite ici à la population masculine. Aussi, une fois les tables ajustées dans chacun de ces scénarios (pour mémoire, pour le premier scénario l'ajustement a déjà été réalisé dans les sections précédentes), on calcule et on compare les espérances de vie générationnelles partielles. Les résultats sont présentés sur la figure suivante (une illustration des espérances de vie issues des tables de référence y est également présentée).

Figure 48 – Évolution des esp. de vie selon le jeu de tables de référence (h.)



En premier lieu, il apparaît logiquement que l'espérance de vie issue des données d'expérience est, dans les deux scénarios, supérieure à celle issue des tables INED de la population générale (compte tenu des effets de sélection et d'anti-sélection au sein de la population d'expérience) et inférieure à celle issue des tables réglementaires TGH 05 de la population assurée (compte tenu de la prudence *a priori* intégrée dans les tables prospectives réglementaires).

Au niveau de la comparaison des deux scénarios, il apparaît en outre que le positionnement par rapport aux tables prospectives de la population générale (INED) conduit à des espérances de vie générationnelles plus faibles que celles issues du positionnement par rapport aux tables prospectives de la population assurée (tables réglementaires TGH 05). En d'autres termes, le choix d'un positionnement par rapport aux tables INED minore la baisse tendancielle de l'évolution de la mortalité de la population d'expérience étudiée (en comparaison avec le choix d'un positionnement par rapport aux tables TGH 05).

L'impact de l'écart dans les estimations selon les deux scénarios est présenté dans le tableau ci-dessous (pour l'année $t = t_0 = 1996$ uniquement). En particulier, l'impact du risque d'avis d'expert au titre du choix de la table de référence sur l'espérance de vie est comparé avec l'impact du risque d'estimation au titre des fluctuations d'échantillonnage sur l'espérance de vie.

Tableau 46 – Esp. de vie selon le jeu de tables de référence retenu (h.)

Esp. de vie gén. partielle (pour t=1996) Age: 67 / Age fin: 95		Ecart relatif avec (1)
Données d'expérience / Positionnement table population assurée (TGH05) (1)	16,9	0,0%
Quantile à 2,5% (simulations et positionnement TGH05)	16,5	-1,8%
Quantile à 97,5% (simulations et positionnement TGH05)	17,2	2,0%
Données d'expérience / Positionnement table population générale (INED H)	16,7	-1,1%
Table population assurée (TGH05)	18,4	9,4%
Table population générale (INED H)	15,4	-8,8%

Il apparaît que le changement de table de référence engendre un écart d'environ 1 % dans l'espérance de vie générationnelle partielle, soit un écart égal à la moitié environ des bornes de l'intervalle de probabilité à 95 % au titre du risque d'estimation (ces bornes correspondant à des écarts d'environ 2 %, à la hausse comme à la baisse).

Dans notre exemple, le risque d'avis d'expert au titre du choix des tables prospectives de référence (tables TGH 05 représentant le niveau et l'évolution tendancielle la mortalité des assurés ou tables INED représentant le niveau et l'évolution tendancielle la mortalité de la population générale) ressort ainsi égal à la moitié du risque d'estimation au titre des fluctuations d'échantillonnage.

7. Synthèse et conclusion

Dans ce chapitre, on propose, dans le cadre de la construction d'une table prospective, une mesure du risque d'estimation lié aux fluctuations d'échantillonnage des données et, dans une moindre mesure, du risque d'avis d'expert lié au choix de la table de référence externe.

En pratique, l'étude propose ainsi dans un premier temps (section 3) une démarche simple pour construire des intervalles de confiance ponctuels (pour un âge) et des bandes de confiance simultanées (pour une tranche d'âge) sur les taux bruts au titre des fluctuations d'échantillonnage. Ce travail permet notamment d'illustrer l'intérêt des bandes de confiance pour encadrer les taux bruts sur un ensemble d'âge (puisque dans nos illustrations les taux ajustés sont parfois en dehors des intervalles de confiance ponctuels à 95 %, alors qu'ils se situent toujours entre les bandes de confiance simultanées à 95 %).

Dans un second temps (section 4), et dans la continuité de l'encadrement simultané des taux bruts sur plusieurs âges, l'étude menée propose un encadrement des taux ajustés sur plusieurs âges (tables de mortalité), au titre des fluctuations d'échantillonnage, à partir de l'encadrement des espérances de vie générationnelles partielles. Dans notre exemple, cette méthode d'encadrement permet notamment de constater un écart d'environ 2 % entre la

table issue des données et les bornes de l'intervalle de probabilité à 95 % des tables issues des fluctuations d'échantillonnages simulées.

Dans une troisième étape (section 5), l'impact du risque d'estimation sur la provision *best estimate* de l'assureur est mesuré. À cet effet, on considère un régime de rente. Il ressort un écart d'environ 1,5 % entre la provision *best estimate* calculée à partir de la table issue des données et celles calculées à partir des tables correspondant aux bornes de l'intervalle de probabilité à 95 % des tables issues des simulations. Cet écart est proche de celui observé dans la section 4 pour les espérances de vie, mais ressort plus faible compte tenu notamment de l'effet actualisation.

Dans une dernière étape (section 6), le risque systématique d'estimation lié aux fluctuations d'échantillonnage est comparé au risque systématique d'avis d'expert lié au choix de la table de référence externe. Cette comparaison illustre que dans notre exemple le poids dans l'espérance de vie de l'incertitude lié au risque d'avis d'expert (en remplaçant la table de référence relative à la population assurée par une table de référence relative à la population générale) représente la moitié du poids de l'incertitude lié au risque d'estimation. En effet, dans le cas du risque d'avis d'expert les bornes de l'intervalle de probabilité à 95 % présentent un écart d'environ 1 % par rapport à l'espérance de vie issue des données, contre un écart de 2 % dans le cas du risque d'estimation.

Tous ces indicateurs fournissent des informations utiles pour quantifier l'impact et le coût du risque systématique d'estimation, au titre des fluctuations d'échantillonnage, et d'avis d'expert. Il apparaît notamment que ces risques systématiques ne sont pas négligeables (de l'ordre de 3 % sur les espérances de vie, et de plus de 2 % sur les provisions).

En introduction du présent travail, nous rappelions que « dans un régime de rentes, le risque viager est transféré de l'assuré à l'assureur » (*cf.* également Brouhns et Denuit [2002]). Cette présentation pourrait toutefois être précisée en indiquant que le transfert vers l'assureur porte en priorité sur le risque viager mutualisable, la mutualisation des risques entre assurés étant le métier « de base » de l'assureur. Dans ce contexte, une réflexion semble alors nécessaire sur la prise en charge du risque systématique, en particulier pour les risques systématique d'estimation et d'avis d'expert illustrés dans ce travail.

À titre d'exemple, aujourd'hui, dans les pays industrialisés la prise en charge du risque systématique d'estimation est principalement assumée par l'assureur, qui se constitue à cet effet un capital de solvabilité approprié (*cf.* Olivieri [2001]). Demain toutefois, dans les marchés en développement (et notamment au sein de la zone CIMA/FANAF) il pourrait être envisagé une prise en charge différente des risques systématiques, au sein d'une configuration consistant, par exemple, à « mutualiser » les risques systématiques de plusieurs portefeuilles d'assurés (en faisant intervenir un ensemble d'organismes assureurs, une autorité publique, un réassureur, les marchés financiers, etc.) ou à introduire un tarif, une cotisation et/ou une prestation dynamique, évoluant *a posteriori* au titre du risque systématique.

Bibliographie

- Abdi H. [2007] « Bonferroni and Sidak corrections for multiple comparisons », *N. J. Salkind (ed.). Encyclopedia of Measurement and Statistics*, Thousand Oaks, CA: Sage.
- Bland J. M., Altman D. G. [1995] « Multiple significance tests: the Bonferroni method », *BMJ*, Vol. 310.
- Booth H., Tickle L. [2008] « Mortality modeling and forecasting: A review of methods », *The Australian Demographic and Social Research Institute*, WP3.
- Brouhns N., Denuit M. [2002] « Risque de longévité et rentes viagères – I. Évolution de la mortalité en Belgique de 1880 à nos jours », *Belgian Actuarial Bulletin*, Vol. 2, No. 1.
- Brouhns N., Denuit M., Vermunt J. K. [2002] « A Poisson log-bilinear regression approach to the construction of projected lifetables », *Insurance: Mathematics and Economics*, No. 31.
- Delwarde A., Denuit M. [2006] « Construction de tables de mortalité périodiques et prospectives », *Economica*.
- Delwarde A., Kachakhidze D., Olié L., Denuit M. [2004] « Modèles linéaires et additifs généralisés, maximum de vraisemblance local et méthodes relationnelles en assurance sur la vie », *Bulletin Français d'Actuariat*, Vol. 6, No. 12.
- Hall W. J., Wellner J. A. [1980], « Confidence bands for a survival curve from censored data », *Biometrika*, Vol. 67, No. 1.
- Kamega A., Planchet F. [2011] « Hétérogénéité : mesure du risque d'estimation dans le cas d'une modélisation intégrant des facteurs observables », *Bulletin Français d'Actuariat*, Vol. 11, No. 21.
- Kamega A., Planchet F. [2010] « Mesure du risque d'estimation associé à une table d'expérience », *Les cahiers de recherche de l'ISFA*, WP2136.
- Klein J. P., Moeschberger M. L. [2005] « Survival Analysis – Techniques for Censored and Truncated Data », *Springer*, 2nd edition.
- Lee R. [2000] « The Lee-Carter Method for Forecasting Mortality, with Various Extensions and Applications », *North American Actuarial Journal*, Vol. 4, No. 1.
- Loisel S., Serant D. [2007] « In the core of longevity risk: hidden dependence in stochastic mortality models and cut-offs in prices of longevity swaps », *Cahiers de recherche de l'ISFA*, WP2044.
- Nair V. N. [1984], « Confidence Bands for Survival Functions With Censored Data: A Comparative Study », *Technometrics*, Vol. 26, No. 3.
- Olivieri A. [2001] « Uncertainty in mortality projections: an actuarial perspective », *Insurance Mathematics & Economics*, No. 29.

- Pitacco E. [2004] « Survival models in a dynamic context: a survey », *Insurance Mathematics & Economics*, No. 35.
- Planchet F. [2006], « Tables de mortalité d'expérience pour les portefeuilles de rentiers – (Tables TGH 05 et TGF 05) – Notice de présentation », *Institut des actuaires*.
- Planchet F., Kamega A. [2011], « Construction d'une table de mortalité prospective pour un régime de rentes : prise en compte du risque d'estimation », *Cahiers de recherche de l'ISFA, WP2140*.
- Planchet F., Lelieur V. [2007] « Utilisation des méthodes de Lee-Carter et log-Poisson pour l'ajustement de tables de mortalité dans le cas de petits échantillon », *Bulletin Français d'Actuariat*, Vol. 7, No. 14.
- Planchet F., Thérond P. [2007] *Pilotage technique d'un régime de rentes viagères*, Paris : Economica.
- Planchet F., Thérond P. [2006] « Modèles de Durée – Applications actuarielles », *Economica*.
- Quashie A., Denuit M. [2005] « Modèles d'extrapolation de la mortalité aux grands âges », *Institut des Sciences Actuarielles et Institut de Statistique, Université Catholique de Louvain*, WP.
- Renshaw A., Haberman S. [2003] « Lee-Carter mortality forecasting with age-specific enhancement », *Insurance: Mathematics and Economics*, Vol. 33.
- Renshaw A., Haberman S. [2005] « Mortality Reduction Factors Incorporating Cohort Effects », *Actuarial Research Paper (Cass Business School)*, No. 160.
- Thatcher A.R. [1999] « The Long-term Pattern of Adult Mortality and the Highest Attained Age. », *Journal of the Royal Statistical Society* 162 Part 1: 5-43.

SYNTHESE ET CONCLUSION GENERALE

Synthèse de l'étude

Dans un marché de l'assurance vie en Afrique subsaharienne francophone à la traîne, mais promis à un bel avenir en cas d'émergence de solutions techniques et commerciales endogènes (à ce sujet, le premier chapitre – partie I – de cette étude présente un état des lieux et les défis du marché), le travail présenté ici propose des outils théoriques et opérationnels adaptés à son développement. Cette démarche s'inscrit en parallèle des actions entreprises par l'autorité de contrôle régionale (la CIMA) pour fournir aux assureurs de la région des outils adaptés.

En effet, la CIMA a initié des travaux pour la construction de nouvelles tables réglementaires d'expérience, ce qui a permis de fournir des références fiables et pertinentes pour la mortalité d'expérience de la région (les méthodes et hypothèses de construction de ces tables sont présentées au deuxième chapitre – partie I –). Toutefois, certaines problématiques techniques utiles n'ont pas été développées dans ces travaux de construction. Les parties II et III de ce travail leur accordent alors une attention particulière.

Ainsi, d'une part, la partie II du travail présenté ici permet de fournir des outils pour tenir compte des différences de mortalité entre pays de la région, tout en limitant les risques systématiques liés aux fluctuations d'échantillonnage (dues à la petite taille des échantillons de données par pays). Il apparaît notamment que si la modélisation indépendante de chaque pays n'est pas appropriée (comme indiqué dans le troisième chapitre – partie II –), les modèles d'hétérogénéité à facteurs observables, tels que le modèle de Cox ou de Lin et Ying, permettent d'atteindre cet objectif (comme indiqué dans le quatrième chapitre – partie II –). On précise toutefois ici que ces modèles d'hétérogénéité ne permettent pas de supprimer le risque systématique lié aux fluctuations d'échantillonnage lors de l'estimation du modèle, ils engendrent seulement une réduction de ce risque en contrepartie d'une augmentation du risque systématique lié au choix du modèle.

D'autre part, la partie III permet de fournir des outils pour modéliser la mortalité d'expérience future dans la région. En absence de données sur les tendances passées de la mortalité d'expérience, ni le modèle classique de Lee-Carter ni ses extensions ne sont applicables. Une solution basée sur un ajustement paramétrique, une hypothèse sur la forme de l'évolution du niveau de mortalité (évolution linéaire ou exponentielle) et un avis d'expert sur l'espérance de vie générationnelle à un âge donné est alors proposée (ces travaux sont présentés au cinquième chapitre – partie III –, et s'appuient sur le modèle de Bongaarts). Ensuite, dans un second temps, en supposant disposer de données sur les tendances passées (ce qui pour mémoire n'est pas le cas à ce stade dans la région, mais devrait l'être dans les prochaines années), on présente (dans le sixième chapitre – partie III –) la modélisation de la mortalité future à partir d'une référence de mortalité externe et une analyse des risques

systematiques associés (risques liés aux fluctuations d'échantillonnage et au choix de la référence de mortalité).

Au final, en présence de données limitées, il apparaît qu'en cas de modélisation de l'hétérogénéité ou de la mortalité future, la segmentation de la population (par pays dans le premier cas ou par année dans le second cas) conduit à favoriser l'émergence de risques systematiques (liés aux fluctuations d'échantillonnage, au choix du modèle, à l'avis d'expert sur les références de mortalité, à l'avis d'expert sur l'espérance de vie attendue, etc.) dans les évaluations réalisées.

Si la prise en compte d'une segmentation est nécessaire pour disposer d'une mesure fine de la mortalité (conformément à ce qui est présenté dans Planchet et Leroy [2009]), en pratique il apparaît donc que cette démarche expose à certains risques d'erreur non négligeables, croissants avec la finesse de la segmentation retenue. Une réflexion s'impose alors sur le niveau optimal de segmentation à retenir.

Perspectives pour la recherche future : choix du niveau de segmentation pertinent pour modéliser la mortalité

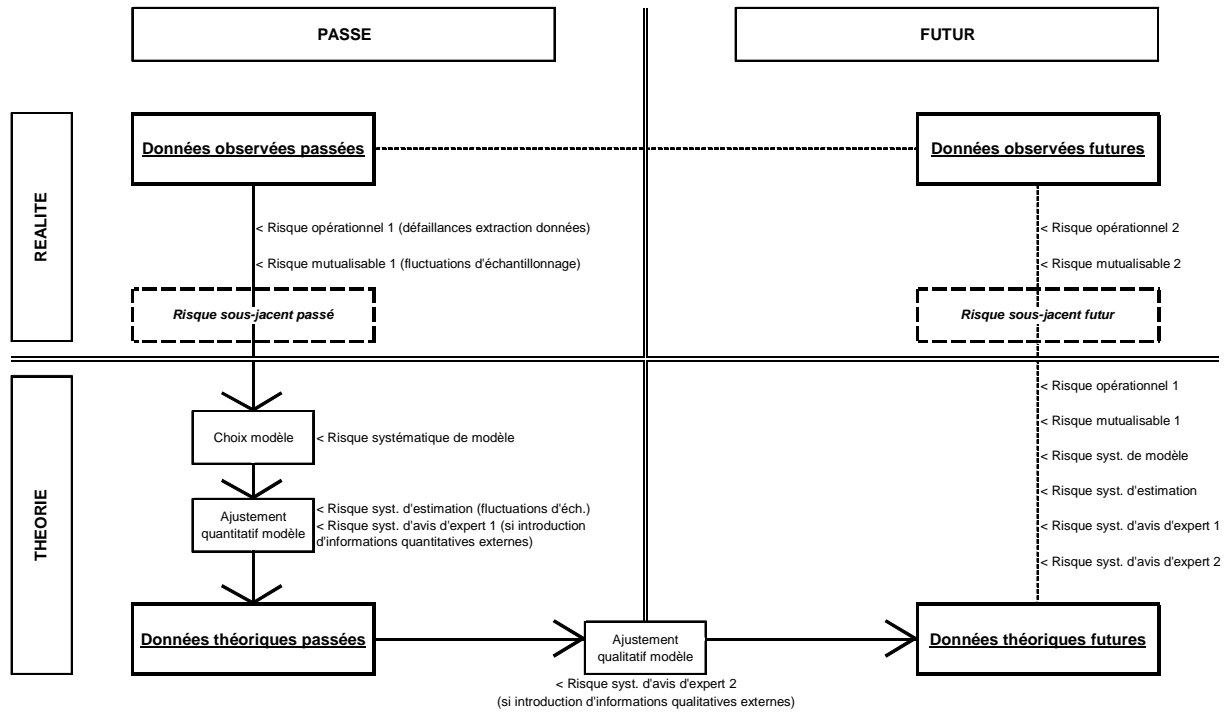
Dans le travail présenté ici, des solutions sont proposées dans la partie II pour optimiser le niveau de segmentation de la population au titre de la prise en compte de l'hétérogénéité.

Le critère d'optimisation retenu à cet effet est le niveau du risque d'estimation (lié aux fluctuations d'échantillonnage). En d'autres termes, les solutions proposées ici pour segmenter la population permettent de limiter le niveau du risque systematique d'estimation. En pratique toutefois, ces solutions développent en contrepartie un risque systematique de modèle, comme évoqué ci-dessus.

Les solutions retenues ici pour segmenter la population sont donc satisfaisantes en termes de risque d'estimation, mais ne le sont pas en termes de risque de modèle. Dans le cadre de travaux futurs, il pourrait ainsi être utile d'envisager des solutions permettant d'optimiser le niveau de segmentation d'une population sur la base d'un ensemble de critères, et non uniquement sur la base du seul risque systematique d'estimation.

À cette fin, la Figure 49 ci-après présente les différentes sources de risques associées à la modélisation de la mortalité d'expérience.

Figure 49 - Organisation de la modélisation de la mortalité et présentation des risques associés



Dans le cadre d'une modélisation de la mortalité il apparaît ainsi plusieurs natures de risques, dont (*cf.* détails dans Booth et Tickle [2008] par exemple) :

- un risque opérationnel, au titre de procédures défectueuses, dues au personnel ou aux systèmes d'informations, dans le cadre de l'extraction des données utilisées pour la modélisation de la mortalité (sur ce point on pourra notamment consulter l'article de Decroocq [2011] relatif à l'importance de la qualité des données) ;
- un risque mutualisable, dû aux fluctuations de l'espérance empirique autour de l'espérance mathématique, compte tenu des fluctuations d'échantillonnage ;
- un risque systématique de modèle, dû à une mauvaise spécification du modèle ;
- un risque systématique d'estimation des paramètres, dû à une mauvaise estimation des paramètres du modèle, compte tenu par exemple des fluctuations d'échantillonnage ;
- un risque systématique d'avis d'expert, dû à un mauvais jugement de l'expert lors de la prise en compte d'informations externes (soit dans le cadre de l'ajustement statistique du modèle, soit dans le cadre de l'ajustement qualitatif du modèle par avis d'expert).

On comprend facilement que le poids de ces risques dans la mesure de la mortalité est croissant avec la finesse de la segmentation de la population. En effet, à titre d'exemple, une segmentation fine expose à une plus grande exigence dans la précision des données (risque opérationnel), à une plus faible mutualisation de la population (risque mutualisable), à une plus grande complexité des modèles (risque de modèle), à une dégradation de la qualité statistique des estimateurs (risque d'estimation) et à une augmentation de la prise en compte d'informations externes (risque d'avis d'expert).

La prise en compte de ces différentes sources de risque dans le cadre d'une optimisation du niveau de segmentation est donc justifiée. À cet effet, la mesure de ces risques doit être traitée avec soin.

Perspectives pour la recherche future : mesure et gestion des risques associés à la modélisation de la mortalité

La recherche d'un niveau de segmentation pertinent pour évaluer finement la mortalité renvoie donc à des problématiques de mesure des risques associés (à l'image de ce que l'on observe en gestion actif-passif dans le cadre de la recherche d'un niveau de segmentation pertinent pour les allocations stratégiques d'actifs, *cf.* Meucci [2007] ou Faleh [2010]). Une attention particulière doit donc être accordée à la mesure du risque opérationnel, du risque mutualisable et des risques systématiques.

En particulier, concernant les risques systématiques, principale source de risque étudiée dans cette étude, il existe de nombreuses perspectives de recherches suite aux travaux présentés, que ce soit pour le risque systématique d'estimation, de modèle ou d'avis d'expert.

Le travail présenté ici s'attache en de nombreux points à mesurer le risque systématique d'estimation lié aux fluctuations d'échantillonnage (dans un contexte de modélisation de l'hétérogénéité et de la mortalité future). À cet effet, une méthode de ré-échantillonnage des taux de décès bruts ou des nombres de décès a été retenue, sans tenir compte de l'interdépendance entre les âges ou les périodes dans les décès. Aussi, dans le cadre de développements futurs, une analyse du risque d'estimation pourrait être réalisée en tenant compte de ces interdépendances dans le ré-échantillonnage, étant donné qu'elles sont vérifiées empiriquement par plusieurs études (*cf.* par exemple Booth et Tickle [2008] et Loisel et Serant [2007]).

Par ailleurs, dans le travail réalisé ici, une illustration introductive du risque de modèle est également présentée (dans le quatrième chapitre, au sein duquel pour mémoire trois modèles sont comparés dans un contexte d'hétérogénéité). Dans le cadre de développements futurs, une analyse approfondie du risque systématique de modèle pourrait être utile, sachant qu'il s'agit d'un sujet qui peut se révéler complexe (comme le rappellent Booth et Tickle [2008]).

De même, il semble également exister un espace important pour la recherche future sur la mesure du risque systématique d'avis d'expert (sur le niveau et la forme de l'évolution attendue de la durée de vie humaine, sur le choix de références externes, etc.). Ici également, il s'agit d'un sujet complexe, qui de surcroît fait intervenir des jugements personnels (par définition subjectifs) et des informations externes (par construction mal maîtrisées).

Dans la continuité des travaux réalisés ici, il apparaît ainsi d'importantes perspectives de recherche pour les trois natures de risques systématiques identifiés : le risque d'estimation, le risque de modèle et le risque d'avis d'expert. En tant que risques systématiques, ces risques ne se mutualisent pas : il s'agit ainsi de risques potentiellement dangereux, et susceptibles d'avoir un impact non négligeable pour l'assureur. En pratique toutefois, une réflexion pourrait être utile sur la responsabilité de l'assureur dans la prise en charge de ces risques systématiques. En effet, si aujourd'hui ces risques sont régulièrement gérés par l'assureur à travers la constitution de provisions prudentes et d'un capital de solvabilité approprié, demain il pourrait être envisagé que les assureurs privilégient un partage et/ou un transfert de tout ou partie de ces risques systématiques (à travers les marchés financiers, les réassureurs, les autorités publiques, les assurés eux-mêmes, etc.), afin de ne conserver que la prise en charge du risque mutualisable (sur ces problématiques associées au partage des risques, on pourra notamment consulter l'article de Planchet et Leroy [2011]).

Perspectives pour la recherche future : autres modélisations et autres risques

La mesure des risques liés à la modélisation de la durée de vie humaine est principalement présentée dans ce travail dans le cadre de l'évaluation *best estimate* des engagements d'assurance, c'est-à-dire dans le cadre de la meilleure évaluation possible des flux de trésorerie associés aux engagements d'assurance.

Cette nouvelle approche pour évaluer les engagements d'assurance s'inscrit dans le cadre de la triple mutation que connaît aujourd'hui le secteur à l'échelle internationale : mutation prudentielle (Solvabilité 2, en Europe), comptable (IFRS) et de la communication financière (MCEV, en Europe). L'Afrique subsaharienne francophone ne saurait échapper aux principes de ces évolutions internationales, bien que le contexte y soit très différent, notamment en termes de marchés financiers, de système d'information et d'identification des risques. Aussi, des adaptations à la réalité du marché de la région sont indispensables, et c'est dans ce contexte que de nombreuses perspectives pour la recherche future existent, en complément de celles présentées ci-avant.

Ainsi par exemple, en se limitant aux dispositions quantitatives du référentiel prudentiel Solvabilité 2, il apparaît d'importantes pistes de réflexion pour la recherche future sur l'identification et l'analyse de risques susceptibles d'impacter significativement la solvabilité des organismes privés d'assurance vie dans la région¹ :

- risques opérationnels (risque de perte résultant de procédures internes défectueuses ou inadaptées, dues au personnel ou aux systèmes d'informations, ou d'événements extérieurs), compte tenu de son importance dans la pérennité des organismes (*cf.* Jezzini [2005]) et de la forte escroquerie à l'assurance, du manque de personnel qualifié, de l'insuffisance des données, des problèmes de gouvernance, de l'évolution du cadre réglementaire, etc. dans la région (concernant les méthodes de modélisation du risque opérationnel on pourra notamment consulter Gamonet [2009], Karam et Planchet [2011] ou encore Jezzini [2005]) ;
- risque catastrophe en vie (risque de perte résultant de l'incertitude importante, liée à des événements extrêmes ou irréguliers, qui pèse sur les hypothèses retenues en matière de prix et de provisionnement), compte tenu des risques élevés de conflits armés, de crises sanitaires, de crises alimentaires, d'accidents dans les transports, etc. dans la région ;
- risque de rachats (risque de perte résultant de fluctuations affectant le niveau, l'évolution tendancielle ou la volatilité des taux de cessation, d'échéance, de renouvellement et de rachats des polices), compte tenu de la spécificité du comportement microéconomique des agents dans la région (sur ce point on pourra consulter Direction des Assurances (Bénin) [2007] pour un exemple d'analyse sur les comportements de rachats dans la région) ;
- risque de change (sensibilité de la valeur des actifs et des passifs aux changements affectant le niveau ou la volatilité du taux de change), compte tenu de l'incertitude dans la parité du franc CFA (monnaie de la région) avec l'euro et de ses conséquences en termes d'inflation ;

¹ La définition des risques présentés ici entre parenthèses est reprise de l'article 105 de la Directive dans Parlement européen et Conseil [2009].

- risque de taux d'intérêt (sensibilité de la valeur des actifs et des passifs aux changements affectant le niveau ou la volatilité des taux d'intérêt), compte de la forte incomplétude des marchés financiers dans la région ;
- etc.

En outre, il apparaît une difficulté transversale, susceptible de concerner toutes ces différentes sources de risques : l'imperfection des données. Aussi, dans tous les cas, cette problématique devra être traitée avec le plus grand soin et une attention particulière devra être accordée à la recherche de solutions adaptées à cette situation, à l'image de ce qui est proposé dans Félix et Planchet [2009], article dans lequel les auteurs proposent une démarche pour modéliser, entre autres, les risques de souscription d'un portefeuille de garanties décès en présence d'informations incomplètes.

Bibliographie

- Booth H., Tickle L. [2008] « Mortality modeling and forecasting: A review of methods », *The Australian Demographic and Social Research Institute, WP3*.
- Decroocq J.-F. [2011], « La qualité des données », *VaRM*.
- Direction des Assurances (République du Bénin) [2007], « Étude sur les ruptures anticipée de contrats d'assurance sur la vie », *République du Bénin, Ministère du développement, de l'économie et des finances, Direction générale de l'économie*.
- Faleh A. [2010] *Allocation stratégique d'actifs et niveau de segmentation pertinent*, Support de groupe de travail, ENASS.
- Félix J.-P., Planchet F. [2009] « Mesure des risques de marché et de souscription vie en situation d'information incomplète pour un portefeuille de prévoyance », *Bulletin Français d'Actuariat, Vol. 9, No. 18*.
- Gamonet J. [2009] *Modélisation du risque opérationnel dans l'assurance*, Mémoire d'actuariat, CEA.
- Jezini M. [2005], « Revue de la littérature : Risque Opérationnel », *Université d'Avignon et des Pays de Vaucluse*.
- Karam E., Planchet F. [2011], « Operational Risks in Financial Sectors », *Cahiers de recherche de l'ISFA, WP21XX*.
- Loisel S., Serant D. [2007], « In the core of longevity risk: hidden dependence in stochastic mortality models and cut-offs in prices of longevity swaps », *Cahiers de recherche de l'ISFA, WP2044*.
- Meucci A. [2007] « Risk and Asset Allocation », *Springer, Reprint of the 2007 edition*.
- Parlement européen, Conseil [2009] « Directive du Parlement européen et du Conseil sur l'accès aux activités de l'assurance et de la réassurance et leur exercice (Solvabilité II) (Refonte) », *Union européenne, PE-CONS 3643/6/09, REV 6*.
- Planchet F., Leroy G. [2009] « Quel niveau de segmentation pertinent ? », *La Tribune de l'Assurance, n. 142*.
- Planchet F., Leroy G. [2011] « Problématiques associées au partage des risques », *La Tribune de l'Assurance, n. 160*.

SIGLES

BCEAO : Banque Centrale des États de l’Afrique de l’Ouest. Banque centrale des États de l’UEMOA.

BEAC : Banque des États de l’Afrique Centrale. Banque centrale des États de la CEMAC.

BRVM : Bourse Régionale des Valeurs Mobilières (*cf.* « www.brvm.org »). Couvre les États de l'UEMOA.

BVMAC : Bourse des Valeurs Mobilières de l'Afrique Centrale (*cf.* « www.bvm-ac.com »). Couvre les États de la CEMAC.

CEMAC : Communauté Économique et Monétaire de l'Afrique Centrale (*cf.* « www.cemac.int »). Union économique et monétaire dont les États membres sont les suivants : Cameroun, Centrafrique, Congo, Gabon, Guinée Équatoriale, Tchad.

CICA : Conférence Internationale de Contrôle des Assurances. Organisation au sein de laquelle trois conventions sont signées en 1962, 1973 et 1990, avant d’être remplacée par la CIMA. Les États signataires des conventions sont les suivants : Bénin, Burkina Faso, Cameroun, Centrafrique, Côte d'Ivoire, France (signataire de la première convention uniquement), Congo Brazzaville, Gabon, Guinée Bissau, Guinée Équatoriale, Mali (adhère à la CICA en 1975, après la deuxième convention), Mauritanie (sort de l’organisation après la deuxième convention en 1973), Madagascar (sort de l’organisation après la deuxième convention en 1973), Niger, Sénégal, Tchad et Togo (adhère à la CICA en 1968, après la première convention).

CIAN : Conseil français des investisseurs en Afrique (*cf.* « www.cian.asso.fr »).

CIMA : Conférence Interafricaine des Marchés de l'Assurance (*cf.* « www.cima-afrique.org »). Organisme communautaire de l’industrie des assurances dans les États africains, dont les États membres sont les suivants : Bénin, Burkina Faso, Cameroun, Centrafrique, Congo Brazzaville, Côte d'Ivoire, Gabon, Guinée Bissau (adhère à la CIMA en 2002), Guinée Équatoriale, Mali, Niger, Sénégal, Tchad et Togo. Pour information, les Comores sont signataires mais n’ont pas encore ratifié le Traité de la CIMA.

CIPRES : Conférence Interafricaine de la Prévoyance Sociale (*cf.* « www.lacipres.org »). Organisme de contrôle et d’appui technique aux Caisses Afrique de Sécurité sociale, dont les États membres sont les suivants : Bénin, Burkina Faso, Cameroun, Centrafrique, Comores, Congo Brazzaville, Côte d'Ivoire, Gabon, Guinée Équatoriale, Mali, Niger, République Démocratique du Congo, Sénégal, Tchad et Togo.

CNUCED : Conférence des Nations Unies sur le Commerce et le Développement. Organisation qui vise à intégrer les pays en développement dans l'économie mondiale de façon à favoriser leur essor.

CRCA : Commission Régionale de Contrôle des Assurances. Organe régulateur de la CIMA qui assure le contrôle des sociétés, assure la surveillance générale et concourt à l'organisation des marchés nationaux.

DSX : Douala Stock Exchange (cf. « www.douala-stock-exchange.com »). Couvre le Cameroun (État membre de la CEMAC par ailleurs).

FANAF : Fédération des sociétés d'Assurances de droits National AFricaines (cf. « www.fanaf.com »). Association professionnelle réunissant, au 31 décembre 2007, 139 Sociétés d'Assurances et de Réassurances plus un Fonds de Garantie Automobile (FGA) opérant dans les pays suivants: Algérie, Afrique du Sud, Bénin, Burkina Faso, Burundi, Cameroun, Centrafrique, Congo Brazzaville, Côte d'Ivoire, Gabon, Ghana, Guinée Conakry, Madagascar, Mali, Maroc, Mauritanie, Nigeria, Niger, République Démocratique du Congo, Rwanda, Sénégal, Tchad, Togo, Tunisie.

FFSA : Fédération Française des Sociétés d'Assurances (cf. « www.ffsa.fr »). Association professionnelle réunissant 90 % du marché français de l'assurance.

FGA : Fonds de Garantie Automobile.

GFA : Groupement français d'assurances.

IFRS : International Financial Reporting Standards (cf. « www.ifrs.org »).

IIA : Institut International des Assurances (cf. « www.iiayaounde.com »). Établissement de formation professionnelle à caractère international. Il a son siège à Yaoundé (Cameroun) et est ouvert aux ressortissants des États membres de la CIMA.

MCEV : Market Consistent Embedded Value (cf. « www.cfoforum.nl »).

MCR : Mutuelle Centrale de Réassurance.

OIT ; Organisation Internationale du Travail.

OMS : Organisation Mondiale de la Santé.

PIB : Produit Intérieur Brut.

UEMOA : Union Économique et Monétaire Ouest Africaine (cf. « www.uemoa.int »). Union économique et monétaire dont les États membres sont les suivants : Bénin, Burkina Faso, Côte d'Ivoire, Guinée Bissau, Mali, Niger, Sénégal, Togo.

BIBLIOGRAPHIE GÉNÉRALE

- Abdi H. [2007] « Bonferroni and Sidak corrections for multiple comparisons », *N. J. Salkind (ed.). Encyclopedia of Measurement and Statistics*, Thousand Oaks, CA: Sage.
- ABS [2009] « Causes of death », *Australian Bureau of Statistics*.
- Adamah-Tassah D. A. [2000] *Assurance et secteur informel : quelles stratégies de pénétration ?*, Mémoire d'assurance, IIA.
- Agbohoun N. [2008] « Le franc CFA et l'euro contre l'Afrique », *Solidarité mondiale, 2^{ème} édition*.
- Aku Sama A. [1996] *Le marché togolais de l'assurance vie*, Mémoire d'assurance, IIA.
- Alho J. M. [2007] « Méthodes utilisées pour établir les projections relatives à la mortalité – Distributions prédictives de la mortalité future », *ISSA, Quinzième Conférence internationale des actuaires et statisticiens de la sécurité sociale*.
- Alla K. A. [2000] *L'escroquerie à l'assurance et ses implications, quelles solutions ?*, Mémoire d'assurance, IIA.
- BAfD/OCDE (Banque Africaine de Développement/Organisation de Coopération et de Développement Économique) [2008] « Perspectives économiques en Afrique 2007/2008 », *OCDE*.
- Banque de France [2008] « Rapport Zone franc - 2007 », *Banque de France - La politique et les agrégats monétaires dans l'UEMOA*.
- Barbi E., Caselli G., Vallin J. [2003] « Hétérogénéité des générations et âge extrême de la vie », *Institut Nationale d'Études Démographiques / Population, Vol. 58, No. 2006-1*.
- Barriou P., Bensusan H., El Karoui N., Hillairet C., Loisel S., Ravanelli C., Salhi Y. [2010] « Understanding, Modelling and Managing Longevity Risk: Key Issues and Main Challenges », *Scandinavian Actuarial Journal*.
- Betoe Minko M. [1986] *La promotion de l'assurance vie en Afrique : l'exemple du Cameroun*, Mémoire d'assurance, IIA.
- Bland J. M., Altman D. G. [1995] « Multiple significance tests: the Bonferroni method », *BMJ*, Vol. 310.
- Bleu V. [1984] *Influence des paramètres socio-économiques sur le développement de l'assurance vie – Ex : Cameroun et Côte d'Ivoire*, Mémoire d'assurance, IIA Yaoundé.
- Bongaarts J. [2004] « Long-Range Trends in Adult Mortality: Models and Projections Methods », *Population Council, WP192*.

- Booth H., Tickle L. [2008] « Mortality modeling and forecasting: A review of methods », *The Australian Demographic and Social Research Institute, WP3*.
- Bourbonnais R. [2004] *Économétrie*, Dunod, 5^e édition.
- Brouhns N., Denuit M. [2002] « Risque de longévité et rentes viagères – I. Évolution de la mortalité en Belgique de 1880 à nos jours », *Belgian Actuarial Bulletin, Vol. 2, No. 1*.
- Brouhns N., Denuit M., Vermunt J. K. [2002] « A Poisson log-bilinear regression approach to the construction of projected lifetables », *Insurance: Mathematics and Economics, No. 31*.
- Cao H. [2005] *A comparison Between the Additive and Multiplicative Risk Models*, Mémoire d'actuariat, Université Laval.
- Cox D. R. [1972] « Regression Models and Life-Tables », *Journal of the Royal Society. Series B (Methodological), Vol. 34, No. 2*.
- CIAN [2008] *Résister aux sollicitations indues dans le domaine fiscal*, Programme de la réunion de la Commission Anti-corruption du 20 novembre 2008, CIAN.
- CIMA [1992], « Traité instituant la CIMA », *CIMA*.
- CIMA [2011] « Note sur la confection des nouvelles tables de mortalité », *CIMA*.
- CIPRES [1993], « Traité instituant la CIPRES », *CIPRES*.
- CIPRES [2007], « Les systèmes de protection sociale en Afrique : Zone CIPRES », *CIPRES*.
- Conférence internationale des statisticiens du travail [1993] « Résolution concernant les statistiques de l'emploi dans le secteur informel adoptée par la quinzième Conférence internationale des statisticiens du travail », *OIT*.
- Daniel J.-P. [2007] « L'assurance islamique gagne du terrain », *La Tribune de l'assurance 111*.
- Debayou A. S. [1988] *Quelle assurance vie pour l'Afrique ?*, Mémoire d'assurance, IIA.
- Decroocq J.-F. [2011], « La qualité des données », *VaRM*.
- Delwarde A., Denuit M. [2006] « Construction de tables de mortalité périodiques et prospectives », *Economica*.
- Delwarde A., Kachakhidze D., Olié L., Denuit M. [2004] « Modèles linéaires et additifs généralisés, maximum de vraisemblance local et méthodes relationnelles en assurance sur la vie », *Bulletin Français d'Actuariat, Vol. 6, No. 12*.
- Denuit M., Charpentier A. [2004] « Mathématiques de l'assurance non-vie – Tome I : principes fondamentaux de théorie du risque », *Economica*.

- Diakit  D. [2008] « Universalit  des valeurs et id al d'humanit  en Afrique : t moignages d'explorateurs », *extrait de* Konar  A. B. (sous la direction de) [2008] « Petit pr cis de remise   niveau sur l'histoire africaine   l'usage du pr sident Sarkozy », *La D couverte*.
- Diallo D. S. [2008] « LA CIMA ou l'histoire d'un mod le d'int gration dans les assurances ».
- DID (Desjardins – D veloppement international) [2011], «  tude sur la microassurance dans la zone CIMA », *Desjardins – D veloppement international*.
- Diop C. A. [1979] « Nations n gres et culture », *Pr sence Africaine*, 4^{ me}  dition.
- Direction des Assurances (R publique du B nin) [2007], «  tude sur les ruptures anticip e de contrats d'assurance sur la vie », *R publique du B nin, Minist re du d veloppement, de l' conomie et des finances, Direction g n rale de l' conomie*.
- Droesbeke J.-J., Fichet B., Tassi P. [1989] ( diteurs) « Analyse statistique des dur es de vie », *Economica*.
- Ewald F. [2010] « La naissance du risque social », *Risques, Les cahiers de l'assurance* 81-82.
- FANAF [2011], « Le march  de l'assurance en Afrique – Donn es 2005   2009 », *FANAF*.
- FFSA [2011], « Rapport annuel 2009 », *FFSA*.
- Faleh A. [2010] *Allocation strat gique d'actifs et niveau de segmentation pertinent*, Support de groupe de travail, ENASS.
- F lix J.-P., Planchet F. [2009] « Mesure des risques de march  et de souscription vie en situation d'information incompl te pour un portefeuille de pr voyance », *Bulletin Fran ais d'Actuariat, Vol. 9, No. 18*.
- Fr my D., Fr my M. [2005] « Quid 2006 », *Robert Laffont*.
- Gamonet J. [2009] *Mod lisation du risque op rationnel dans l'assurance*, M moire d'actuariat, CEA.
- Georges I. [2000] *Assurance et secteur informel : quelles strat gies de p n tration*, M moire d'assurance, IIA.
- Guibert Q. [2010] *Analyse de la solvabilit  d'un r gime de retraite suppl mentaire*, M moire d'actuariat, ISFA.
- Haman Bello I. [1984] *Essai d'analyse de la faiblesse de l'assurance vie en Afrique*, M moire d'assurance, IIA.
- Hampat  B  A. [1972] « aspects de la civilisation africaine », *Pr sence Africaine*.
- Hall W. J., Wellner J. A. [1980], « Confidence bands for a survival curve from censored data », *Biometrika*, Vol. 67, No. 1.

- Hill C., Com-Nougé C., Kramar A., Moreau T., O'Quigley J., Senoussi R., Chastang C. [1990] « Analyse statistique des données de survie », *Inserm, Médecine – Sciences, Flammarion*.
- Hugon P. [2009] « L'économie de l'Afrique », *La Découverte, 6^{ème} édition*.
- Jezzini M. [2005], « Revue de la littérature : Risque Opérationnel », *Université d'Avignon et des Pays de Vaucluse*.
- Juompan-Yakam C. [2010] « Retraites – Un continent en quête de cotisants », *Jeune Afrique 2590*.
- Kamega A. [2006] *Impact de la sursuicidité en Bretagne sur les garanties temporaires décès*, Mémoire d'actuariat, EURIA.
- Kamega A. [2010] « Défis de l'assurance vie en Afrique noire francophone », *La Tribune de l'Assurance 143*.
- Kamega A. [2011] *Quelle anticipation de l'espérance de vie dans les portefeuilles d'assurance ?*, Chaire transitions démographiques, transitions économiques, Intervention du 27/01/2011.
- Kamega A., Planchet F. [2010], « Mesure du risque d'estimation associé à une table d'expérience », *Cahiers de recherche de l'ISFA, WP2136*.
- Kamega A., Planchet F. [2011A], « Analyse et comparaison des populations générale et assurée en Afrique subsaharienne francophone pour anticiper la mortalité future », *Cahiers de recherche de l'ISFA, WP2138*.
- Kamega A., Planchet F. [2011B], « Hétérogénéité : mesure du risque d'estimation dans le cas d'une modélisation intégrant des facteurs observables », *Bulletin Français d'Actuariat, Vol. 11, No. 21*.
- Karam E., Planchet F. [2011], « Operational Risks in Financial Sectors », *Cahiers de recherche de l'ISFA, WP21XX*.
- Kassim Z. [2008] « Takaful - Concepts et pratique », *Le Réassureur Africain, 22^{ème} édition*.
- Ki-Zerbo J. [1978] « Histoire de l'Afrique noire - D'hier à demain », *Hatier*.
- Klein J. P., Moeschberger M. L. [2005] « Survival Analysis – Techniques for Censored and Truncated Data », *Springer, 2nd edition*.
- Kone M.G.K. [2007A] « État des lieux de l'assurance vie dans les pays de la CIMA », *CIMA*.
- Kone M.G.K. [2007B] « État généraux de l'assurance vie - Rapport final », *CIMA/FANAF*.
- Lee R. [2000], « The Lee-Carter Method for Forecasting Mortality, with Various Extensions and Applications », *North American Actuarial Journal, Vol. 4, No. 1*.
- Lin D. Y., Wei L.J., Ying Z. [1993] « Checking the Cox model with cumulative sums of martingale-based residuals », *Biometrika, n. 80*.

- Lin D. Y., Ying Z. [1994] « Semiparametric analysis of the additive risk model », *Biometrika*, n. 81.
- Loisel S., Serant D. [2007], « In the core of longevity risk: hidden dependence in stochastic mortality models and cut-offs in prices of longevity swaps », *Cahiers de recherche de l'ISFA*, WP2044.
- Lopez O. [2007] *Réduction de dimension en présence de données censurées*, Thèse de doctorat, Université de Rennes 1.
- Martinussen T., Scheike T. H. [2006] « Dynamic Regression Models for Survival Data », *Springer*.
- McCullagh P., Nelder J. A. [1989] « Generalized Linear Models », *Monographs on Statistics and Applied Probability 37*, Chapman & Hall/CRC, 2nd edition.
- Meslé F., Vallin J. [2007] « De la transition épidémiologique à la transition sanitaire », *Médecine Tropicale*, No. 67.
- Meucci A. [2007] « Risk and Asset Allocation », *Springer, Reprint of the 2007 edition*.
- Nabeth M. [2006] « Micro-assurance – Défis, mise en place et commercialisation », *L'Argus de l'assurance*.
- Nabeth M. [2007] « Assurance et microassurance en Afrique », *Risques, Les cahiers de l'assurance 71*.
- Nair V. N. [1984], « Confidence Bands for Survival Functions With Censored Data: A Comparative Study », *Technometrics*, Vol. 26, No. 3.
- Ndembou S. [2007] « La solidarité africaine relève de la générosité », *extrait de Courade G. (ouvrage dirigé par) [2007] « L'Afrique des idées reçues », Belin*.
- Ngbwa J.-C. [2007] « Apport de la CIMA à la solvabilité des entreprises d'assurances », *Risques, Les cahiers de l'assurance 71*.
- Nonyu Moutassie É. [2005] « Présentation de la CIMA », *CIMA*.
- Olivieri A. [2001] « Uncertainty in mortality projections: an actuarial perspective », *Insurance Mathematics & Economics*, No. 29.
- Omran A. R. [2005] « The Epidemiologic Transition: A Theory of the Epidemiology of Population Change », *The Milbank Quarterly*, Vol. 83, No. 4, Reprinted.
- Omvane Nkouélé F.-X. [1988] *L'assurance sur la vie et le financement de l'économie dans les pays de la CICA*, Mémoire d'assurance, IIA Yaoundé.
- Ouédraogo A. [2008] « Faiblesses actuelles des compagnies d'assurances et moyens susceptibles d'être mobilisés pour leur consolidation pérenne », *CIMA*.
- Ouédraogo O. [1999] « Plaidoyer macroéconomique pour l'Afrique », *Karthala*.

- Parlement européen, Conseil [2009] « Directive du Parlement européen et du Conseil sur l'accès aux activités de l'assurance et de la réassurance et leur exercice (Solvabilité II) (Refonte) », *Union européenne*, PE-CONS 3643/6/09, REV 6.
- Pitacco E. [2004] « Survival models in a dynamic context: a survey », *Insurance Mathematics & Economics*, No. 35.
- Planchet F. [2006], « Tables de mortalité d'expérience pour les portefeuilles de rentiers – (Tables TGH 05 et TGF 05) – Notice de présentation », *Institut des actuaires*.
- Planchet F. [2007], « Prospective models of mortality with forced drift – Application to the longevity risk for life annuities », *Proceedings of the 11th IME Congress*.
- Planchet F. [2009] « Provisionnement *best estimate* et risque arrêt de travail », *La Tribune de l'Assurance*, n. 140.
- Planchet F., Juillard M., Faucillon L. [2006] « Quantification du risque systématique de mortalité pour un régime de rentes en cours de service », *Assurance et gestion des risques*, Vol. 75.
- Planchet F., Kamega A. [2011A], « Construction d'une table de mortalité prospective pour un régime de rentes : prise en compte du risque d'estimation », *Cahiers de recherche de l'ISFA*, WP2140.
- Planchet F., Kamega A. [2011B], « Mortalité prospective en cas de petits échantillons : modélisation à partir d'informations externes en utilisant l'approche de Bongaarts », *Cahiers de recherche de l'ISFA*, WP2141.
- Planchet F., Kamega A., Ziguélé M. [2010], « Confection des tables de mortalité réglementaires pour la zone CIMA », *CIMA/WINTER & Associés (confidentiel)*.
- Planchet F., Lelieur V. [2007] « Utilisation des méthodes de Lee-Carter et log-Poisson pour l'ajustement de tables de mortalité dans le cas de petits échantillon », *Bulletin Français d'Actuariat*, Vol. 7, No. 14.
- Planchet F., Leroy G. [2009] « Quel niveau de segmentation pertinent ? », *La Tribune de l'Assurance*, n. 142.
- Planchet F., Leroy G. [2011] « Problématiques associées au partage des risques », *La Tribune de l'Assurance*, n. 160.
- Planchet F., Théron P. [2006] « Modèles de Durée - Applications actuarielles », *Economica*.
- Planchet F., Théron P. [2007] *Pilotage technique d'un régime de rentes viagères*, Paris : Economica.
- Quashie A., Denuit M. [2005] « Modèles d'extrapolation de la mortalité aux grands âges », *Institut des Sciences Actuarielles et Institut de Statistique, Université Catholique de Louvain*, WP.

- Rakotomalala R. [2008] « Tests de normalité – Techniques empiriques et tests statistiques », *Université Lumière Lyon 2, Version 2.0*.
- Renshaw A., Haberman S. [2003] « Lee-Carter mortality forecasting with age-specific enhancement », *Insurance: Mathematics and Economics, Vol. 33*.
- Renshaw A., Haberman S. [2005] « Mortality Reduction Factors Incorporating Cohort Effects », *Actuarial Research Paper (Cass Business School), No. 160*.
- Saporta G. [2006] « Probabilités, analyses des données et statistique », *Technip, 2^{ème} édition*.
- Sow A.-M. L. [2007] « Ethnies et sociétés islamique en Afrique, un paradoxe ? – Le cas du Fuuta Dyalöö guinéen du XVI^e au XX^e siècles », *L'Harmattan*.
- Sylla Ndoye I. [1981] *Le développement de l'assurance vie dans un contexte islamique. Expérience de la société nationale d'assurance mutuelle vie*, Mémoire d'assurance, IIA.
- Tabeau E., Ekamper P., Huisman C., Bosch A. [2001a] « Predictiong Mortality from Period, Cohort or Cause-Specific Trends: a Study of Four European Countries », *extrait de* Tabeau E., Van Den Berg Jeths A., Heathcote C. (edited by) [2001b] « Forecasting Mortality in Developed Countries – Insights from a Statistical, Demographic and Epidemiological Perspective », *Kluwer*.
- Tarone R. E. [1981] « On Summary Estimators of Relative Risk », *Journal of Chronic Diseases, Vol. 34*.
- Thatcher A. R. [1999] « The long-term pattern of adult mortality and the highest attained age », *Journal of the Royal Statistical Society, 162*.
- Therneau T. M., Grambsch P. M. [2000] « Modeling Survival Data – Extending the Cox Model », *Springer*.
- Trani J.-F. [2007] « L'économie informelle est la voie pour un développement à l'africaine », *extrait de* Courade G. (ouvrage dirigé par) [2007] « L'Afrique des idées reçues », *Belin*.
- Van Den Berg Jeths A., Hoogenveen R., De Hollander G., Tabeau E. [2001] « A Review of Epidemiological Approaches to Forecasting Mortality and Morbidity », *extrait de* Tabeau E., Van Den Berg Jeths A., Heathcote C. (edited by) [2001b] « Forecasting Mortality in Developed Countries – Insights from a Statistical, Demographic and Epidemiological Perspective », *Kluwer*.
- Vaupel J. W. [2002] « Life Expectancy at Current Rates vs. Current Conditions: A Reflexion Stimulated by Bongaarts and Feeney's "How Long Do We Live?" », *Demographic Research, Vol. 7, Art. 8*.
- Vaupel J. W., Manton K. G. Stallard E. [1979] « The impact of heterogeneity in individual frailty on the dynamics of mortality », *Demography, Vol. 16, No. 3*.

- Viallon V. [2006] *Processus empiriques, estimation non paramétrique et données censurées*, Thèse de doctorat, Université Paris 6.
- Wilmoth J. R. [1995] « Are mortality projections always more pessimistic when disaggregated by cause of death? », *Mathematical Population Studies*, Vol. 5(4).
- Yeatman J. [2004] « L'assurance française en Afrique noire francophone », *Risques, Les cahiers de l'assurance* 57.
- Yeatman J. [2007] « L'assurance en Afrique : une émergence difficile », *Risques, Les cahiers de l'assurance* 71.
- Ziguélé M. [2008] « Comment renforcer les compagnies d'assurances africaines de la zone CIMA », *CAPafrique – Analyses sectorielles*.