Mémoire présenté le :

pour l'obtention du Diplôme Universitaire d'actuariat de l'ISFA et l'admission à l'Institut des Actuaires

Par : Yahia SALHI

Titre Tables de mortalité best estimate : une approche de crédibilité pour les portefeuilles de petite taille

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

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

Membre présents du jury de l'Institut des Actuaires	signature	*Entreprise :*
		Nom : ISFA
		Signature :
Membres présents du jury de l'ISFA		*Directeur de mémoire en entreprise :*
		Nom : Pierre THEROND
		Signature :
		Invité :
		Nom :
		Signature :
		***Autorisation de publication et de mise en ligne sur un site de diffusion de documents actuariels** (après expiration de l'éventuel délai de confidentialité)*
		Signature du responsable entreprise
		Signature du candidat

Mémoire d'Actuaire

Institut de Science Financière et d'Assurances

Par : Yahia SALHI

Titre : Tables de mortalité *best estimate* : une approche de crédibilité pour les portefeuilles de petite taille

---

Tables de mortalité *best estimate* : une approche de crédibilité pour les portefeuilles de petite taille

Yahia SALHI

2017

Résumé

L'avènement du nouveau régime prudentiel des assureurs européens a considérablement renforcé l'intérêt pour ceux-ci de disposer d'hypothèses de mortalité les plus pertinentes et proches de l'expérience possible. En effet, l'utilisation de tables de mortalité trop prudentes (par exemple celles imposées par la réglementation locale dans la tarification des produits d'assurance vie, en France) pour l'évaluation des provisions techniques Solvabilité II conduit à un double effet, à la fois sur les provisions techniques best estimate mais aussi sur la charge de capital du risque de mortalité.

Néanmoins un certain nombre d'acteurs se heurtent à des problématiques techniques liées à la taille de leurs portefeuilles et leur hétérogénéité en termes de garanties (pour un même risque). Par exemple, en termes de risque en cas de décès, un assureur peut disposer de quelques dizaines de milliers d'assurés bénéficiant de contrats de prévoyance décès, de garanties plancher sur des contrats en unités de comptes et de contrats d'assurance de prêt. Dans de tels contextes, il est délicat de construire des tables de mortalité au moyen de la seule expérience de chaque type de produit. D'autant plus que la révision dans le temps de celles-ci risque d'induire des impacts significatifs dans l'évolution des provisions techniques

Dans ce contexte ce mémoire propose une approche de crédibilité consistant à réviser, au fur et à mesure que de nouvelles observations arrivent, les paramètres d'un ajustement Makeham ou encore d'un modèle de lissage non-paramétrique basé sur la vraisemblance locale. Dans le premier cas, un tel ajustement permet en effet de rajouter de la structure ce qui s'avère utile lorsque les portefeuilles sont de taille limitée et le processus de révision proposé intègre la bonne représentation aux différents âges, ce qui est un point crucial compte tenu de la fréquente hétérogénéité du coût des prestations décès en fonction de l'âge. Dans le second cas, la méthode du lissage appliquée à des portefeuilles de petite taille ne parvient généralement pas à proposer une table d'expérience suffisamment lisse ni pertinente compte tenu du peu de décès observés. La méthode de crédibilité développée dans ce mémoire permet dans ce cas également de proposer une table d'expérience qui pondère l'information provenant des différents portefeuilles ou contrats en fonction de la pertinence des différentes informations sous-jacentes.

Crédibilité ; mortalité ; assurance vie ; construction de table ; Makeham ; lissage ; vraissemblance locale ; prediction

Abstract

Recently, there has been an increasing interest of life insurers to assess their portfolios own mortality risk. The new European prudential regulation, namely Solvency II, emphasized the need to use mortality and life tables that best capture and reflect the experienced mortality, and thus policyholders' proper risk profile, in order to adequately quantify the underlying risk. Therefore, building a mortality table based on the experience from the portfolio is highly recommended and, for this purpose, various approaches have been introduced in the literature. Although, such approaches succeed in capturing the main feature, it remains difficult to assess the mortality when the underlying portfolio lacks of sufficient exposure.

In this report, we propose to graduate the mortality curve using an adaptive procedure based on the local likelihood. The latter has the ability to model the mortality patterns even in presence of complex structures and avoid to rely on experts opinion. However, such a technique fails at proposing a consistent yet regular structure for portfolios with limited deaths. Although the technique borrows the information from the adjacent ages, it is sometimes not sufficient to produce a robust life tables. In the presence of such a bias, we propose to adjust the corresponding curve, at the age level, based on a credibility approach. This consists on reviewing, as new observations arrive, the assumption on the mortality curve.

We derive the updating procedure and investigate its benefits of using the latter instead of a sole graduation based on real datasets. Moreover, we look at the divergences in the mortality forecasts generated by the classical credibility approaches including Hardy-Panjer, the Poisson-Gamma model and Makeham framework on portfolios originating from various French insurance companies.

Credibility ; Mortality ; Life Insurance ; Graduation ; Makeham ; Smoothing ; Local Likelihood ; Prediction

Table des matières

Introduction

Motivation

L'avènement du nouveau régime prudentiel des assureurs européens a considérablement renforcé l'intérêt pour ceux-ci de disposer d'hypothèses de mortalité les plus pertinentes et proches de l'expérience possible. En effet, le référentiel Solvabilité 2 (2009) repose sur une approche économique de valorisation. Ainsi, un engagement dit *best estimate* "*correspond à la moyenne pondérée par leur probabilité des flux de trésorerie futurs compte tenu de la valeur temporelle de l'argent estimée sur la base de la courbe des taux sans risque pertinente, soit la valeur actuelle attendue des flux de trésorerie futurs*" (cf. l'article 77 de la directive Solvabilité 2 (2009) mais aussi l'article R351-2 du Code des assurances (2017)). Les hypothèses biométriques doivent donc représenter le plus fidèlement possible l'expérience propre à chaque contrat ou portefeuille. Celles-ci doivent bien évidemment être prudente selon la directive. Néanmoins, l'utilisation de tables de mortalité trop prudentes ; par exemple celles imposées par la réglementation locale dans la tarification des produits d'assurance vie, en France ; ou pour l'évaluation des provisions techniques Solvabilité 2 (2009) conduit à un double effet :

- Augmentation des provisions techniques *best estimate* (et donc diminution des fonds propres - *basic own-funds* - permettant de couvrir les exigences de capitaux),
- Augmentation de l'assiette servant au calcul de la charge de capital du risque de mortalité (scénario d'augmentation de 15% des taux conditionnels de décès dans la formule standard)

Par conséquent, la question de savoir quelle table de mortalité peut être considérée à des fins de tarification ou de provisionnement est d'une importance capitale. Une première tentative pour répondre à cette question consiste en l'utilisation des données disponibles au niveau du portefeuille et donc construire une table de mortalité spécifique. Pour cela, deux méthodologies de construction sont disponibles :

Approche paramétrique : Il s'agit de construire un indicateur de survie (ou de décès) sans aucune hypothèse sur la distribution des temps de survie (ou de décès). En pratique, cette approche est indispensable lorsque les connaissances sur le comportement ou la forme des lois d'intérêt ne sont pas précises. Il s'agit donc de l'estimation la plus naturelle des lois inhérentes du phénomène de décès. En assurance vie, entre autres, deux estimateurs sont souvent utilisés : l'estimateur de *Kaplan-Meier* introduit par Kaplan and Meier (1958), et celui de *Nelson-Aalen* considéré par Nelson (1972) et Aalen (1978). Ces deux estimateurs s'intéressent aux données incomplètes. Autrement dit, lorsque les survies (ou les décès) ne sont pas complètement accessibles dus au phénomène de censure car un individu peut être en vie à la fin de la date d'observation. Plus tard, nous caractérisons les données incomplète d'un point de vue plus formel.

Approche non-paramétrique : Il s'agit d'imposer une forme, avec des paramètres inconnus, à l'indicateur biométrique qui nous intéresse. Cela implique une connaissance parfaite de la lois sous-jacente au phénomène de survie (ou de décès). Depuis le début du 19$^{\text{ème}}$ siècle, la mortalité a fait l'objet de multiple études empiriques en démographie (humaine). Celles-ci tentaient, entre autres, de décrire l'évolution d'indicateur biométrique en fonction de l'âge (se référer par exemple aux travaux de Gompertz (1825), Makeham (1867, 1890), Weibull (1951) et Breslow et al. (1986)). Ces différents travaux académiques se sont donc attelés à expliquer le phénomène du vieillissement (sénescence) au travers une fonctionnel de l'âge atteint. Cette dernière s'avère avoir *une forme en "S"* caractérisant la mortalité adulte au delà d'un certain âge. Cette forme impose donc une croissance de la mortalité avec l'âge avec un changement de convexité. Pour les très grands âges plusieurs travaux académiques indiquent que la mortalité tend à avoir *un plateau*. En d'autres termes, l'augmentation de la mortalité décélère pour ces âges. Par conséquent, plusieurs formes paramétriques ont été introduites dans littérature démographique. Elles ont toutes un point commun qui est celui de reproduire cette évolution en "S" de la mortalité. On peut citer par exemple la loi de Gompertz (1825) et celle de Makeham (1867).

Approche semi-paramétrique : Il s'agit principalement du modèle de Cox (1972) dit à "*hasard proportionnel*" faisant intervenir des variables explicatives. Ce modèle ne requiert pas la formulation d'une hypothèse sur la forme de la mortalité, hypothèse qui est au coeur de l'approche paramétrique. Cependant, l'estimation des paramètres du modèle et tout particulièrement des coefficients des variables explicatives passe par la maximisation d'une fonction de vraisemblance dite partielle considérée par exemple dans le livre de Tibshirani and Hastie (1987). Le principe du modèle repose sur la décomposition de la mortalité comme le produit de deux éléments. Le premier dit fonction ou risque de base identique pour tous les assurés alors que le second dépend des caractéristiques des individus. Dans le cas où les individus sont seulement identifiés par leur âge et leurs sexe, la second composante est une fonctionnelle de l'âge permettant d'ajuster la mortalité localement en fonction de l'information disponible à cet même mais aussi celles des tranches d'âges adjacentes. Cette approche est dite de régression locale. L'aspect semi-paramétrique réside la conjugaison d'une estimation non-paramétrique avec une technique de maximisation de la vraisemblance partielle. Ce mécanisme a été retenu par exemple par le *groupe de travail "mortalité"* de l'*institut des actuaires*[*] ayant pour vocation de *fournir un ensemble de références et d'outils destinés à permettre aux organismes assureurs de calibrer leur risque de mortalité en respectant la logique best estimate des textes.*

La taille a de l'importance

Pour la mise en place d'une des approches citées supra, un certain nombre d'acteurs se heurtent à des problématiques techniques liées à la taille de leurs portefeuilles et leur hétérogénéité en termes de garanties (pour un même risque). Par exemple, en termes de risque en cas de décès, un assureur peut disposer de quelques dizaines de milliers d'assurés bénéficiant de contrats de prévoyance décès, de garanties plancher sur des contrats en unités de comptes et de contrats d'assurance de prêt. Dans de tels contextes, il est délicat de construire des tables de mortalité au moyen de la seule expérience de chaque type de produit. D'autant plus que la révision dans le temps de celles-ci risque d'induire des impacts significatifs dans l'évolution des provisions techniques. En effet, si l'on construit des tables d'expérience pour des groupes de contrat ou de portefeuilles de taille modeste on se heurte à un risque d'estimation fort important. A titre d'exemple, nous représentons dans la Figure 1 l'estimation des probabilités conditionnelles de décès en se basant sur une méthode simple d'ajustement paramétrique. Nous avons, en effet, considéré un portefeuille de 10000 assurés de sexe masculin d'âge entre 60 et 100 ans distribués selon la Figure 1 (graphique gauche). Nous avons supposé que la mortalité de ces assurés et proportionnelle à la mortalité d'une table de référence (réglementaire), en l'occurrence la table *TH00-02*. Le coefficient d'abattement étant le même pour tous les âges. Dans un second, nous avons estimés ce même coefficient et nous avons représenté l'incertitude sur l'estimation des probabilités de décès, cf. le graphique à gauche de la Figure 1. L'intérêt de cet exercice est de montrer que l'incertitude est inversement proportionnelle à l'exposition et donc que la difficulté d'apprécier la mortalité est d'autant plus importante que le portefeuille est petit ou l'exposition est relativement petite. Ce cas de figure pose un problème, car quand on s'intéresse à la mortalité par portefeuille le nombre d'assurés est peu élevé. D'un autre côté, l'exposition est aussi hétérogène en fonction de l'âge. Par exemple, pour le portefeuille considéré dans la Figure 1 les âges supérieurs à 80 ans souffrent d'une sous-représentation ce qui induit une forte incertitude des estimations. Il faut donc palier à ce problème en essayant d'emprunter une information supplémentaire des autres tranches d'âges mais aussi des autres portefeuilles avec une qualité mais aussi une quantité de données suffisamment importante.

Dans ce contexte ce mémoire propose une approche de crédibilité consistant à réviser, au fur et à mesure que de nouvelles observations arrivent, les paramètres d'un ajustement non-paramétrique de type *Makeham* ou encore d'un modèle de lissage semi-paramétrique basé sur la vraisemblance

[*]. Les conclusions et recommandation du groupe de travail sont disponibles sur ce lien http://www.ressources-actuarielles.net/gtmortalite

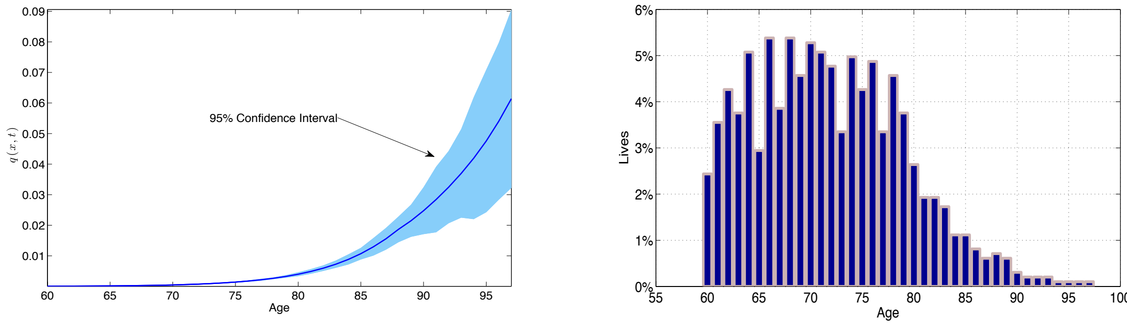

FIGURE 1 – Estimation des probabilités de décès (conditionnelles) par une méthode d'ajustement paramétrique

locale.

Approche paramétrique Dans ce cas, un tel ajustement permet en effet de rajouter de la structure ce qui s'avère utile lorsque les portefeuilles sont de taille limitée. En effet, un estimateur non-paramétrique de type Kaplan-Meier produit des estimations de mortalité ne respectant pas une cohérence dite *biologique*. Cela consiste en l'augmentation de la mortalité avec l'âge due au vieillissement. Il est d'usage dans ce cas de *lisser* ces mêmes taux de mortalité *a posteriori* pour leur donner la structure souhaitée. Par contre, la démultiplication des traitement que subissent les taux estimés augmente *de facto* l'incertitude des tables d'expérience et par conséquent l'adéquation aux observations. Pour cela, l'utilisation du modèle de Makeham permet de contourner cette seconde étape en proposant directement des taux lissés. Néanmoins, pour les portefeuilles de petite taille les estimateurs des différents paramètres modèle présentent une forte variance. Le processus de révision proposé intègre la bonne donc représentation aux différents âges, ce qui est un point crucial compte tenu de la fréquente hétérogénéité du coût des prestations décès en fonction de l'âge. Cette procédure de révision des paramètres intégrant la crédibilité des données provenant de la seule observation du portefeuille en question. En l'absence d'une information fiable et suffisamment *abondante* la procédure de révision emprunte alors de l'information aux autres portefeuilles permettant ainsi d'améliorer l'estimation des paramètres du modèle. Il est alors naturel d'accorder une importante crédibilité (proche de 1 sur une échelle de 0 à 1) à un portefeuille suffisamment grand et baser l'estimation des paramètres sur la seule information provenant de ce portefeuille. D'un autre côté, l'estimation des paramètres pour les petits portefeuilles requiert l'inclusion des estimations de ces même paramètres sur d'autres portefeuilles avec une pondération qui dépendant, bien évidemment, de la taille de ceux-ci.

Approche semi-paramétrique Dans le cas semi-paramétrique, nous considérons la méthode préconisée par le groupe de travail de l'*institut des actuaires*. Celle-ci propose sur un lissage local de la fonction de survie où la mortalité d'un portefeuille étant dépendante d'une mortalité de référence. L'aspect local repose sur l'adaptation de l'ajustement à chaque âge tout en empruntant l'information disponible aux âges adjacents. Plus précisément, la mortalité à un certain âge donnée la méthode du lissage appliquée à des portefeuilles de petite taille ne parvient généralement pas à proposer une table d'expérience suffisamment lisse ni pertinente compte tenu du peu de décès observés. La méthode de crédibilité développée dans ce mémoire permet dans ce cas également de proposer une table d'expérience qui pondère l'information provenant des différents portefeuilles ou contrats en fonction de la pertinence des différentes informations sous-jacentes.

Source de Données

Nous disposons de données de portefeuilles transmis par 14 assureurs Français pour la construction des tables *best estimate* dans le cadre du *groupe de travail "mortalité"* de l'*institut des actuaires*[*]. Cette base de donnée d'environ 8 millions de lignes a été utilisée dans plusieurs articles académiques (voir les travaux de Tomas and Planchet (2013, 2014)) et récemment dans le cadre d'un mémoire d'actuariat par Musset (2016). Se sont des données *par assuré* ayant un contrat individuel, collectif à adhésion obligatoire ou collectif à adhésion facultative. Pour chaque individu, nous avons plusieurs informations relatives à : sa date de naissance, son sexe, sa date d'entrée dans le portefeuille, sa date de sortie et son statut de sortie. Ce dernier représente une indicatrice de la survenance du décès ou non de l'individu durant sa présence dans la base. Il est alors possible de construire les différents indicateurs dont nous avons besoin pour mener à bien la construction d'une table d'expérience pour chaque portefeuille. En l'occurrence, nous serons en mesure de retracer *la vie d'un assuré dans le portefeuille*. Notamment, la duration d'un *séjour* dans le portefeuille, élément clé de la construction d'une table, sera quantifiée. La préparation de ces données pour la construction sera développée dans la section 1.2 du chapitre 1 mais aussi la section 2.2 du chapitre 2.

∗. Les conclusions et recommandation du groupe de travail sont disponibles sur ce lien http://www.ressources-actuarielles.net/gtmortalite

Chapitre 1

Approche paramétrique : Modèle de Makeham

1.1 Introduction

Therefore, the question of which mortality table can be considered for pricing and reserving purposes is of substantial importance. A first attempt, to handle this issue, is to use the available data at the portfolio level and build a specific mortality table. However, practitioners may face technical difficulties related to the size of the portfolio and the heterogeneity of the guarantees (for the same underlying risk). For instance, an insurer may detain a fairly big portfolio but with insured holding different policies : pure endowment contracts, unit-linked contracts with minimum death guarantees, loan insurance and so on. In such a case, it is difficult to build mortality tables only based on the sole experience of each policy. Especially since it may induce significant impacts on the technical reserves if the table has to be updated more frequently over time. In this chapter, we consider an insurer with exposures to different policies and aiming at establishing an experience-based mortality table for each policy.

In the academic literature, various methodologies have been proposed to built and graduate mortality rates at the insured portfolio level. They are usually divided into non-parametric and parametric techniques. The latter are very useful in practice especially when there is sufficient data, see Forfar et al. (1988a) for a comprehensive introduction to the use of parametric models for graduation. These approaches fit the parametric structure to the mortality of interest over a given period. The graduated mortality is then used to project future liabilities related to the underlying population. By doing so, the evolution of the flow of data related to latest available information is not taken into account. This should be, for example, used to update the graduated mortality. However, if one decides to re-calibrate the parametric model each year, the forecasts are likely to be unstable. This is mainly due to the instability of parameters estimation due to the lack of sufficient data.

In this context and following the work of Bühlmann and Gisler Bühlmann and Gisler (2005) and Hardy and Panjer Hardy and Panjer (1998), we propose a credibility approach which consists on reviewing, as new observations arrive, the parameters of a Makeham fit. The framework considered in Hardy and Panjer (1998) focuses on the update of the aggregate deaths recorded over the whole portfolio. However, such an approach may be not effective in situations where the insurer liability is highly dependent on the age structure of the underlying portfolio. Thus, using an adjustment makes possible to add a structure in the mortality pattern which is useful when portfolios are of limited size so as to ensure a good representation over the entire age-band considered. Note that, adding an age structure is also beneficial given the heterogeneity observed in the cost of the guarantees according to the age. To recap, as we can see in Section 1.5, the proposed adjustment approach is intended to enhance the predictive ability of the credibility-based revisions at the age-level and not on the aggregate portfolio level.

The remainder of the chapter is organized as follows. Section 1.2 has still an introductory purpose. It specifies the notation, assumptions and the Makeham settings used in the following. Section 1.3 introduces the Makeham credibility approach and assess the estimation of the credibility model. Section 1.4 describes the classical credibility approaches of mortality including Hardy and Panjer Hardy and Panjer (1998) and the Poisson-Gamma model. Section 1.5 presents an application with experience data originating from French insurance companies. Finally, some remarks in Section 1.6 conclude the chapter.

1.2 A Credibility Model for Makeham's Law

1.2.1 Data Structure and Notation

We suppose that we have at our disposal age-specific mortality statistics originating from n portfolios. For each portfolio $i \in \{1, \cdots, n\}$, we observe the deaths of exposures over a period T_i. Denote the number of individuals at attained age x during calendar year $t = 1, \cdots, T_i$ by $L^i_{x,t}$ and $D^i_{x,t}$ represents the number of deaths recorded. We also introduce the following notation,

$$D^i_{x,\bullet} = \sum_{t=1}^{T_i} D^i_{x,t}, \quad L^i_{x,\bullet} = \sum_{t=1}^{T_i} L^i_{x,t}, \quad \text{and} \quad D^i_{\bullet,t} = \sum_{x=\underline{x}}^{\overline{x}} D^i_{x,t}, \quad L^i_{\bullet,t} = \sum_{x=\underline{x}}^{\overline{x}} L^i_{x,t},$$

which refer respectively to the aggregate deaths and individuals over the age-band $\{\underline{x}, \underline{x}+1, \ldots, \overline{x}\}$ and calendar years 1 to T_i for each portfolio i. Henceforth, the "$\bullet$" indexation refers to the summation over the index of interest. For example, $D^\bullet_{x,\bullet}$ refers to the aggregate deaths over the period $[1, T_i]$ and over the n portfolios, i.e. $D^\bullet_{x,\bullet} = \sum_{i=1}^{n} \sum_{t=1}^{T_i} D^i_{x,t}$.

1.2.2 Mortality Law

We consider the (first) Makeham law of mortality, which generalizes the Gompertz law. Omitting the time dependency, Makeham Makeham (1867) assumes that the force of mortality φ_x^i at attained age x during calendar year t has the following form :

$$\varphi_x^i = A^i + B^i \times (C^i)^x, \tag{1.1}$$

with A^i, B^i and C^i are some constants. These parameters capture the essential properties of the progression of mortality. For instance, the dominant effect, i.e. the aging effect, over the age is captured by the multiplicative component factor $B^i \times (C^i)^x$. The non-age dependent parameter A^i can be interpreted as the non-senescent mortality, for instance, due to accidents. Both of these capture the exponential increase in the forces of mortality observed for adult mortality, see Bongaarts (2005) for more details.

Various modification of the above law have been proposed, especially, to encounter for the time dependency of the mortality, see e.g. Keyfitz (1981) among others. Indeed, as soon as age-specific mortality patterns over time are concerned, the time series records of the latter show a discernible downward trend with minor fluctuations around. In order to correct this deficiency in the model (1.1), we suppose that the time trend is incorporated in the parameter B^i denoted henceforth B_t^i. Therefore, the force of mortality $\varphi_{x,t}^i$ for portfolio i writes now as the following expression

$$\varphi_{x,t}^i = A^i + B_t^i (C^i)^x. \tag{1.2}$$

This model should capture the behavior of the probability of death over years through the time-dependent parameter B_t^i. This also make possible the prediction of future mortality, i.e. for $t = T+1, T+2, \ldots$, through the study of the time series B_t^i for $t = 1, \ldots, T$. When it comes to small portfolios, the model in (1.2) is not easy to implement. Indeed, as discussed later in this chapter, the temporal behavior the factor B^i cannot be accurately extracted. Nevertheless, one can use the estimated values of B_t^i even over the few periods to predict the future behavior of B^i.
Note that in order to estimate the parameters of the model 1.2, given the growth of the forces of mortality with the age, we must have a constant C greater than 1 and a positive B. Then,

$$\begin{aligned} q_{x,t}^i &= 1 - \exp\left(-\int_x^{x+1} \varphi_{y,t}^i \, dy\right) = 1 - \exp\left(-\int_x^{x+1} A^i + B_t^i \times (C^i)^y \, dy\right) \\ &= 1 - \exp(-A^i) \exp\left(-\frac{B_t^i}{\ln C^i} (C^i)^x (C^i - 1)\right), \end{aligned} \tag{1.3}$$

where $q_{x,t}^i$ denotes the one-year probability of death at attained age x during calendar year t for portfolio i. Consistent estimates $\widehat{A^i}, \widehat{B_t^i}$ and $\widehat{C^i}$ of the parameters are obtained by minimizing the following weighted distance :

$$\sum_{x=\underline{x}}^{\overline{x}} \frac{L_{x,t}^i}{q_{x,t}^i (1 - q_{x,t}^i)} \, (q_{x,t}^i - \widehat{q}_{x,t}^i)^2,$$

with $\widehat{q}_{x,t}^i = D_{x,t}^i / L_{x,t}^i$ is the crude mortality rates.

1.2.3 Differential Mortality Law

It is common in modeling specific portfolio's mortality to consider an adjustment with regard to a baseline mortality. Generally, this implicitly assumes that both populations share common features up to a random effect. Relational models stipulate a deterministic relationship in the form $q_x^i = f(q_x^{\mathrm{b}})$ links the two mortalities, where q_x^{b} refers to the *baseline mortality*. The function $f : [0, 1] \to [0, 1]$ is a known and deterministic function, see Delwarde et al. (2004) for more details. A simple example would suggest that the death rate is common for all companies. Specifically, $q_x^i = q_x^{\mathrm{b}}$ for any $i \in \{1, \cdots, n\}$. However, such an assumption does not appreciate the specific characteristic of each portfolio's mortality profile. In other words, portfolios having lives in poorer or better conditions than the baseline mortality do not behave in a similar fashion than the baseline mortality. This implies that one should encounter for differential mortality that arises due to portfolio specific features, e.g. particular socioeconomic groups involved, average income level, etc. However, when it comes to the study of the mortality at a single portfolio level, some specific issues arise :
(i) Size of populations : Insured population are generally of small size, so none or very few deaths are observable at some ages. This may not only bias the estimation of the force of mortality but also lead to a mis-estimation of the parameters in (1.3). This may cause high fluctuations for $q_{x,t}^i$ and consequently for A^i, B_t^i and C^i.
(ii) Length of historical data : Available age-specific mortality statistics lacks of deepness. This makes difficult to isolate a possible time trend as it may be captured by B_t^i. The latter may be fluctuating due to the small size of the dataset as noted before.
(iii) Scale of available data : Insured portfolios show a typical behavior compared to a national mortality. The mortality of insured population is significantly lower than the national population from which it is drawn. This could make the use of a baseline mortality based on national demographic statistics as a substitute useless as it may not have the same characteristics of the initial population.

All these characteristics make forecasting of future mortality evolution problematic. In order to overcome these issues when implementing and fitting the model (1.3) for each portfolio $i \in \{1, \ldots, n\}$ we will make the following assumptions :

(i) The baseline mortality $q_{x,t}^{\mathrm{b}}$ is described by the Makeham model in (1.3).
(ii) The age effect is similar on the n portfolios and companies specific model is assumed to share the same parameters A^i and C^i. Those are set equal to the baseline ones, i.e. $A^i = A^{\mathrm{b}}$ and $C^i = C^{\mathrm{b}}$ for any $i \in \{1, \cdots, n\}$.
(iii) The time-dependent parameter B_t^i is fitted at each period. This is given by the following formula :

$$B_t^i = \frac{D_{\bullet,t}^i - A^{\mathrm{b}} L_{\bullet,t}^i}{\sum_{x=\underline{x}}^{\overline{x}} (C^{\mathrm{b}})^x L_{x,t}^i}.$$

The assumptions (i) and (ii) allows to overcome potential estimation bias of the parameters A^i and C^i. Indeed, basing the estimation on a large population allows to avoid erroneous inferences of the parameters. Also, if the portfolio i is a subset of the baseline population composed of the aggregated portfolios, we may think that both the non-senescent factor A^i and the slope C^i are equivalent and thus normalized with the baseline mortality. Empirical evidence of a normalized slope can be found in Thatcher (1999). It is shown that relative rate of increase is the same at all ages and is a shared feature with over subset populations, see also the empirical study of Zhu and Li (2013). As for the non-senescent parameter, the assumption is relevant to the extent that this effect is generally of small impact and sometimes ignored (especially for industrialized countries), see Gavrilova and Gavrilov (2011). The unique parameter that captures the specific mortality at the portfolio level is B_t^i, which would *a priori* not be the same over companies due to the heterogeneity of the underlying populations as explained above. This can be regarded as an unobservable random factor and similar to the so-called *frailty factor*. Such a methodology is widely understood in the literature as well as in life insurance practice. Assumption (iii) gives an estimate of the time-dependent parameter. By time-dependent we only track the fluctuation of B_t^i over time that might be caused by the small size and length of data. Thus our main aim is to *sequentially* adjust the estimation of B_t^i over time in view of the flow of information at our disposal.

1.3 Credibility of the Makeham Mortality

1.3.1 Next Period Prediction

In the following, we are interested in the behavior, over time, of the random variable

$$X_t^i = \frac{B_t^i}{B_t^{\mathrm{b}}}, \tag{1.4}$$

and specifically on its next period prediction X_{T+1}^i merging information from other portfolios $j = 1, \cdots, n$ with $j \neq i$. Specifically, suppose that we are at the end of the year T, i.e. at time $T+1$, and we want to predict the next period deaths $D_{x,T+1}^i$ in the portfolio (equivalently the probability of death $q_{x,T+1}^i$). Naturally, we can assume that this ratio is constant over time and thus invoke a widespread practice that applies a single factor of reduction/increase to the baseline mortality. On the other hand, one could propose a dynamic model on the same line as Plat (2009). The latter proposes a modeling framework of the relative ratio of an experienced mortality (death rates) to a baseline and consider that this can be diffused using either an autoregressive model or a decomposition similar to the one introduced by Lee and Carter Lee and Carter (1992). Other methodologies have been also proposed, see Ngai and Sherris (2011) and Hyndman et al. (2013) among others. However, random effects that constitute the decomposition of the experienced mortality have to be projected using their temporal and statistical features. In our case, we are not only interested in handling populations of small size but also with potentially limited historic period

of observation. Therefore, such a methodology would typically not be useful in our setting as it requires a long experience.
Note that the behavior of X_t^i is broadly related to the so-called basis risk. This refers to the fact that the evolution of the policyholders mortality is usually different from that of the national population (baseline), due to some selection effects. This selection effect has different impacts on different insurance companies portfolios, as mortality improvements and accelerations are very heterogeneous in the insurance industry, see Barrieu et al. (2012).

1.3.2 Heterogeneity and Makeham's Law Adjustment

As noted above, we are interested in the accurate adjustment of the portfolio-dependent parameter in (1.3), i.e. B_t^i. Given the specific parameterization of the problem, one may think of the n portfolios as a subset of the reference population and thus each population is characterized by a risk profile Θ_i. In addition, it is beneficial to borrowing information across the different portfolios to enhance the knowledge and estimation of the mortality at the single portfolio level. Furthermore, these subpopulations may, for example, share a common mortality feature, while showing some specificity in their mortality profile. This can be seen as a random variable effect or heterogeneity characterizing the specific profile of each portfolio, for $i = 1, \cdots, n$. Therefore, we implicitly assume that each portfolio is endowed by a risk profile θ_i which is a realization of a random variable Θ.

In view of the various stylized facts presented above and in order to predict $D_{x,T+1}^i$, for each age x, we focus on the projection of X_{T+1}^i. Therefore, we suppose that this relative trend level of portfolio i with respect to the baseline mortality (trend) is characterized by the risk profile θ_i which is a realization of Θ_i. In other words, X_t^i is viewed as a function of a random element Θ_i representing the unobserved characteristics of the portfolio mortality trend (with respect to the baseline). By doing so, we implicitly take into account the heterogeneity of the portfolio i's portfolio mortality profile. It thus remains to predict X_{T+1}^i taking into account this random heterogeneity. By doing so, we naturally invoke the use of a credibility approach to estimate X_{T+1}^i.

1.3.3 Credibility Based Adjustment

As noted above, the objective is to estimate the next period projection of the relative ratio for each portfolio i. More precisely, in view the available data up to time T_i, i.e. X_t^i, for $t = 1, \cdots, T_i$, one aims to find the best estimate of $\mathbb{E}[X_{T_i+1}^i|\Theta_i] = \mu(\Theta_i)$, which is unknown. Let $\widehat{\mu}(\Theta_i)$ be this estimation. For this purpose and using the usual credibility setting, we shall make the following hypotheses :

(H1) Conditionally on Θ_i, the random variables X_t^i, for $t \in \{1, \ldots, T_i\}$, are independent with mean and variance given as follows

$$\mathbb{E}\big[X_t^i \,|\, \Theta_i\big] = \mu\big(\Theta_i\big) \quad \text{and} \quad \mathbb{V}\text{ar}\big[X_t^i \,|\, \Theta_i\big] = \frac{\sigma^2\big(\Theta_i\big)}{\omega_t^i} \quad ,$$

for some functions $\mu\big(\Theta_i\big)$ and $\sigma^2\big(\Theta_i\big)$ and where

$$\omega_t^i = \frac{\sum_{x=\underline{x}}^{\overline{x}} (C^{\mathrm{b}})^x \, L_{x,t}^i}{\sum_{i=1}^n \sum_{x=\underline{x}}^{\overline{x}} (C^{\mathrm{b}})^x \, L_{x,t}^i},$$

measures the *weight* given to the period t experience from the portfolio i.

(H2) The pairs $\big(\Theta_i, X_t^i\big)$, $\big(\Theta_k, X_l^k\big)$, $k \neq i$ are independent and identically distributed.

The first assumption (H1) implies that for each risk profile i (portfolio), the *true* relative ratio $\mu\big(\Theta_i\big)$ (conditionally on the knowledge of the risk profile Θ_i) does not change over time and its variance given Θ_i, $\mathbb{V}\text{ar}\big[X_t^i \,|\, \Theta_i\big]$ changes in proportion to the relative size of the portfolio ω_t^i. The latter expresses different concerns outlined earlier. Specifically, it links the variability of the estimation of the parameter B_t^i to the size of the underlying population : very small portfolios are subject to larger variability on the estimation of B_t^i and vice versa.

The second assumption (H2) means that the risk profiles are independent. The successive realizations of the relative ratio X_t^i for any portfolio are independent of each other except through the risk parameter Θ_i. Moreover, using the random variable X_t^i instead of B_t^i permits to avoid data adjustment.

Intuitively, assumption (H2) implicitly suggests that portfolios are comparable as they are random sub-groups of a reference (national) population, but not entirely similar which induces the conditional independence.
In view of these assumptions, the following results are straightforward :
(i) The expected prediction of X_{T+1}^i unconditionally on the risk profile Θ_i is given by $\mathbb{E}\big[X_{T+1}^i\big] = \mathbb{E}\big[\widehat{X}_{T+1}(\Theta_i)\big] = 1$. In other words, in the absence of any information on the heterogeneity level on the parameter B_t^i, the best next-period prediction of the latter is the reference one, i.e. $\mathbb{E}\big[B_t^i\big] = B_t^{\mathrm{b}}$.
(ii) Using the law of total variance, the dependence structure of portfolio i associated risk factor over time, is, for $l, t \in \{1, \dots, T\}$,

$$\begin{aligned}
\mathbb{C}\text{ov}\big(X_l^i, X_t^i\big) &= \mathbb{C}\text{ov}\big(\mathbb{E}\big[X_l^i|\Theta_i\big], \mathbb{E}\big[X_t^i|\Theta_i\big]\big) + \mathbb{E}\big[\mathbb{C}\text{ov}\big(X_l^i, X_t^i|\Theta_i\big)\big] \\
&= \mathbb{V}\text{ar}\big[\mu\big(\Theta_i\big)\big] + \mathbb{E}\big[\mathbb{C}\text{ov}\big(X_l^i, X_t^i|\Theta_i\big)\big] \\
&= \begin{cases} \tau^2 & \text{if } l \neq t \\ \tau^2 + \dfrac{\sigma^2}{\omega_t^i} & \text{if } l = t, \end{cases}
\end{aligned} \tag{1.5}$$

where $\mathbb{V}\text{ar}\big[\mu\big(\Theta_i\big)\big] = \mathbb{V}\text{ar}\big[\Theta_i\big] := \tau^2$, while $\mathbb{E}\big[\sigma^2\big(\Theta_i\big)\big] = \mathbb{E}\big[\Theta_i\big] := \sigma^2$.

1.3.4 Credibility Estimator

Following the Bühlmann-Straub credibility approach, the aim is to find the *best estimate* of the actual to expected mortality ratio $\mathbb{E}\big[X^i_{T_i+1}\,|\,\Theta_i\big] = \mu\big(\Theta_i\big)$ which is linear in the observations. For each portfolio, due to the assumption (H2), $\widehat{\mu}(\Theta_i)$ depends only on the observations and the linear credibility estimator is of the form

$$\widehat{\mu}\big(\Theta_i\big) = \widehat{a}^i_0 + \sum_{t=1}^{T_i} \widehat{a}^i_t\, X^i_t\ , \tag{1.6}$$

where the coefficients $\widehat{a}^i_t$, for $t = 0, \cdots, T_i$, are those minimizing the mean squared errors criterion

$$\big(\widehat{a}^i_t\big)_{t=0,\cdots,T_i} = \underset{(a^i_t)_{(t=0,\cdots,T_i)}}{\operatorname{argmin}} \Big\{\mathbb{E}\Big[\big(\widehat{\mu}\big(\Theta_i\big) - a^i_0 - \sum_{t=1}^{T_i} a^i_t X^i_t\big)^2\Big]\Big\}.$$

In view of (1.5), taking the derivatives of the above criterion with respect to the $a_{i,t}$'s and equating to zero gives,

$$\widehat{a}^i_0 = 1 - \frac{\tau^2\,\omega^i_\bullet}{\sigma^2 + \tau^2\,\omega^i_\bullet} \quad \text{and} \quad \widehat{a}^i_t = \frac{\tau^2\,\omega^i_t}{\sigma^2 + \tau^2\,\omega^i_\bullet}, \quad \text{with} \quad \omega^i_\bullet = \sum_{t=1}^{T_i} \omega^i_t\ . \tag{1.7}$$

Then, substituting (1.7) into (2.18), leads to the following the Bühlmann-Straub credibility estimator of $X^i_{T_i+1}$

$$\widehat{X}^i_{T_i+1}(\Theta_i) = \alpha^i\, X^i_\bullet + (1 - \alpha^i), \quad \text{with } \alpha^i = \omega^i_\bullet \tau^2/(\omega^i_\bullet \tau^2 + \sigma^2), \tag{1.8}$$

where $X^i_\bullet = (\sum_{t=1}^{T_i} \omega^i_t X^i_t)/\omega^i_\bullet$. Note that the ratio σ^2/τ^2 represents the credibility coefficient. The parameter α_i is called the *credibility factor* or *credibility weight* for portfolio i and takes values in $[0, 1]$. For each portfolio i, note that the larger the volume of historical data, the larger α_i will be, see Equation (1.8).

1.3.5 Estimators of the Structure Parameters

As the risk parameters, Θ_i, for $i \in \{1, \ldots, n\}$, are assumed to be identically distributed, their moments are identical. Therefore τ^2 and σ^2 are the same for all portfolios and measure the residual heterogeneity of the risk profiles and the pure randomness respectively. These parameters are the key determinants of the credibility estimator, i.e. Equation (1.8). In the following, special attention is addressed to the estimation of these quantities. Recall the definition of the structure parameters,

$$\sigma^2 = \mathbb{E}\big[\sigma^2\big(\Theta_i\big)\big] = \omega^i_t\, \mathbb{E}\big[\mathbb{V}\mathrm{ar}\big[X^i_t|\Theta_i\big]\big], \quad \text{and} \quad \tau^2 = \mathbb{V}\mathrm{ar}\big[\mathbb{E}[X^i_t|\Theta_i]\big].$$

Then, it is reasonable to propose the estimators $\widehat{\sigma}^2$ and $\widehat{\widehat{\tau}}^2$ in the same vein as Bühlmann and Gisler (2005) based on the observations X_t^i :

$$\widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} s_i^2, \quad \text{with} \quad s_i^2 = \frac{1}{T_i - 1}\sum_{t=1}^{T_i} \omega_t^i \big(X_t^i - X_\bullet^i\big)^2, \tag{1.9}$$

$$\text{and} \quad \widehat{\widehat{\tau}}^2 = \frac{\omega_\bullet^\bullet}{(\omega_\bullet^\bullet)^2 - \sum_{i=1}^{n}(\omega_\bullet^i)^2}\left\{\sum_{i=1}^{n}\omega_\bullet^i\big(X_\bullet^i - X_\bullet^\bullet\big)^2 - (n-1)\widehat{\sigma}^2\right\},$$

$$\text{with} \quad X_\bullet^\bullet = \frac{1}{\omega_\bullet^\bullet}\sum_{i=1}^{n}\omega_\bullet^i X_\bullet^i \quad \text{and} \quad \omega_\bullet^\bullet = \sum_{i=1}^{n}\omega_\bullet^i.$$

These estimators are unbiased and consistent, see Bühlmann and Gisler (2005) for more details. Note that $\widehat{\widehat{\tau}}^2$ can be negative. This would mean that there would be no difference between the risks. In this case, $\widehat{\tau}^2$ is set to 0. Hence we use as estimator $\widehat{\tau}^2 = \max\big(\widehat{\widehat{\tau}}^2, 0\big)$.

1.3.6 Empirical Credibility Estimator

The empirical credibility estimator is obtained from the credibility formula (2.18) by replacing the structural parameters σ^2 and τ^2 by their estimators derived in Subsection 1.3.5. Hence, we have

$$\begin{cases} \widehat{\widehat{X}}^i_{T_i+1} & = \widehat{\alpha}^i X_\bullet^i + (1 - \widehat{\alpha}^i), \\ \widehat{\alpha}^i & = \dfrac{\widehat{\tau}^2 \omega_\bullet^i}{\widehat{\sigma}^2 + \widehat{\tau}^2 \omega_\bullet^i}. \end{cases} \tag{1.10}$$

It follows from Equation (1.8), that the mortality time varying coefficient is successively updated as follows

$$\widehat{B}^i_{T_i+1} = \widehat{B}^{\mathrm{b}}_{T+1}\big(1 + \widehat{\alpha}^i\,(X_\bullet^i - 1)\big), \tag{1.11}$$

and similarly, the forces of mortality and the probabilities of death are given respectively by

$$\widehat{\varphi}^i_{x,T_i+1} = \Big\{\widehat{\alpha}^i\big(1 - X_\bullet^i\big)\Big\}\widehat{A}^{\mathrm{b}} + \Big\{\widehat{\alpha}^i\big(X_\bullet^i - 1\big) + 1\Big\}\widehat{\varphi}^{\mathrm{b}}_{x,T+1},$$

$$\text{and} \quad \widehat{q}^i_{x,T_i+1} = \widehat{q}^{\mathrm{b}}_{x,T+1}\left(\frac{1 - \widehat{q}^{\mathrm{b}}_{x,T+1}}{\exp(-\widehat{A}^{\mathrm{b}})}\right)^{\dfrac{1}{\widehat{\alpha}^i(X_\bullet^i - 1)}}. \tag{1.12}$$

1.4 Classical Credibility Approaches to Mortality

Next, we wish to compare our model to the Hardy and Panjer Hardy and Panjer (1998) and Poisson-Gamma credibility analysis to mortality. The actual to expected mortality ratio is the key observation that is the focus of the two following approaches. Specifically, the *a priori* expected number of deaths for portfolio i in calendar year t in the age-band $\left[\,\underline{x}\,,\overline{x}\,\right]$ is denoted by

$$\omega_t^i = \mathbb{E}\big[D^i_{\bullet,t}\big] = \sum_{x=\underline{x}}^{\overline{x}} q^{\mathrm{b}}_{x,t}\, L^i_{x,t}\,.$$

The actual to expected mortality ratios denoted by X_t^i are computed for each calendar year t in aggregate for each portfolio i,

$$X_t^i = \frac{D^i_{\bullet,t}}{\mathbb{E}\big[D^i_{\bullet,t}\big]} = \frac{D^i_{\bullet,t}}{\omega_t^i}\,.$$

Both the Hardy and Panjer Hardy and Panjer (1998) and the Poisson-Gamma credibility approaches are using the Bühlmann-Straub set-up, see Section 1.3.3. Again, the key determinants of the credibility estimator (1.8) are the structure parameters, i.e. the variance part of the credibility premium $\mathbb{E}\big[\sigma^2\big(\Theta_i\big)\big]$ denoted by σ^2 and the fluctuation part $\mathbb{V}\mathrm{ar}\big[\mu\big(\Theta_i\big)\big]$ denoted by τ^2.

1.4.1 The Hardy-Panjer Approach

As in general we have no knowledge or, at least, no exact knowledge of the parametric distributions for the number of deaths or of the structure distribution, we need estimators for the two components of the credibility estimator (1.8), i.e. estimators for τ^2 and σ^2. The Hardy and Panjer Hardy and Panjer (1998) credibility approach to mortality estimates the structure parameters from the aggregated data using the estimators derived in Lourdes (1989). They estimate $\mathbb{E}\big[\sigma^2\big(\Theta_i\big)\big]$ using the following estimator, denoted by $\widehat{\sigma}_0^2$:

$$\widehat{\sigma}_0^2 = \frac{1}{\sum_{i=1}^n C_i}\sum_{i=1}^n C_i\, s_i^2, \quad \text{where} \quad C_i = \frac{1}{1+\frac{2}{T_i-1}\,\phi} \quad \text{with} \quad \phi = \frac{\mathbb{E}\big[\sigma^4\big(\Theta_i\big)\big]}{\mathbb{V}\mathrm{ar}\big[\sigma^2\big(\Theta_i\big)\big]}.$$

Again, as the risk parameters, $\big\{\Theta_i\big\}_{i=1}^n$, are assumed to be identically distributed, the factors $\mathbb{E}\big[\sigma^4\big(\Theta_i\big)\big]$ and $\mathbb{V}\mathrm{ar}\big[\sigma^2\big(\Theta_i\big)\big]$ are independent of the portfolio. Hence, the only portfolio dependent variable in C_i is T_i, the number of years data available for the portfolio.

Both methods give the same result. In addition, the latter approach, derived from Lourdes (1989), allows to obtain a credibility estimator for the variance part of the credibility premium which has the form : $\widetilde{\sigma_i}^2 = C_i\, s_i^2 + (1-C_i)\,\widehat{\sigma}_0^2$.

The estimate of $\mathbb{V}\mathrm{ar}\big[\mu\big(\Theta_i\big)\big]$ denoted by $\widehat{\tau}^2$ is

$$\widehat{\tau}^2 = \frac{\omega_{\bullet}^{\bullet}\, W - \widehat{\sigma}^2}{\omega_{\bullet}^{\bullet}\, \Omega}\,, \quad \text{where} \ \Omega = \frac{1}{\big(\sum_{i=1}^n T_i\big)-1}\sum_{i=1}^n \frac{\omega_{\bullet}^i}{\omega_{\bullet}^{\bullet}}\left(1-\frac{\omega_{\bullet}^i}{\omega_{\bullet}^{\bullet}}\right),$$

$$\text{and} \quad W = \frac{1}{\big(\sum_{i=1}^n T_i\big)-1}\sum_{i=1}^n\sum_{t=1}^{T_i}\frac{\omega_t^i}{\omega_{\bullet}^{\bullet}}\big(X_t^i - X_{\bullet}^{\bullet}\big)^2,$$

$$\text{with} \quad \omega_{\bullet}^{\bullet} = \sum_{i=1}^n \omega_{\bullet}^i\,, \quad \text{and} \quad X_{\bullet}^{\bullet} = \frac{1}{\omega_{\bullet}^{\bullet}}\sum_{i=1}^n \omega_{\bullet}^i X_{\bullet}^i\,.$$

Then, the estimate of $\mathbb{E}\big[\sigma^4(\Theta_i)\big]$ denoted by $\widehat{\sigma}^4$ is

$$\widehat{\sigma}^4 = \frac{1}{\sum_{i=1}^n (T_i+1)} \sum_{i=1}^n (T_i-1)\big(s_i^2\big)^2 ,$$

and the estimate of $\mathbb{V}\mathrm{ar}\big[\sigma^2(\Theta_i)\big]$ denoted by $\widehat{\upsilon}_{\sigma^2}$ is

$$\widehat{\upsilon}_{\sigma^2} = \frac{1}{R}\left(\sum_{i=1}^n (T_i-1)\big(s_i^2-\beta^2\big)^2 - 2\,\widehat{\sigma}^4(n-1)\right),$$

$$\text{where}\quad R = \sum_{i=1}^n (T_i-1) - \frac{\sum_{i=1}^n (T_i-1)^2}{\sum_{i=1}^n (T_i-1)}\,,\ \text{and}\quad \beta^2 = \frac{1}{n}\sum_{i=1}^n s_i^2 .$$

1.4.2 The Poisson-Gamma Approach

A priori, we could assume that $\mathbb{E}\big[\Theta_i\big]=1$ so that the baseline mortality produces the a priori expected number of deaths,

$$\mathbb{E}\big[D^i_{\bullet,t}\big] = \mathbb{E}\big[\omega^i_t\,\Theta_i\big] = \omega^i_t .$$

We suppose here that the parametric distribution for the number of deaths $D^i_{\bullet,t}$ is Poisson conditional to the relative risk level Θ_i, so that

$$\mathbb{E}\big[D^i_{\bullet,t}\,|\Theta_i\big] = \mathbb{V}\big[D^i_{\bullet,t}\,|\Theta_i\big] = \omega^i_t\,\Theta_i .$$

Then, under assumption (H1) in Section 1.3.3, the conditional mean and variance of the actual to expected mortality ratios become :

$$\mathbb{E}\big[X^i_t\,|\,\Theta_i\big] = \mu(\Theta_i) = \Theta_i \quad\text{and}\quad \mathbb{V}\mathrm{ar}\big[X^i_t\,|\,\Theta_i\big] = \frac{\sigma^2(\Theta_i)}{\omega^i_t} = \frac{\Theta_i}{\omega^i_t}\,,$$

and the p.d.e with respect to $a_{i,0}$ and $a_{i,t}$, (see Equation (1.7)) are :

$$a_{i,0} = 1 - \frac{\tau^2\,\omega^i_\bullet}{1+\tau^2\,\omega^i_\bullet} \quad\text{and}\quad a_{i,t} = \frac{\tau^2\,\omega^i_t}{1+\tau^2\,\omega^i_\bullet}\,,\quad\text{since}\quad \sigma^2 = \mathbb{E}\big[\Theta_i\big]=1.$$

Then the linear credibility estimator is given by

$$\widehat{\mu}(\Theta_i) = \widehat{X}^i_{T_i+1} = \frac{1}{1+\tau^2\,\omega^i_\bullet} + \frac{\tau^2\,\omega^i_\bullet}{1+\tau^2\,\omega^i_\bullet}\,\frac{1}{\omega^i_\bullet}\sum_{t=1}^{T_i}\omega_{i,t}\,X_{i,t} . \tag{1.13}$$

And, the expected number of deaths for portfolio i for next year T_i+1 is

$$\omega^i_{T_i+1}\,\widehat{X}^i_{T_i+1} = \omega^i_{T_i+1}\frac{1+\tau^2\,D^i_{\bullet,\bullet}}{1+\tau^2\,\omega^i_\bullet} .$$

Then, we need to obtain the structure parameter $\tau^2 = \mathbb{V}\mathrm{ar}\big[\Theta_i\big]$. As the distribution of the total number of deaths in portfolio i is $D^i_{\bullet,\bullet}\sim\mathcal{MP}\big(\omega^i_\bullet\Theta_i\big)$ (mixed Poisson distribution) and using the variance decomposition principle,

$$\begin{aligned}\mathbb{V}\mathrm{ar}\big[D^i_{\bullet,\bullet}\big] &= \mathbb{V}\mathrm{ar}\big[\mathbb{E}\big[D^i_{\bullet,\bullet}\,|\,\Theta_i\big]\big] + \mathbb{E}\big[\mathbb{V}\mathrm{ar}\big[D^i_{\bullet,\bullet}\,|\,\Theta_i\big]\big] = \mathbb{V}\mathrm{ar}\big[\omega^i_\bullet\Theta_i\big] + \mathbb{E}\big[\omega^i_\bullet\Theta_i\big]\\ &= \tau^2\,(\omega^i_\bullet)^2 + \omega^i_\bullet .\end{aligned}$$

Noticing that $\sum_{i=1}^{n} \mathbb{V}\text{ar}\big[D^i_{\bullet,\bullet}\big] = \omega^2 \sum_{i=1}^{n}(\omega^i_{\bullet})^2 + \sum_{i=1}^{n}\omega^i_{\bullet}$, leads to $\tau^2 = \sum_{i=1}^{n}\big(\mathbb{V}\text{ar}\big[D^i_{\bullet}\big] - \omega^i_{\bullet}\big) / \sum_{i=1}^{n}(\delta^i_{\bullet})^2$. Thus, the estimator of τ^2 writes

$$\widehat{\tau}^2 = \frac{\sum_{i=1}^{n}\left(\left(D^i_{\bullet,\bullet} - \omega^i_{\bullet}\right)^2 - D^i_{\bullet,\bullet}\right)}{\sum_{i=1}^{n}(\omega^i_{\bullet})^2}.$$

1.5 Numerical Analysis

1.5.1 Data Quantitative Analysis

The data come from studies conducted by *Institut des Actuaires*. These studies include in total 14 portfolio covering the period 2007-2011 with each companies contributing data for at least 4 of a possible 5 years. Table 1.1 presents the observed characteristics of the male population of the portfolios. For this dataset, we are considering respectively $T_i = 3$ and $T_i = 4$ for all companies. The

Table 1.1 – Observed characteristics of portfolios population.

	Period of observation		**Mean age**		**Average exposure**	**Mean age at death**
	Beginning	**End**	**In**	**Out**		
1	1/1/07	12/31/11	36.96	39.74	2.77	68.78
2	1/1/07	12/31/11	69.3	73.35	4.05	80.34
3	1/1/07	12/31/10	40.16	43.1	2.94	71.77
4	1/1/07	12/31/11	37.5	41.13	3.63	54.08
5	1/1/07	12/31/11	36.9	39.1	2.2	59.31
6	1/1/07	12/31/10	48.5	52.11	3.62	82.34
7	1/1/07	12/31/11	66.65	71.29	4.64	73.68
8	1/1/07	4/13/11	67.51	71.38	3.86	80.72
9	1/1/07	6/30/11	45.97	49.6	3.62	73.17
10	1/1/07	12/31/11	62.97	67.64	4.67	79.77
11	1/1/07	12/31/11	38.89	42	3.11	56.44
12	1/1/07	12/31/11	37.05	39.2	2.15	57.41
13	1/1/07	12/31/11	43.01	46.89	3.88	71.03
14	1/1/07	12/31/11	50.12	54.16	4.04	72.37

remaining years serve to test the predictive feature of the model through an in-sample analysis. The age band for all companies ranges from 30 to 95 years old. Figure 1.1 shows the age distribution of two portfolios. It graphically depicts the heterogeneity observed between the portfolios with insured holding different policies.

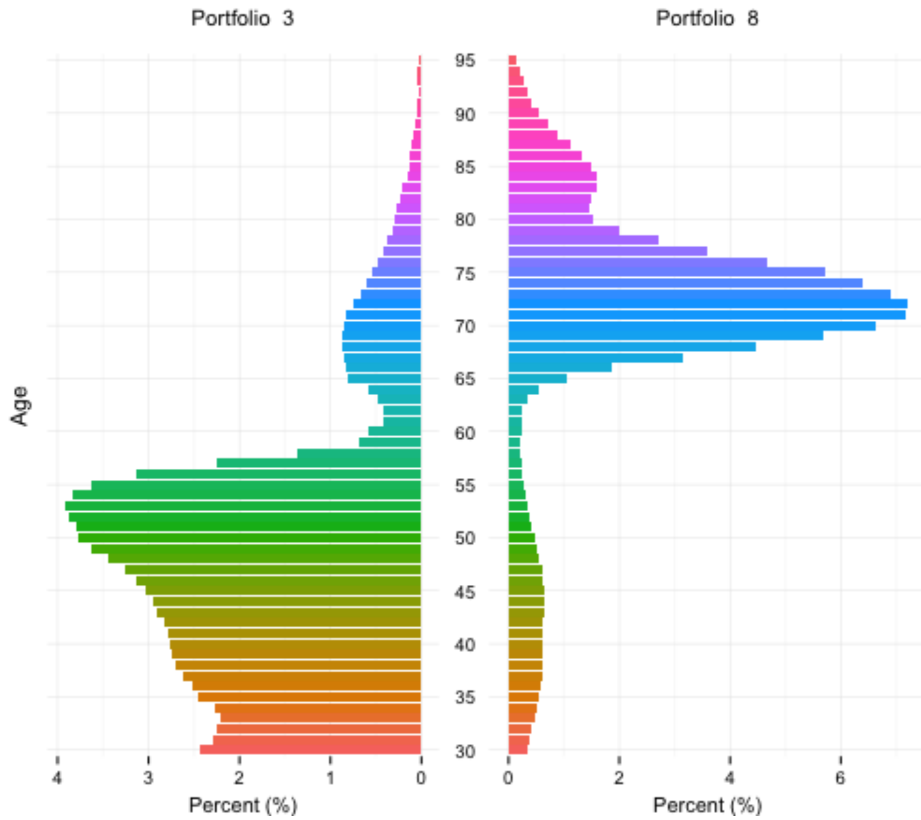

FIGURE 1.1 – Distribution of age groups in portfolios 3 (left panel) and 8 (right panel), male population.

1.5.2 The Baselines Mortality

We consider two prospective tables as baselines for our credibility models. One is the national demographic projections for the French population over the period 2007-2060, provided by the French National Office for Statistics, INSEE, see Blanpain and Chardon (2010). These projections are based on assumptions concerning fertility, mortality and migrations. We choose the baseline scenario among a total of 27 scenarios. The baseline scenario is based on the assumption that until 2060, the total fertility rate is remaining at a very high level (1.95). The decrease in sex and age-specific mortality rates is greater for men over 85 years old. The baseline assumption on migration consists in projecting a constant annual net-migration balance of $100,000$ inhabitants. The second external reference table, denoted IA2013, is a market table constructed for the French insurance market provided by *Institute des Actuaires*, see Tomas and Planchet (2013). It is worth to mention that this table is derived on mortality trends originating from the INSEE table and covers the period 2007-2060.

Following, assumption (i) in Section 1.2.3, the baseline mortality $q^{\mathrm{b}}_{x,t}$ is described by the Makeham model in (1.3). Table 1.2 presents the estimated parameters for each of the baselines considered.

TABLE 1.2 – Estimated parameters of the Makeham model (1.3) for the baselines of mortality considered, male population.

	INSEE		**IA2013**	
	2007–2009	**2007–2010**	**2007–2009**	**2007–2010**
$\widehat{A}^{\mathrm{b}}_T$	$4.2835e-03$	$4.2787e-03$	$2.1577e-04$	$2.4355e-04$
$\widehat{B}^{\mathrm{b}}_T$	$7.9564e-07$	$7.7199e-07$	$4.0863e-06$	$3.9935e-06$
$\widehat{C}^{\mathrm{b}}_T$	1.1484	1.1487	1.1211	1.1213

1.5.3 Adjustment of the Makeham model

Following assumptions (ii) in Section 1.2.3, we fit the Makeham model (1.3) for the baselines of mortality considered so as to estimate B^{b}_t for each calendar year while the parameters $\widehat{A}^{\mathrm{b}}_t = \widehat{A}^{\mathrm{b}}_T$ and $\widehat{C}^{\mathrm{b}}_t = \widehat{C}^{\mathrm{b}}_T$ remain fixed. Table 1.3 presents the estimated parameters for each year and baselines considered.

TABLE 1.3 – Estimated parameters of the Makeham model (1.3) for each year and baselines of mortality considered, male population.

	INSEE		**IA2013**	
	2007–2009	**2007–2010**	**2007–2009**	**2007–2010**
$\widehat{A}^{\mathrm{b}}_T$	$4.2835e-03$	$4.2787e-03$	$2.1577e-04$	$2.4355e-04$
$\widehat{B}^{\mathrm{b}}_{2007}$	$8.0826e-07$	$7.9035e-07$	$4.1740e-06$	$4.1204e-04$
$\widehat{B}^{\mathrm{b}}_{2008}$	$7.9554e-07$	$7.7790e-07$	$4.0843e-06$	$4.0319e-06$
$\widehat{B}^{\mathrm{b}}_{2009}$	$7.8318e-07$	$7.658e-07$	$4.0009e-06$	$3.9496e-06$
$\widehat{B}^{\mathrm{b}}_{2010}$	–	$7.5406e-07$	–	$3.8729e-06$
$\widehat{C}^{\mathrm{b}}_T$	1.1484	1.1487	1.1211	1.1213

1.5.4 Proximity Between the Observations and the Model

We assess the overall deviation with the observed mortality by comparing criteria measuring the distance between the observations and the models with the χ^2 Forfar et al. (1988a), the mean average percentage error (MAPE) Felipe et al. (2002) as well as the standardized mortality ratio (SMR) and the number of standardized residuals larger then 2 and 3, see Tomas and Planchet (2014). In addition, we find useful to use the SMR test proposed in Liddell (1984) and the likelihood ratio test. The tests and quantities summarizing the proximity between the observations and the model are described in the following. The χ^2 allows to measure the quality of the fit of the model. It writes,

$$\chi^2 = \sum_{(x,t)} \frac{\big(D_{x,t} - L_{x,t}\, \widehat{q}_x(t)\big)^2}{L_{x,t}\, \widehat{q}_x(t)\big(1 - \widehat{q}_x(t)\big)}.$$

The MAPE is the average of the absolute values of the deviations from the observations,

$$\text{MAPE} = \frac{\sum_{(x,t)} \big|\big(D_{x,t}/L_{x,t} - \widehat{q}_x(t)\big)/\big(D_{x,t}/L_{x,t}\big)\big|}{\sum_{(x,t)} D_{x,t}} \times 100.$$

We can also determine if the fit corresponds to the underlying mortality law (null hypothesis $\mathcal{H}_0$) with the likelihood ratio test. The statistic, ξ^{LR}, writes

$$\xi^{\text{LR}} = \sum_{(x,t)} \left(D_{x,t} \ln\left(\frac{D_{x,t}}{L_{x,t}\, \widehat{q}_x(t)}\right) + \big(L_{x,t} - D_{x,t}\big) \ln\left(\frac{L_{x,t} - D_{x,t}}{L_{x,t} - L_{x,t}\, \widehat{q}_x(t)}\right).\right).$$

If $\mathcal{H}_0$ is true, this statistic follows a χ^2 law with a number of degrees of freedom equal to the number of observations n : $\xi^{\text{LR}} \sim \chi^2(n)$. Hence, the null hypothesis $\mathcal{H}_0$ is rejected if $\xi^{\text{LR}} > \chi^2_{1-\alpha}(n)$, where $\chi^2_{1-\alpha}(n)$ is the $(1-\alpha)$ quantile of the χ^2 distribution with n degrees of freedom. The p-value is the lowest value of the type I error (α) for which we reject the test. We will privilege the model having the p-value $= \mathbb{P}\big[\chi^2_{1-\alpha}(n) > \xi^{\text{LR}}\big] = 1 - \text{F}_{\chi^2(n)}(\xi^{\text{LR}})$ closest to 1.
The SMR is computed as the ratio between the observed and fitted number of deaths :

$$\text{SMR} = \frac{\sum_{(x,t)} D_{x,t}}{\sum_{(x,t)} L_{x,t}\, \widehat{q}_x(t)}.$$

Hence, if SMR > 1, the fitted deaths are under-estimated and vice-versa if SMR < 1. Note that we can consider the SMR as a global criterion which does not take the age structure into account, compared to the chi2 and MAPE for instance. We can also apply a test to determine if the SMR is significantly different from 1. Liddell Liddell (1984) proposes to compute the statistic,

$$\xi^{\text{SMR}} = \begin{cases} 3 \times D^{\frac{1}{2}}\big(1 - (9D)^{-1} - (D/E)^{\frac{1}{3}}\big) & \text{If SMR } > 1, \\ 3 \times D^{*\,\frac{1}{2}}\big((D^*/E)^{\frac{1}{3}} + (9D^*)^{-1} - 1\big) & \text{If SMR } < 1, \end{cases}$$

where $D = \sum_{(x,t)} D_{x,t}$, $D^* = \sum_{(x,t)} D_{x,t}+1$ and $E = \sum_{(x,t)} L_{x,t}\, \widehat{q}_x(t)$. If the SMR is not significantly different from 1 (null hypothesis $\mathcal{H}_0$), this statistic follows a standard Normal law, $\xi^{\text{SMR}} \sim \text{N}(0,1)$.

Thus, the null hypothesis $\mathcal{H}_0$ is rejected if $\xi^{\text{SMR}} > \text{N}_{1-\alpha}(0,1)$, where $\text{N}_{1-\alpha}(0,1)$ is the $(1-\alpha)$ quantile of the standard Normal distribution. The p-value is given by p-value $= 1 - \text{F}_{\text{N}(0,1)}(\xi^{\text{SMR}})$.

1.5.5 In-Sample Numerical Analysis

We fitted the approaches over a history covering 3 and 4 years (2007-2009 and 2007-2010 respectively) and compared the overall deviation between the observations and the models (for the year 2010 and 2011 respectively). Table 1.4 displays the estimates of the structure parameters for the three approaches.

TABLE 1.4 – Estimates of the structure parameters, male population.

		Hardy-Panjer		**Poisson-Gamma**		**Makeham-Credibility**	
		INSEE	IA2013	INSEE	IA2013	INSEE	IA2103
2007-09	$\widehat{\mu}_0$	3.5521	16.3290	1	1	1	1
	$\widehat{\sigma}^2$	44.4032	92.1668	1	1	4.0552e-04	2.3198e-03
	$\widehat{\tau}^2$	6.8368	44.0092	10.7485	367.5029	0.1935	3.5960e-02
2007-10	$\widehat{\mu}_0$	3.6495	15.7865	1	1	1	1
	$\widehat{\sigma}^2$	65.9649	116.0159	1	1	5.1034e-04	2.6285e-03
	$\widehat{\tau}^2$	7.0772	43.4966	10.9684	338.4440	0.2217	5.0281e-02

Table 1.5 presents the tests and quantities summarizing the overall deviation between the observations and the credibility analysis for the male population of portfolio 1 obtained by the Hardy-Panjer, Poisson-Gamma and the Makeham credibility approaches with the two baselines mortality considered for the year 2010. Tables A.1, A.2 and A.3, A.4 in Appendix A.1 and A.2 display the results for all the portfolios and for the years 2010 and 2011 respectively.

The Hardy-Panjer and Poisson-Gamma approaches produce relatively similar graduations. However, we notice some differences with the Makeham credibility model which displays more favorable results whatever the baseline mortality considered for the two periods fitted.
It is also apparent that using the market baseline mortality IA2013 produces better results than the national demographic projections originating from INSEE, see Section 1.5.2. It illustrates the

TABLE 1.5 – Tests and quantities summarizing the deviation between the observations and the models for portfolio 1, calendar year 2010, male population.

		INSEE			**IA2103**		
		Hardy-Panjer	Poisson-Gamma	Makeham-Credibility	Hardy-Panjer	Poisson-Gamma	Makeham-Credibility
Standardized	> 2	60	60	35	46	46	15
residuals	> 3	48	48	28	32	32	5
χ^2		5481.86	5542.82	3569.97	1705.25	1747.25	208.81
MAPE	(%)	233.22	230.94	373.89	117.01	115.42	42.35
Likelihood	ξ^{LR}	946.98	947.72	443.16	463.48	468.46	88.03
ratio test	p-value	0	0	0	0	0	0.0364
	SMR	1.1792	1.1919	0.5265	1.7629	1.7957	1.0532
SMR test	ξ^{SMR}	4.0379	4.2939	12.1893	13.0352	13.4202	1.2845
	p-value	0	0	0	0	0	0.0995

importance of using an adequate baseline mortality when adjusting the models.
When looking at criteria and quantities which take the age structure of the error into account, the Makeham credibility approach is a benefit. The quality of the fit increases, sometimes drastically, compared to the Hardy-Panjer and Poisson-Gamma model in terms of having the minimum χ^2 and MAPE values. The Makeham credibility model leads to the lowest number of standardized residuals lower than 2 and 3. It exhibits as well the highest p-value for the likelihood ratio test.
Even when we consider a global indicator of the quality of the fit such as the SMR which does not take the age structure into account, the Makeham credibility model seems to perform better than the Hardy-Panjer and Poisson-Gamma approaches. The statistic ξ^{SMR} of the SMR test is the smaller 8 times over 14 for the year 2010, see Tables A.1 and A.2 in Appendix A.1, and 6 times over 12 for the year 2011, see Tables A.3 and A.4 in Appendix A.2.
We also notice that the the Makeham credibility model has tendency to over-estimate the total number of deaths, having a SMR lower than 1 for 9 portfolios over 14 in 2010 and for 8 portfolios over 12 in 2011.

In the following, these quantitative diagnostics are supplemented by a range of visual comparisons. Besides the tests and quantities, the comparison involves graphical analysis. It consists of representing graphically the fitted values against the observations for the years 2010 and 2011. For clarity, the graphical comparisons only consider the market baseline mortality IA2013 as it leads to better results than using the national demographic projections.

Figure 1.2 (top panel) displays the the fitted probabilities of death in the log scale for portfolio 1 for the year 2010. Figures A.1 and A.2 in Appendices A.3 and A.4 display the comparisons for all the portfolios and for the years 2010 and 2011 respectively. It gives us the opportunity to visualize the similarities and differences between the fits obtained by the approaches. It is again apparent that the Hardy-Panjer and Poisson-Gamma models lead to similar results. In addition, we observe

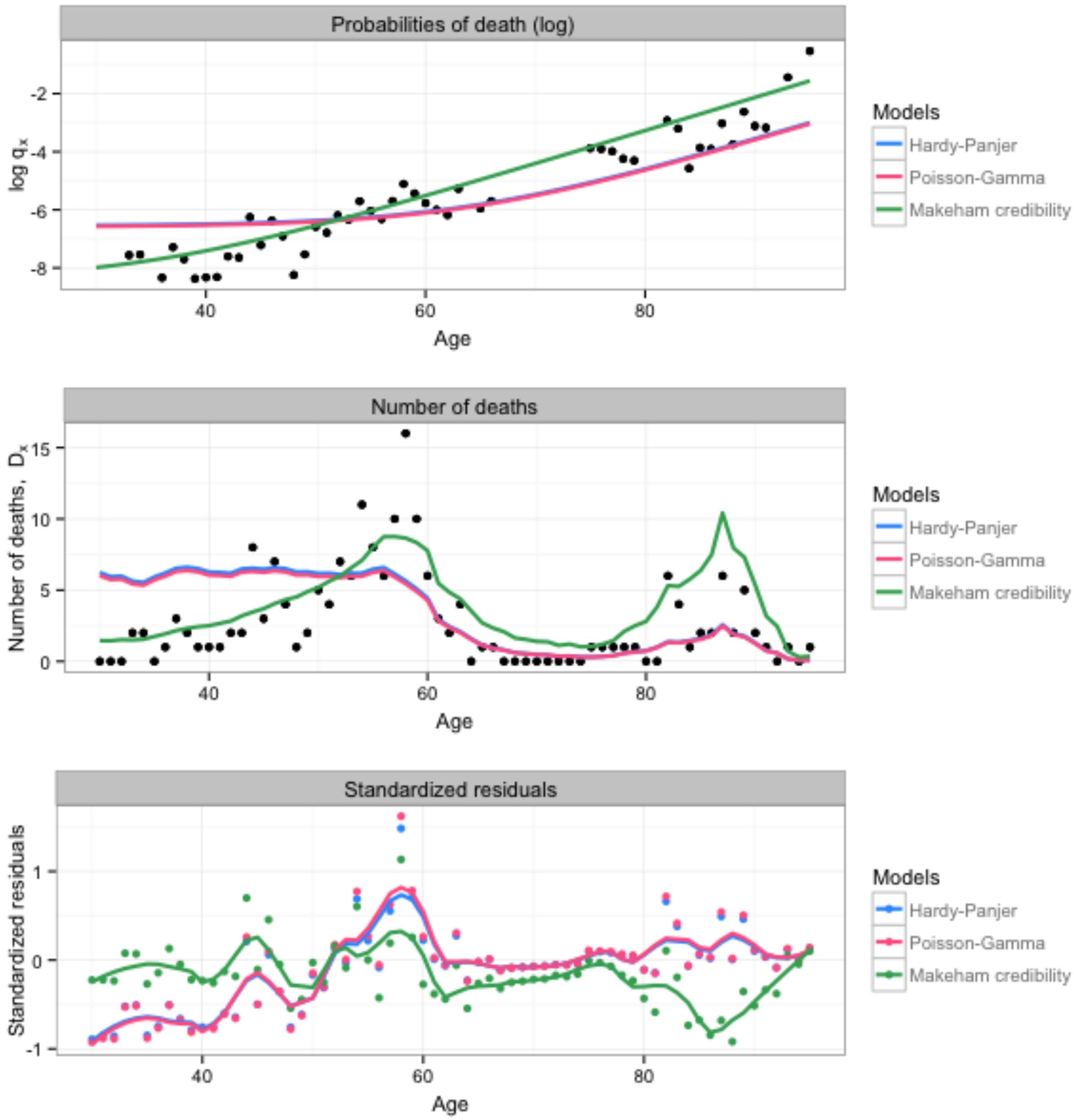

FIGURE 1.2 – Fitted values against the observations for portfolio 1 for the year 2010, male population. Top figure : Fitted probabilities of death in the log scale. Middle : Fitted number of deaths. Bottom : Standardized residuals.

that these approaches have a tendency to strongly overestimate the probabilities of death for the age band $[30, 60]$ and reciprocally underestimate them for the age band $[60, 95]$. This is explained by the fact that the age structure is not taken in account by the Hardy-Panjer and Poisson-Gamma approaches, contrary to the Makeham credibility model. We can visualize this lack of fit in the plots of the fitted number of deaths, Figure 1.2 (middle) for portfolio 1 and Figures A.3 and A.4 for all portfolios in Appendices A.5 and A.6.

In conjunction with looking to the plots of the fits, we should study the residuals plots. Such residual plots provide a powerful diagnostic that nicely complements the analysis. The diagnostic plots can show lack of fit locally and we have the opportunity to judge the lack of fit based on our knowledge on the data and of the performance of the models. We superimposed a smooth curve on the standardized residuals. This smoothness helps search for clusters of residuals that may indicate a lack of fit. The plots of the standardized residuals, for the male population, are display in Figure 1.2 (bottom panel) for portfolio 1 and Figure A.5 and A.6 in Appendix A.7 and A.8 for all the portfolios and for the years 2010 and 2011 respectively.
The standardized residuals, obtained by the Hardy-Panjer and Poisson-Gamma models, present a high curvature for most of the portfolios in Figures A.5 and A.6. It indicates a clear lack of fit. These models overestimate the number of deaths for the age band $[30, 60]$ et underestimate them for the age band $[60, 95]$, as observed in the plots of the fits previously. Conversely, no strong patterns appear in the standardized residuals retrieved for the Makeham credibility model. The smooth curves over the standardized residuals is meanly flat, meaning that no systematic reproducible lack of fit has been detected and that the Makeham credibility model captures adequately the variability of the data.

1.6 Concluding Remarks

We considered the periodic adjustment of a mortality graduated curve using a Makeham parametric model. This relies on the revision of a single parameter the two remaining been fixed. The framework considered here is closely related to the one introduced in Hardy and Panjer (1998). The main difference is the age-structure included through the parametric Makeham model. By doing so, we showed that adding an age structure enhances the predictive ability of the death forecast when we consider age-sensitive proxies. If one is only interested in predicting deaths at the aggregate portfolio level our methodology yields to the same forecast as in the Hardy and Panjer Hardy and Panjer (1998) framework. Moreover, we should note that in our methodology especially using the ratio of the considered Makeham parameters allows to overcome the de-trending step recommended in Hardy and Panjer (1998).

In order to assess the predictive power of our methodology, various other measures of risk and goodness-of-fit should be taken into account. Especially, we should consider the age-structure's impact on the prices and reserves and potential benefit of our model compared to the current market practice. There are also several piratical we do not address here which we openly acknowledge and leave for future research.

Chapitre 2

Approche semi-paramétrique

2.1 Introduction

A natural and straightforward approach to handle the issue introduced in chapitre is to use the available data at the portfolio level and build an entity-specific mortality table. However, practitioners may face technical difficulties related to the size of the portfolios and the heterogeneity of the guarantees (for the same underlying risk, say mortality risk). For instance, an insurer may have a fairly big portfolio but with policyholders holding different insurance contracts : pure endowment contracts, unit-linked contracts with minimum death guarantees, loan insurance and so on. In such a case, it is difficult to build a mortality table based on the sole experience of each product or guarantee. More precisely, the constructed table would not be able to represent the mortality profile of the policyholders thus failing in capturing the underlying risk. This should also be the case even if the mortality table is periodically updated with the incoming new data. If one draws a mortality table only based on the experience stemming from one product or guarantee, she shall have to face a problem of sample size. The latter arises not only at the portfolio level but also for individual ages. In fact, the mortality profile is highly dependent on the age of the individuals and some age groups being *poorly* represented may alter the quantification of the mortality risk at each individual age.

In this chapter, we consider an insurer with exposures to different coverages and aiming at establishing an experience-based mortality table for each policy and age level, as individuals may have different risk profiles (as showed by some empirical mortality studies, e.g. see Vaupel et al. (1979) and Hougaard (1984) among others). As a first step, we consider a graduation principle to build mortality rates at the insured portfolio level. There are usually two sorts of methods : nonparametric and parametric, see Forfar et al. (1988a) and Debón et al. (2006) for a comprehensive introduction to the use of both graduation techniques. The nonparametric framework is very useful in practice especially when there is sufficient data. This method relies on the use of kernel estimation techniques which were first used for graduation by Copas and Haberman (1983) and Ramlau-Hansen (1983). The literature on this subject has a long history and we may observe two schools. First, there is a continuous approach that defines data sampling via stochastic counting processes, which considers the lifetimes of individuals to be continuous random variables subject to random censorships, i.e. left truncation and right censoring. In our case, the mortality data that we use are divided into discrete yearly numbers of death occurrences and exposures. Therefore, these data only allow to use an approach based on an approximation of the continuous filtered model. Both the continuous and the discrete formulations have been intensively explored in the literature, see e.g. Fan and Gijbels (1995); Jiang and Doksum (2003); Nielsen et al. (2009) and more recently Gámiz et al. (2016). In these models the hazard rate is estimated using a nonparametric kernel method. A number of commonly used smoothing methods such as smoothing splines, kernel estimates and local polynomial fitting can be used to implement the basic step of the graduation of a mortality table. More recently, estimators based on local polynomial fitting, discussed in earlier works of Cleveland (1979) and Lejeune (1985), among others, have become more popular. This keen interest turned out, in particular for their good performance and analytical tractability, see for example the monograph by Fan and Gijbels (1996).

In the approach proposed here, local polynomial fitting methods are used as implementation of smoothing methods. This allows to model the mortality patterns even in presence of complex structures and avoid to rely on experts opinion. In Tomas (2011), the author used the same adaptive smoothing procedure applied to the dataset used throughout the chapter.

The graduated mortality can be then used as such to project future insurance liabilities related to the underlying population. However, the evolution of the flow of data related to latest available information is not taken into account. This should be, for example, used to update the graduated mortality. However, if one decides to redo a graduation procedure including the new data, the forecasts are likely to be unstable ; adding potential volatility to the underlying reserves and capital charges. Therefore, the primary contribution of this chapter is the incorporation of sample bias into the graduated mortality table model by introducing an unobserved variate for individual differences in each attained age. Such an approach has been considered in Salhi et al. (2016) but with different graduated curve. The latter used a parametric model, i.e. Makeham law, to first build the mortality curve and then applies a credibility procedure to a portfolio-sensitive parameter. Other approach have been also introduced in the literature but work directly on the aggregate death counts, e.g. Hardy and Panjer (1998). Unlike the classical approaches that focus on the update of the aggregate deaths recorded over the whole portfolio, the proposed adjustment approach is intended to enhance the predictive ability of the graduated mortality using a credibility-based revision at the age-level and not on the aggregate portfolio level, while borrowing information from other portfolios with sufficient information. More formally, our methodology is based on a discretization of the Nielsen and Sandqvist (2000) credibility approach, which was applied to the operational risk. The latter, however, did not consider a multidimensional credibility as the underlying risk does not exhibit an extra dimension rather than the observation date, see section 2.3.

The rest of the chapter is organized as follows. section 2.2 specifies the notation and assumptions used throughout the chapter. It also introduces the smoothing model in its general and continuous form. A discretization of the latter is considered as mentioned earlier and we also recall some statistical inference results used in the sequel. section 2.3 introduces the credibility approach to the graduated mortality. We specify the model and make connection with the recent literature. Furthermore, we derive the main tools needed to fully characterize the next period prediction of mortality rates when a (multiplicative) credibility factor is taken into account. section 2.4 presents an application with experience data originating from some French insurance companies. Finally, some remarks in section 2.5 conclude the chapter.

2.2 Notation, Assumptions and Preliminaries

2.2.1 Notation, Assumptions and Continuous Time Local Smoothing

Assume that we have at our disposal mortality statistics originating from K portfolios (or companies) over the time interval $[0, T_i]$, $i \in \{1, \cdots, K\}$. We suppose that the portfolios are composed of I_i individuals for which we associate a triplet $(Y_e^i, Z_e^k, \Delta_e^i)$, for $e = 1, \cdots, I_i$, where Y_e^i is the age that an individual enters the portfolio during the considered period, Z_e^i the age she leaves the portfolio and Δ_e^i an indicator of the censoring status. In other terms, Δ_e^i is equal to 1 when the individual deceases during the period $[0, T_i]$ and 0 when she leaves for other reasons, e.g. surrendering her policy. Based on this triplet, which can be observed in most life insurance portfolios, we let $N_e^i(x) = \Delta_e^i \mathbf{1}_{\{Z_e^i \leq x\}}$ be the counting process indicating the death of the individual e before age x. Similarly, we define the process $L_e^i(x) = \mathbf{1}_{\{Y_e^i \leq x < Z_e^i\}}$ that indicates if the insured is at risk at age x. For all the portfolios, we are considered with mortality behavior over an age interval $[x_1, x_{n_i}]$. Moreover, under usual conditions, we assume Cox (1972)'s multiplicative model where the random intensity of death φ_x^i, at age x of the portfolio i is related to a reference φ_x^{ref} as follows :

$$\varphi_x^i = \exp[f^i(x)]\varphi_x^{\text{ref}}, \tag{2.1}$$

where f^i is an unspecified, smooth and deterministic function of the age x. The latter allows to link the mortality of the company i to the baseline at the attained age level x. Here, we adopt a parametric form for the functional f^i and denote β^i this vector of parameters which will be specified later on this section.

Remark. *(i) In this assumption, the baseline mortality is shared over portfolios. However, the functional f^i is not common as it is supposed to adjust to the particular feature of each portfolio. That is the form as well as the parameters may depend on the sample size and particularly over ages. The form of the latter will be common and will defined to be of polynomial form. However, the degree will be adapted to each portfolio. (ii) As we can see later on this chapter, the reference mortality φ_x^{ref} is constructed using the aggregate data stemming from the portfolios, which will* *underpin the use of the common baseline. However, it may be of interest to consider a full Cox model taking into account the specific features of each portfolio. This is, for instance, investigated in Nielsen and Sandqvist (2005), where it is taken into account that mortality rates should not be around a common mean, but around a Cox regression instead. By doing so, it allows the approach to be used even when the lines of mortality are different, as long as they fit into a proportional hazard framework, see Gustafsson et al. (2006, 2009).*

The specification in Équation 2.1 is a simple variation of the Cox's proportional hazards regression model. This was considered, for example, in Anderson and Senthilselvan (1980); Gray (1990) using a known link function but with covariates that adjust the mortality given the observed heterogeneity. The general Cox's model, in the presence of covariates, with unknown link function is considered in Wang (2001, 2004) who proposed a local likelihood approach to estimate the function f^i. Formally, under the above assumptions, the likelihood functional $\mathcal{L}(\varphi^i;\beta^i)$ in the presence of the left-truncation and right-censoring is given as follows :

$$\mathcal{L}(\varphi^i;\beta^i) = \prod_{e|Y_e^i \leq Z_e^i} \left[(\varphi^i_{Z_e^i})^{\Delta_e^i} \exp\left(\int_{Y_e^i}^{Z_e^i} \varphi_s^i ds \right) \right].$$

Therefore,

$$\begin{aligned} \log \mathcal{L}(\varphi^i;\beta^i) &= \sum_{e|Y_i^k \leq Z_i^k} \left[\Delta_i^k \log\left(\varphi^k_{Z_i^k}\right) - \int \mathbf{1}_{\{Y_i^k \leq s < Z_i^k\}} \varphi_s^k ds \right] \\ &= \int \log(\varphi_s^k) dN^k(s) - L^k(s)\varphi_s^k ds, \end{aligned}$$

where $N^i(x) = \sum_{e=1}^{I_i} N_e^i(x)$ and $L^i(x) = \sum_{e=1}^{I_i} L_e^i(x)$. In the light of the foregoing, we consider the local likelihood model which fits a polynomial model locally within a smoothing window. To this end, the localized log-likelihood at an age x can be written as follows :

$$\log \mathcal{L}^{\text{loc}}(\varphi_x^i;\beta^i) = \int \omega_h(s-x)\log(\varphi_s^i)dD^i(s) - \omega_h(s-x)L^i(s)\varphi_s^i ds, \tag{2.2}$$

where $\omega_h(u)$ is a weight function with a bandwidth parameter $h > 0$ that assigns largest weights to observations close to x. These considerations will yield the local kernel weighted log-likelihood estimation of the polynomial function f^i. Such a formulation complies with the literature on local polynomial hazard estimation, see Fan and Gijbels (1995); Jiang and Doksum (2003) and Gámiz et al. (2016). We assume that $f^i(x_j)$ is a p^{th} degree polynomial in x_l's, where x_l is an element in the neighborhood of x_j. Formally, denoting $\mathbf{x}_l = \left(1, x_l - x_j, \cdots, (x_l - x_j)^p\right)^\top$ and $\beta^i = \left(\beta_0^i, \cdots, \beta_p^i\right)^\top$ we can write $f^i(x_j)$ in the following form $f^i(x_j) = \mathbf{x}_l^\top \beta^i$.

Remark. *Various forms of the function f have been considered in empirical actuarial science. For example, in Currie (2013), the function f has the parametric form $f^i(x) = \beta_0^i + \beta_1^i x$ for some unknown parameters β_0^i and β_1^i. Other examples were considered in Renshaw et al. (1996).*

2.2.2 Local Likelihood Smoothing of Mortality in Discrete Time

Up to now, we considered the lifetimes of individuals to be continuous random variables subject to random censorships. In our case, the mortality data at our disposal are divided into discrete yearly numbers of death occurrences and exposures. Therefore, these data only allow to use an approach based on an approximation of the continuous filtered model in Équation 2.2. As noted before, both the continuous and the discrete formulations have been intensively explored in the literature, see e.g. Fan and Gijbels (1995); Jiang and Doksum (2003) and more recently Gámiz et al. (2016). The latter provides a theoretical treatment of local linear mortalities and it is also describing in detail the relationship between discrete and continuous sampling. In actuarial literature, early works based on discrete data date a long way back to Gram (1879, 1883) who develops local polynomial hazard estimators that are not far in spirit from our work.

The discretization of Équation 2.2 relies on an aggregation of the lifetimes into intervals. In this subsection, we describe a modification of the local linear estimator for discrete data in Équation 2.2. We suppose that the following yearly aggregated values of occurrences and exposures are available :

$$D^i_{x_j} = \sum_{e=1}^{I_i} \int_{x_j}^{x_{j+1}} dN^i_e(s), \quad E^i_{x_j} = \sum_{e=1}^{I_i} \int_{x_j}^{x_{j+1}} L^i_e(s)ds. \tag{2.3}$$

These refer, respectively, to the number of deaths and the number of individuals who are at risk in the age interval $[x_j, x_{j+1}[$. Moreover, we assume a piecewise constant hazard rate φ^i_x in the sense that $\varphi^i_x = \varphi^i_{x_j}$ for any $x \in [x_j, x_{j+1}[$. Then, a natural approximation of the localized likelihood function in a neighborhood of x_j, i.e. Équation 2.2, would be

$$\begin{aligned} \log \mathcal{L}(\varphi^i_{x_j}; \beta^i) &= \sum_{l=1}^{m} \omega_h(x_l - x_j) \log(\varphi^i_{x_l}) D^i_{x_l} - \omega_h(x_l - x_j) \varphi^k_{x_l} E^i_{x_l} \\ &= \sum_{l=1}^{m} \omega_{lj} \mathbf{x}_l^\top \beta^i D^i_{x_l} - \omega_{lj} \varphi^{\text{ref}} e^{\mathbf{x}_j^\top \beta^i} E^i_{x_l} + C^i, \end{aligned} \tag{2.4}$$

where C^i is a constant offset, which does not depend on the parameter vector β^i.

Remark. *The true likelihood given in Équation 2.4 can be recovered, up to a constant offset, using the hypothesis of Poisson distributed death occurrences. In fact, if the parameter of the Poisson distribution is assumed to be $E^i_x \varphi^i_x$ where the intensity φ^i_x is as in Équation 2.1 then, one can write the problem as a generalized linear model (GLM) such that the first moment of D^i_t can be written as follows :*

$$\log \mathbb{E}[D^i_x] = \log E^i_x + \log \varphi^i_x = \log E^i_x + \log \varphi^{ref}_x + f^i(x),$$

where the term $\log E^i_x$ is an offset. Then, in the presence of unknown link function f^i, we can rely on a localized likelihood version which add a weight to the observations at each age. Such an approach was used to graduate life tables with attained age context in Delwarde et al. (2004), Debón et al. (2006) and Tomas (2011).

In Équation 2.4, the non-negative weights, i.e. ω_{lj}, depend on the distance between the observations and the fitting point x_j and can be characterized using the kernel ω_h as follows

$$\omega_{lj} = \begin{cases} \omega(|x_l - x_j|/h), & \text{if } |x_l - x_j| \leq h, \\ 0, & \text{otherwise}, \end{cases} \tag{2.5}$$

with ω is Gaussian kernel, and h is a smoothing parameter determining the radius of the neighborhood of x_j used in the smoothing. It gives the bandwidth of the neighbor used in the kernel. For instance, the smallest h the thiner is the neighborhood that contributes to the likelihood at each attained age.

In order to estimate the parameters vector β^i we maximize the log-likelihood in (2.4). To this end, we let $\mathbf{D}^i = (D^i_{x_1}, \cdots, D^i_{x_m})$ and $\varphi^i = (\varphi^i_{x_1}, \cdots, \varphi^i_{x_m})$. Then, taking the derivative with respect to β^i, yields the following system of equations,

$$(\mathbf{X}^j)^\top \mathbf{W}^j (\mathbf{D}^i - \varphi^i) = 0, \tag{2.6}$$

where $\mathbf{X}^j$ is the $m \times (p+1)$ matrix

$$\mathbf{X}^j = \begin{pmatrix} 1 & x_1 - x_j & (x_1 - x_j)^2 & \cdots & (x_1 - x_j)^p \\ 1 & x_2 - x_j & (x_2 - x_j)^2 & \cdots & (x_2 - x_j)^p \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_m - x_j & (x_m - x_j)^2 & \cdots & (x_m - x_j)^p \end{pmatrix}, \tag{2.7}$$

and $\mathbf{W}^j$ is the $m \times m$ diagonal weight matrix with diagonal elements w_{lj}, for $l = 1, \cdots, m$. Since φ^i is non-linear on β^i. The solution of the above equation, i.e. estimations, must be obtained numerically using, for example, an iterative algorithm as Nelder-Wedderburn, Newton-Raphson algorithms or the fisher scoring methodology, see Loader (2006, Chapter 12) for further development. From these, we can get the estimation of β^i and φ^i denoted, henceforth, by $\widehat{\beta}^i$ and $\widehat{\varphi}^i$.

2.2.3 Inference of the Graduated Mortality

The aim of this subsection is to characterize the statistical feature of the estimators considered above. We recall some well known results in the literature on nonparametric smoothing, see e.g. Tibshirani and Hastie (1987) and Wand and Jones (1994), regarding particularly the variance of the graduated mortality and the expected behavior of these estimations. In fact, using theoretical results concerning bias and variance, the estimator $\widehat{\varphi}^i$ is shown to be asymptotically robust and consistent. It is, for instance, shown in Fan and Gijbels (1996) that the smoothed mortality rates $\widehat{\varphi}^i$ are unbiased estimators of φ^i in the sense that :

$$\mathbb{E}[\widehat{\varphi}^i] \approx \varphi^i. \tag{2.8}$$

This approximation is based on the inspection of the mean squared errors, which are commonly used to assess the bias of the estimation in such a framework. Expressions of the latter are available in classical textbooks and the readers is referred to (a modifier la ref et mettre un livre)(Tomas, 2011, Sec. 3.2.) who provides an approximation to the bias of the estimator $\widehat{\varphi}^i$. Unlike the linear model fitting, there is no exact expression for moments of $\widehat{\varphi}^i_x$ due to the non-linearity in Équation 2.6. Using a multivariate version of Taylor series expansion around $\boldsymbol{\beta}^i$ allows to use classical results on the inference of the estimated parameter $\widehat{\boldsymbol{\beta}}^i$. Note that this approximation depends on the bandwidth of the neighborhood h used in the kernel. More precisely, the bias is decreasing with the bandwidth. This is, particularly, reasonable in practice, because a large bandwidth induces a miss-fitting of the local polynomials and hence also the sum of squared residuals. In other hand, to derive the second order moment of $\widehat{\boldsymbol{\varphi}}^i$, a variance approximation based on Taylor linearization is also generally suggested and shown to be consistent, see Loader (2006). More precisely, we have the following expression for the variance :

$$\mathbb{V}\mathrm{ar}(\widehat{\varphi}^i) \quad = \quad (\mathbf{S}^i \mathbf{S}^{i\top})\widehat{\varphi}^i, \tag{2.9}$$

where the matrix $\mathbf{S}^i$ is given as follows :

$$\mathbf{S}^i = \begin{pmatrix} s_1^i(x_1) & s_2^i(x_1) & s_3^i(x_1) & \cdots & s_n^i(x_1) \\ s_1^i(x_2) & s_2^i(x_2) & s_3^i(x_2) & \cdots & s_n^i(x_2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ s_1^i(x_n) & s_2^i(x_n) & s_3^i(x_n) & \cdots & s_n^i(x_n) \end{pmatrix}, \tag{2.10}$$

with rows $\mathbf{s}^i(x_j)^\top = (s_1^i(x_j), s_2^i(x_j), \cdots, s_n^i(x_j)) = (\mathbf{X}^{j\top}\mathbf{W}^j\mathbf{X}^j)^{-1}\mathbf{X}^{j\top}\mathbf{W}^j$, where $\mathbf{W}^j$ is the weight matrix and $\mathbf{X}^j$ is given in Équation 2.7.

2.2.4 Small Sized Porfolios and Sampling Bias

It is worth mentioning that the relational (proportional) model considered in Équation 2.1 implicitly accounts for differential mortality that may arise due to portfolio specific features, e.g. particular socioeconomic groups involved, income level, etc. This is all the more true if we consider the national mortality as a baseline as insured portfolios show a typical behavior compared to a national mortality. Specifically, the mortality of insured population is significantly lower than the national population from which it is drawn. On the other hand, when it comes to the study of the mortality at a single portfolio level, some stylized facts arise, which might compromise the efficiency of the graduation procedure. For instance, insured population are generally of small size, so none or very few deaths are observable at some ages. Therefore, the use of the model (2.1) and the local-likelihood based estimation procedure advocates using the information stemming from the adjacent ages to construct the mortality curve. This *learning* procedure will enhance the determination of the mortality at a given age. However, when successive ages lack of information, the approach exposed above will need a large bandwidth h for the estimator to access distant ages with sufficient and reliable information. By doing so, we increase the bias surrounding the smoothed curve. Indeed, as noted before, the mean of squared errors measuring the bias due to the local regression increases with the bandwidth h.

Due to these different sources of uncertainty we suppose that the *true* mortality curve φ_x^i, for $x = x_1 \cdots, x_n$, is known up to an unobservable multiplicative factor Θ_x^i. In other words, the portfolios examined should be regarded as a sample of the reference. Estimates based on the data will be subject to sampling errors and the smaller the group is, the bigger will be the relative random errors in the number of deaths and the less reliable will be the resulting estimates. This argument is extended to include the bias stemming at the attained age level due the consideration exposed above. Thus, if one has estimated the curve using the non-parametric approach, the *true* curve is an adjustment of the latter as multiplied by the random and non-observable parameter Θ_x^i. Such a setting is inspired by the credibility approach to hazard estimation of Nielsen and Sandqvist (2000).

2.3 Company-Specific Relative Risk Level

Recall that we have at our disposal K portfolios with individuals ages ranging from x_1 to x_{n_i}. Here, the n_i's could be all different to be in line with the insurance practices. This kind of information structure is similar to the so-called *unbalanced* framework used in actuarial science. For the sake of readability, without loss of generality, we will henceforth assume similar observed age groups for all companies, i.e. $n = n_1 = \cdots = n_K$. With a slight variation to the model, however, it can be easily extended to the unbalanced case.

2.3.1 The credibility model

Given the specific parameterization of the problem, one may think of the K portfolios as subsets of the reference population and each portfolio is characterized by a risk profile. The latter is due to the heterogeneous sizes of the portfolios as well as the underlying guarantees (for the same underlying risk). These sources of heterogeneity might also induce an age varying risk profile within the same portfolio. Therefore, for a company i, we let the vector $\boldsymbol{\Theta}^i = \text{diag}(\Theta_{x_1}^i, \cdots, \Theta_{x_n}^i)$ be her relative risk level. For $x \in \{x_1, \cdots, x_n\}$, each Θ_x^i characterizes the age-specific risk level, which are unobservable random variables.

The primal objective is to characterize the force of mortality of each company i at a specific age x through the proportional relationship introduced in sous-section 2.2.2, i.e.

$$\varphi^i = \mathbf{\Theta}^i \alpha, \tag{2.11}$$

where $\alpha = (\alpha_{x_1}, \cdots, \alpha_{x_n})^\top$ such that for $j = 1, \cdots, n$, we have $\alpha_{x_j} = \exp[f^i(x_j)]\varphi^{\text{ref}}_{x_j}$. This model suggests that for each company i, the age-specific experienced force of mortality varies around the baseline α_x, which can be seen as a reference or best-estimate mortality. This fluctuation is modeled by a heterogeneity parameter Θ^i_x capturing the individual properties (heterogeneity) of each company at attained age x. Thus, using new incoming data should allow to update the next period mortality φ^i_x by adjustment following model in Equation (2.11). The approach is to first find an estimator $\widehat{\varphi}^i_x$ of $\varphi^i_x = \Theta^i_x \alpha_x$ for each company i using the likelihood-based approach introduced in section 2.2. Henceforth, the notation $\widehat{\varphi}^i_x | \Theta^i_x$ refers to the estimation of the quantity $\alpha_x = \exp[f^i(x)]\varphi^{\text{ref}}_x$, which, by abuse of language, referred to as the estimated mortality conditional on the risk profile.

Remark. *This reasoning is build upon the work of Nielsen and Sandqvist (2000, 2005) and Gustafsson et al. (2006), that investigate the quantification fo the operational risk. Nielsen and Sandqvist (2000) considered hazards of different groups assuming that the hazard of each group fluctuates across a common baseline hazard and used continuous sampling of observations as in sous-section 2.2.1. In the current work, we are considering a discretization of the model in Nielsen and Sandqvist (2000) as the mortality data that we will use are divided into discrete yearly numbers of death occurrences and exposures. Moreover, we slightly extend this framework by considering a multivariate setting and allow for the age to influence the estimation of future mortality. Using a multivariate framework will provide a base to catch the sample bias properties at the attained age level. This is even more significant given that the mortality intensities are correlated not only at the portfolio level but also between different portfolios.*

The random variables $\widehat{\varphi}^i_{x_1}, \cdots, \widehat{\varphi}^i_{x_n}$ are assumed to be dependent, namely, the force of mortality of one age does directly impact those of other ages. This is mainly due to the graduation of mortality at a given age, which weights up over the adjacent age groups, see section 2.2. This dependency will be explored later on this section. Finally, in order to characterize the next period mortality level, we make use of the credibility theory. For this purpose and using the usual credibility setting, we shall make the following assumptions :

(A1) The random vectors $\mathbf{\Theta}^i$ are independent across companies and ages. Moreover, for $i = 1, \cdots, K$, Θ^i_x's are identically distributed with $\mathbb{E}[\mathbf{\Theta}^i] = \mathbf{I}_n$ and $\mathbb{V}\text{ar}(\mathbf{\Theta}^i) = \boldsymbol{\sigma}$, where $\boldsymbol{\sigma}$ is a diagonal matrix with elements σ_x and $\mathbf{I}_n$ is the identity matrix.

(A2) The random vectors $(\varphi^i, \mathbf{\Theta}^i)$, $i = 1, \cdots, K$, are independent across companies

(A3) $\varphi^i_{x_1}, \cdots, \varphi^i_{x_n}$ are conditionally independent given $\mathbf{\Theta}^i$.

The first assumption (A1) ensures that the baseline mortality produces the a priori expected number of deaths under the model assumption (2.1), in the sense that $\mathbb{E}[D^i_x] = \mathbb{E}[\Theta^i_x \alpha_x] = \alpha_x$. The assumptions (A2) means that the risk profiles are independent over portfolios. In other words, the successive realizations of the mortality intensity (so as the death counts) for any portfolio are independent of each other except through the risk parameter. This assumption is commonly used in actuarial literature when dealing with mortality risk. The latter is in line practical uses and empirical results on mortality risk. However, note that this only makes sense for mortality-contingent contracts. Thus, we should exclude annuities and pension policies where a dependence over observations is present due, for instance, to the cohort effect. Finally, assumption (A3) translates the dependency of the mortality over ages. It is only captured by the vector $\boldsymbol{\Theta}^i$. Conditionally on the latter the forces of mortality at the age level are independent.

As noted before, $\widehat{\varphi}^i|\boldsymbol{\Theta}^i$ is the *conditional* local-likelihood estimator of the intensity in Equation (2.1) based on the data from the i^{th} portfolio as developed in section 2.2. In view of the the assumptions (A1)-(A3), it is important to recall that conditional on the knowledge of the risk profile $\boldsymbol{\Theta}^i$ the theoretical properties of $\widehat{\varphi}^i$ are identical to those the of local-likelihood estimator considered in sous-section 2.2.3. This will be used, among others, in the following lemma, in order to state some fundamental features of the dependence structure.

Lemma 1. *Under assumptions (A1), (A2) and (A3) and the notation above, we have*
(i) The first order moment of φ^i is given by

$$\mathbb{E}[\varphi^i] = \alpha. \tag{2.12}$$

(ii) The variance matrix of $\widehat{\varphi}^i|\boldsymbol{\Theta}^i$, denoted $\boldsymbol{\Sigma}^i(\boldsymbol{\Theta}^i) = \mathbb{V}ar(\widehat{\varphi}^i|\boldsymbol{\Theta}^i)$, is given by

$$\boldsymbol{\Sigma}^i(\boldsymbol{\Theta}^i) = (\boldsymbol{S}^i\boldsymbol{S}^{i\top})\varphi^i. \tag{2.13}$$

Hence, the variance $\boldsymbol{\Sigma}^i = \mathbb{V}ar(\widehat{\varphi}^i)$ can be written as

$$\boldsymbol{\Sigma}^i = \big(\boldsymbol{S}^i\boldsymbol{S}^{i\top} + \sigma\big)\alpha \tag{2.14}$$

(iii) The covariance of φ^i_x with $\widehat{\varphi}^i_x$ is given by

$$\mathbb{C}ov\left(\varphi^i_x, \widehat{\varphi}^i\right) = \left(\sigma_x\alpha_x\right)^2\mathbf{e}_{\delta_x}, \tag{2.15}$$

with $\delta_x = j$ if $x = x_j$ and $\mathbf{e}_j$ is the vector with all 0's except for a 1 in the j^{th} coordinate.

Démonstration. To show these results, we make an intensive use of the law of total variance.
(i) Équation 2.12 is a direct consequence of assumption (A1) which gives $\mathbb{E}[\varphi^i|\boldsymbol{\Theta}^i] = \boldsymbol{\Theta}^i\alpha$.
(ii) The conditional variance $\boldsymbol{\Sigma}^i(\boldsymbol{\Theta}^i)$ is directly derived from the calculus in sous-section 2.2.3. Hence, to check (2.14), the law of total variance gives

$$\begin{aligned}\boldsymbol{\Sigma}^i &= \mathbb{E}\big[\mathbb{V}\text{ar}(\widehat{\varphi}^i|\boldsymbol{\Theta}^i)\big] + \mathbb{V}\text{ar}\big(\mathbb{E}[\widehat{\varphi}^i|\boldsymbol{\Theta}^i]\big),\\ &= (\boldsymbol{S}^i\boldsymbol{S}^{i\top})\mathbb{E}[\widehat{\varphi}^i] + \mathbb{V}\text{ar}(\boldsymbol{\Theta}^i)\alpha = \big(\boldsymbol{S}^i\boldsymbol{S}^{i\top} + \sigma\big)\alpha.\end{aligned}$$

(iii) Finally, to prove (2.15), notice that $\mathbb{C}\text{ov}(\varphi^i_x, \widehat{\varphi}^i|\boldsymbol{\Theta}^i) = 0$. Thus,

$$\begin{aligned}\mathbb{C}\text{ov}\big(\varphi^i_x, \widehat{\varphi}^i\big) &= \mathbb{C}\text{ov}\big(\mathbb{E}[\varphi^i(x)|\boldsymbol{\Theta}^i], \mathbb{E}[\widehat{\varphi}^i|\boldsymbol{\Theta}^i]\big) + \mathbb{E}\big[\mathbb{C}\text{ov}(\varphi^i_x, \widehat{\varphi}^i|\boldsymbol{\Theta}^i)\big],\\ &= \mathbb{C}\text{ov}(\Theta^i_x\alpha_x, \boldsymbol{\Theta}^i\alpha) = \big(\sigma_x\alpha_x\big)^2\mathbf{e}_{\delta_x},\end{aligned}$$

where the last equality follows from the independence assumption in (A1). □

2.3.2 The next-period linear per-age mortality estimator

The goal is to predict the future force of mortality for each company i at the age level. Therefore, we will be looking for the inhomogeneous credibility predictor corresponding to the linear estimators of φ_x^i. We thus solve the following optimization problems :

$$\min_{c_{0,x}^i, \mathbf{c}_x^i} \mathbb{E}\left[\left(\varphi_x^i - c_{0,x}^i - \mathbf{c}_x^{i\top}\widehat{\varphi}^i\right)^2\right], \tag{2.16}$$

where $c_{0,x}^i \in \mathbb{R}$ and $\mathbf{c}_x^i \in \mathbb{R}^n$. This formulation suggests adjusting the next period force of mortality at a given age using the information stemming from the other age groups. This should enhance the prediction for ages with low or sparse information using the credibility in ages of high information. Based on Proposition 1, we can easily derive the inhomogeneous credibility estimators of φ^i. Indeed, we can state the following proposition.

Proposition 1. *The point estimate of the linear factors in (2.16) can be written as follows*

$$c_{0,x}^i = \left(\mathbf{1}_n - \mathbf{c}_x^i\right)^\top \alpha \quad \text{and} \quad \mathbf{c}_x^i = \left(\sigma_x \alpha_x^i\right)^2 (\boldsymbol{\Sigma}^i)^{-1}\mathbf{e}_{\delta_x}. \tag{2.17}$$

The next period predicted mortality (estimator) $\widetilde{\varphi}^i$ of φ^i is given by

$$\widetilde{\varphi}^i = (\mathbf{I}_n - \left(\alpha(\boldsymbol{\Sigma}^i)^{-1}\sigma\alpha\right)^\top)\alpha + \left(\alpha(\boldsymbol{\Sigma}^i)^{-1}\sigma\alpha\right)^\top\widehat{\varphi}^i. \tag{2.18}$$

Démonstration. Let us first derive the intercept $c_{0,x}^i$. To do this, we develop the expectation in Équation 2.16 and take the derivative with respect to $c_{0,x}^i$. This yields to the following equality :

$$c_{i,0} + (\mathbf{c}_x^i)^\top \mathbb{E}[\widehat{\varphi}^i] = 1.$$

On the other hand, differentiating the expectation in Équation 2.16 with respect to the vector $\mathbf{c}_x^i$ gives rise to the following variance

$$\mathbb{V}\text{ar}\left(\varphi_x^i - (\mathbf{c}_x^i)^\top\widehat{\varphi}^i\right),$$

needed to fully characterize the solution. This can be computed using results in Lemma 1. Indeed, we can write :

$$\mathbb{V}\text{ar}\left(\varphi_x^i - (\mathbf{c}_x^i)^\top\widehat{\varphi}^i\right) = \mathbb{V}\text{ar}(\varphi_x^i) - 2(\mathbf{c}_x^i)^\top\mathbb{C}\text{ov}(\varphi_x^i, \widehat{\varphi}^i) + (\mathbf{c}_x^i)\boldsymbol{\Sigma}^i(\mathbf{c}_x^i)^\top.$$

Taking the derivative with respect to the vector $\mathbf{c}_x^i$ yields

$$2\mathbb{C}\text{ov}(\varphi_x^i, \widehat{\varphi}^i) - 2\boldsymbol{\Sigma}^i\mathbf{c}_x^i = 0.$$

The terms $\boldsymbol{\Sigma}^i$ and $\mathbb{C}\text{ov}(\varphi_x^i, \widehat{\varphi}^i)$ are given in Lemma 1, which concludes the proof. □

Note that we are able to estimate all the components needed to characterize the next-period intensity $\widetilde{\varphi}^i$, except for the variance σ. Remarking that $\widehat{\varphi}^i$ is an estimator of $\varphi^i = \mathbf{\Theta}^i\alpha$, we can write $\widehat{\mathbf{\Theta}}^i = \text{diag}(\widehat{\varphi}^i \oslash \alpha)$, with " $\oslash$ " being the Hadamard division (element-wise) operator. Therefore, a natural choice of the estimator of $\boldsymbol{\sigma}$ is

$$\widehat{\boldsymbol{\sigma}} = \big(\widehat{\mathbf{\Theta}}^i - \mathbf{I}_n\big)^\top \big(\widehat{\mathbf{\Theta}}^i - \mathbf{I}_n\big). \tag{2.19}$$

We can now derive the following estimation of the adjustment factor $\mathbf{\Theta}^i$.

Lemma 2. *The optimal credibility estimator of $\mathbf{\Theta}^i$ is given by*

$$\widetilde{\widetilde{\mathbf{\Theta}}}^i = (\mathbf{I}_n - \big(\alpha(\mathbf{\Sigma}^i)^{-1}\widehat{\sigma}\alpha\big)^\top)\mathbf{1}_n + \big(\alpha(\mathbf{\Sigma}^i)^{-1}\widehat{\sigma}\alpha\big)^\top \widehat{\mathbf{\Theta}}^i,$$

and the next period prediction of φ^i can be approximated by $\widetilde{\widetilde{\mathbf{\Theta}}}^i\alpha$.

Remark. *The adjustment procedure described in Proposition 1 and Lemma 2 can be written for each individual age x in the classical form $\widetilde{\varphi}^i_x = (1 - z^i_x)\alpha_x + z^i_x\widehat{\varphi}^i_x$, where z^i_x is the credibility factor given as follows*

$$z^i_x = (\alpha_x)^2\widehat{\sigma}^2_i \left[(\alpha_x)^2\widehat{\sigma}^2_i + \widehat{\varphi}^i_x\|\mathbf{s}^i(x)\|^2\right]^{-1}.$$

Here, recall that $\|\mathbf{s}^i(x)\|^2 = \sum_{j=1}^n \big(s^i_j(x)\big)^2$ and measures the reduction in variance of the smoothed mortality curve $\widehat{\varphi}^i_x$.

Remark. *All the ingredients required to implement the credibility approach in Lemma 2, in order to predict the next-period estimator, are already determined. However, we still need to characterize an estimation for α. To do this, we borrow the same procedure considered in Nielsen and Sandqvist (2000), which amend to estimate α as a linear weighted average over the portfolios.*

2.4 Numerical Analysis

2.4.1 Source of Data

The data come from studies conducted by *Institut des Actuaires*. These include in total 14 portfolios covering the period 2007 − 2011 with each company's contributing data for at least 4 of a possible 5 years. Tableau 2.1 presents the observed characteristics of the male population of these portfolios. For this dataset, we are considering a period of $T = 4$ years for all companies. The remaining year serves to test the predictive feature of the model using an in-sample analysis. The considered analysis follows similar lines as in Salhi et al. (2016), which also exploit the same dataset. Therefore, the age band for all companies ranges from $x = 30$ to 95 years old. The Figure 2.1 shows the age distribution of the portfolios (in percentage), i.e. the aggregate number of individuals exposed to risk at each attained age. It graphically depicts the size heterogeneity observed between the portfolios with insured holding different coverages. These portfolios are not only of different sizes but also of different age pyramids. For example, the portfolio P1 corresponds to a typical death contingent coverage. In fact, the latter has a concentration on middle aged populations with few exposure at high ages. When portfolios, such as P2, are concerned, we should note that those are not contingent to life annuities but rather correspond to death insurance coverage and saving contracts. These allow for a tax-advantaged investment component for those anticipating their succession and or suitable for estate planning, which typically attract more elderly.

In the sequel, the baseline mortality φ_x^{ref} is a market table, denoted *IA2013*. The latter is derived from mortality trends originating from the INSEE table, French national bureau of statistics, constructed for the French insurance market provided by *Institute des Actuaires*, see Tomas (2011). Before proceeding to the implementation of the methodology developed in the previous sections,

Table 2.1 – Observed characteristics of portfolios population.

	Period of observation		**Mean age**		**Average exposure**	**Mean age at death**	**Size**
	Beginning	**End**	**In**	**Out**			
P1	01/01/07	12/31/11	36.96	39.74	2.77	68.78	616390
P2	01/01/07	12/31/11	69.3	73.35	4.05	80.34	7589
P3	01/01/07	12/31/10	40.16	43.1	2.94	71.77	80086
P4	01/01/07	12/31/11	37.5	41.13	3.63	54.08	93165
P5	01/01/07	12/31/11	36.9	39.1	2.2	59.31	21540
P6	01/01/07	12/31/10	48.5	52.11	3.62	82.34	847469
P7	01/01/07	12/31/11	66.65	71.29	4.64	73.68	89507
P8	01/01/07	04/13/11	67.51	71.38	3.86	80.72	78650
P9	01/01/07	06/30/11	45.97	49.6	3.62	73.17	1556150
P10	01/01/07	12/31/11	62.97	67.64	4.67	79.77	132990
P11	01/01/07	12/31/11	38.89	42	3.11	56.44	420405
P12	01/01/07	12/31/11	37.05	39.2	2.15	57.41	904020
P13	01/01/07	12/31/11	43.01	46.89	3.88	71.03	848757
P14	01/01/07	12/31/11	50.12	54.16	4.04	72.37	233488

we must look deeper into the particular feature of our dataset. Specifically, we must focus on those that may arise specific concerns when it comes to the graduation of a mortality table using the smoothing procedure considered in section 2.2. As previously reported, the experienced mortality does not only suffer from a small sample size but also the under-representation of those within some age groups. This is typically the case of portfolio P2, see Tableau 2.1 and Figure 2.1. In fact, we have a small sample of 7589 individuals with only 2% aged under 60. This is also the case for portfolios P7, P8 and to some extent P10, but with a larger exposure. For these portfolios, the use of the smoothing procedure in sous-section 2.2.2 has the advantage of borrowing the information in age bands where the exposure is substantially larger. This may allow for mortality curve to fulfill some required local properties such as, e.g., smoothness. In fact, enlarging the smoothing window h, giving access to far distant ages, may ensure the increasing of mortality intensity over ages, which is not only a very much sought behavior but also a biologically reasonable quality.

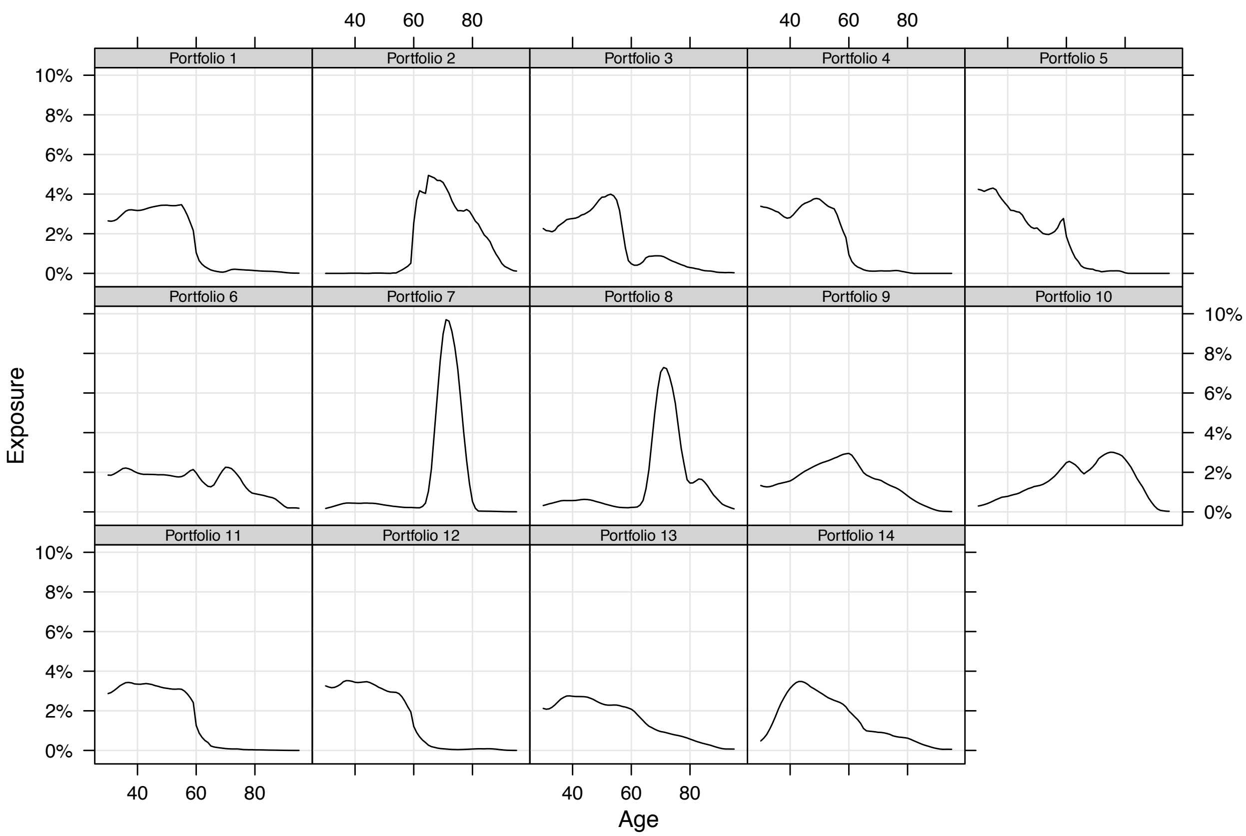

FIGURE 2.1 – Distribution of age groups in the portfolios.

2.4.2 Entity-Specific Graduated Mortality

In order to implement the local likelihood based graduation approach in section 2.2, we need to identify the fitting variables. There are several components of the local fit that must be specified : the bandwidth h, the degree of local polynomial p and the weight function. (i) The latter is assumed be Gaussian kernel as stipulated earlier in this chapter. Other types of kernels can be investigated but this has much less effect on the bias and the variance tradeoff. As noted by Loader (2006), the kernel choice only influences the visual quality of the fitted regression curve. (ii) On the other hand, the bandwidth has a critical effect on the local regression fit. The simplest specification is a constant bandwidth for all ages x. This is, however, not satisfactory in our case. In fact, as mentioned throughout the chapter, for ages where data is available in a sufficient amount small bandwidth will produce a convenient fit with the desired features. In turn, when the population is poorly represented at some ages, large values of h should be needed. Accordingly, one might choose a different bandwidth for each fitting age $x \in \{x_1, \cdots, x_m\}$, taking into account local features such as the local intensity and the amount of data. The problem of choosing the bandwidth h received a lot of attention in the literature. See for example, Fan and Gijbels (1995); Jones et al. (1996); Bagkavos and Patil (2009); Nielsen et al. (2009) and Gámiz et al. (2016) and the references therein.

First of all, when the bandwidth h does not depend on the age level, we can use a scoring procedure based on a generalization of Aïkake Information Criterion (AIC) that uses the deviance function, i.e. the likelihood together with the degrees of freedom of the fitted model, to rank the models. In our case, as we adopt a local rather than a global bandwidth, we advocate using some popular and yet efficient data-driven approaches. Here, we use the selection rule proposed by Jiang and Doksum (2003). The latter can be summarized in the following steps :

Step 1 We choose an initial global bandwidth h. The latter can be based on a modified AIC as described above and advocated by Loader (2006). This is, for instance, the approach used in the empirical work of Tomas (2011). Then, pilot estimators $\widehat{\varphi_x}$ of φ_x are obtained by using the same bandwidth h for ages x and the local likelihood estimator in Équation 2.4.

Step 2 For each age level x, we optimize the likelihood functional in Équation 2.4 being function of the bandwidth. We obtain its minimizer h.

Step 3 We run a local smoother of the bandwidths h over ages using the global bandwidth in **Step 1** and the same kernel ω.

The above rule is the analogue of least-squares cross-validation or the leave-one-out principle, see Mammen et al. (2011), Pérez et al. (2013) and Gámiz et al. (2016). In Gámiz et al. (2016) a precise connection of the cross-validation procedure with our discrete framework is investigated.

Once an estimate of the local bandwidths are obtained, one can estimate the optimal polynomial degree p through the global partial likelihood. In Tableau 2.2 we reported the degree of the polynomial used for smoothing as well as the corresponding degree of freedom and the AIC score. This is intended to represent the global sparsity of the data and the goodness of fit quality. We can see that for some portfolios the optimal choice of the degree controls induce a high level of degrees of freedom, i.e. portfolios P5 and P6. This is to say that the corresponding "smoothed" curves $\widehat{\boldsymbol{\varphi}}^i, i = 5, 6$, will be noisy showing many feature. Indeed, the degree of freedom is a qualitative proxy for the regularity of graduated mortality curve as the *smoothness* evolves inversely to the degree of freedom. This feature can already be deduced from the limited amount of information (exposures) that are at our disposable for these portfolios. However, the sparsity of the data is not only represented by the exposure. In fact, the deaths are of paramount importance in characterizing the survival rate. Indeed, looking at the exposure reported in Tableau 2.1, one could expect a high degree of freedom for the portfolio P2 having only 7589 individuals exposed to risk. However, the death records are concentrated on a small band making the smoothing less noisy as the information needed to estimate the mortality at each individual age is accessible at the immediate adjacent ages. This is, for instance, not the case for the portfolio P5 having a greater exposed individuals but with high sparsity and small number of deaths over few ages. For the remaining portfolios, the degrees of freedom are relatively small. In the following, we will implement the credibility approach described in the last section to assess the impact of the latter on the graduated mortality curves.

TABLE 2.2 – Local-likelihood smoothing parameters' optimal choice

	P1	P2	P3	P4	P5	P6	P7
AIC	57.12	49.34	79.68	78.91	61.10	106.73	68.37
Degree	2	2	1	1	3	3	1
DF	5	3	4	6	16	10	8
	P8	**P9**	**P10**	**P11**	**P12**	**P13**	**P14**
AIC	74.11	63.96	53.44	70.86	74.82	82.70	78.37
Degree	2	1	2	1	1	3	2
DF	4	5	4	6	6	7	5

2.4.3 Next-Period Mortality Rate

Here, we consider the mortality experience over the period $2007-2010$ upon which we calibrate the smoothing procedure considered in the above subsection. For each portfolio, we build a graduated mortality table $\widehat{\varphi}^i$ and aim at adjusting the latter for the next period projection. For each age x, the graduated mortality gives a candidate rate for the next period, i.e. $\widehat{\varphi}^i_x$. The insurer has the possibility of whether to rely on this rate or adjust it given the experience stemming from the other rates at other ages. In other words, the mortality used for the next period forecasts can be adjusted using the credibility formula in Équation 2.18. To do this, we estimate the different quantities needed to implement (2.18) as follows :

(i) Following Lemma 2.3.2, the expected mortality rate α can be estimated as follows :

$$\widehat{\alpha} = \Big(\sum_{x=x_1}^{x_n}\sum_{i=1}^{K} E^i_x \widehat{\varphi}^i_x\Big) / \sum_{i=1}^{K} E^i_{\bullet}.$$

(ii) The weight loading matrix S^i is given as an output of the graduation step and can be estimated using Équation 2.10.

(iii) The diagonal matrix σ relies on the variance of Θ^i_x's which might be estimated given that

$\widehat{\Theta}_x^i = \widehat{\varphi}_x^i/\widehat{\alpha}$, and thus we can write

$$\sigma_x = \sum_{i=1}^{K} \left(\widehat{\varphi}_x^i/\widehat{\alpha}\right)^2/K - \left(\left(\sum_{i=1}^{K} \widehat{\varphi}_x^i/\widehat{\alpha}\right)/K\right)^2.$$

Figure 2.2 depicts, respectively, the graduated mortality over the period 2007−2010 as described above and the next-period (2011) mortality rates using the credibility formula in Équation 2.18. Similarly, Figure 2.3 represents the next-period predicted deaths using the two mortality rates. In these figures we grayed areas (ages) where the relative difference between the smoothed mortality and its adjusted counterpart exceeds a 10% level. More precisely, this corresponds to the ages x where $|\widetilde{\varphi}_x^i - \widehat{\varphi}_x^i|/\widehat{\varphi}_x^i > 0.1$. At a first glance, we remark that the credibility adjustment does change the mortality rate and overall propose a smoother curve compared to the initial one, and this is even evident when dealing with portfolios with small sizes and high degrees of freedom. In fact, when we deal with portfolios such as P5, where the exposure-to-risk as well as the underlying deaths are very limited, the smoothing approach fails to capture the mortality structure and the output of the procedure proposed in section 2.2 are very irregular and noisy. Indeed, as noted above such a procedure need information stemming from adjacent ages when a particular age lacks of sufficient exposure. The case of P5 is very appealing of the limit of the semi-parametric smoothing techniques as the limitation on the information is shared over the ages. This is why the corresponding degree of freedom is high and the AIC is low, see Tableau 2.2, and explains the irregular curve (dashed line) for the smoothed mortality.

On the other hand, the degrees of freedom given as $\text{tr}(\mathbf{S}^i\mathbf{S}^{i\top})$ provide information on the credibility of the smoothed curve $\widehat{\boldsymbol{\varphi}}^i$. In fact, as we can see in Équation 2.18 or in a more tractable way as in Lemma 2.3.2, the higher the degrees of freedom, i.e. $\text{tr}(\mathbf{S}^i\mathbf{S}^{i\top}) = \sum_{x=x_1}^{x_n} \|\mathbf{s}^i(x)\|^2$; the smallest is the weight attached to the smoothed curve $\boldsymbol{\varphi}^i$ (in aggregate). At an age level $x \in \{x_1, \cdots, x_m\}$, one should look at the individual variance $\|\mathbf{s}^i(x)\|^2$. That being said, we can conclude that the parameter driving the adjustment at the age levels is vanishing meaning that the adjusted mortality rate $\widetilde{\varphi}_x^i$ is close to the reference α_x. It comes as no surprise, then, to find that the adjusted curve tends to offset this undesired effect thanks also to the information coming from other ages but from different portfolios.

The visual inspection of the credibility based mortality curve shows that the regularity is preserved avoiding the limitation of the sole smoothing procedure discussed above. For some portfolios, such as portfolio P12, the regularization based on the credibility attached to each age level enhance the prediction of the future mortality. Indeed, the smoothed mortality based on past observations suggest a local distortion of the curve for ages ranging from 60 to 80. This particular feature is however not observed in the mortality curve for the year 2011 and thus the credibility based curve has a better fitting. This can also be observed in Figure 2.3, where the predicted deaths using $\widetilde{\boldsymbol{\varphi}}^i$, for $i = 12$, are (visually) more in line with the observations. The same conclusions, in the grayed area, can be drawn for the other portfolios.

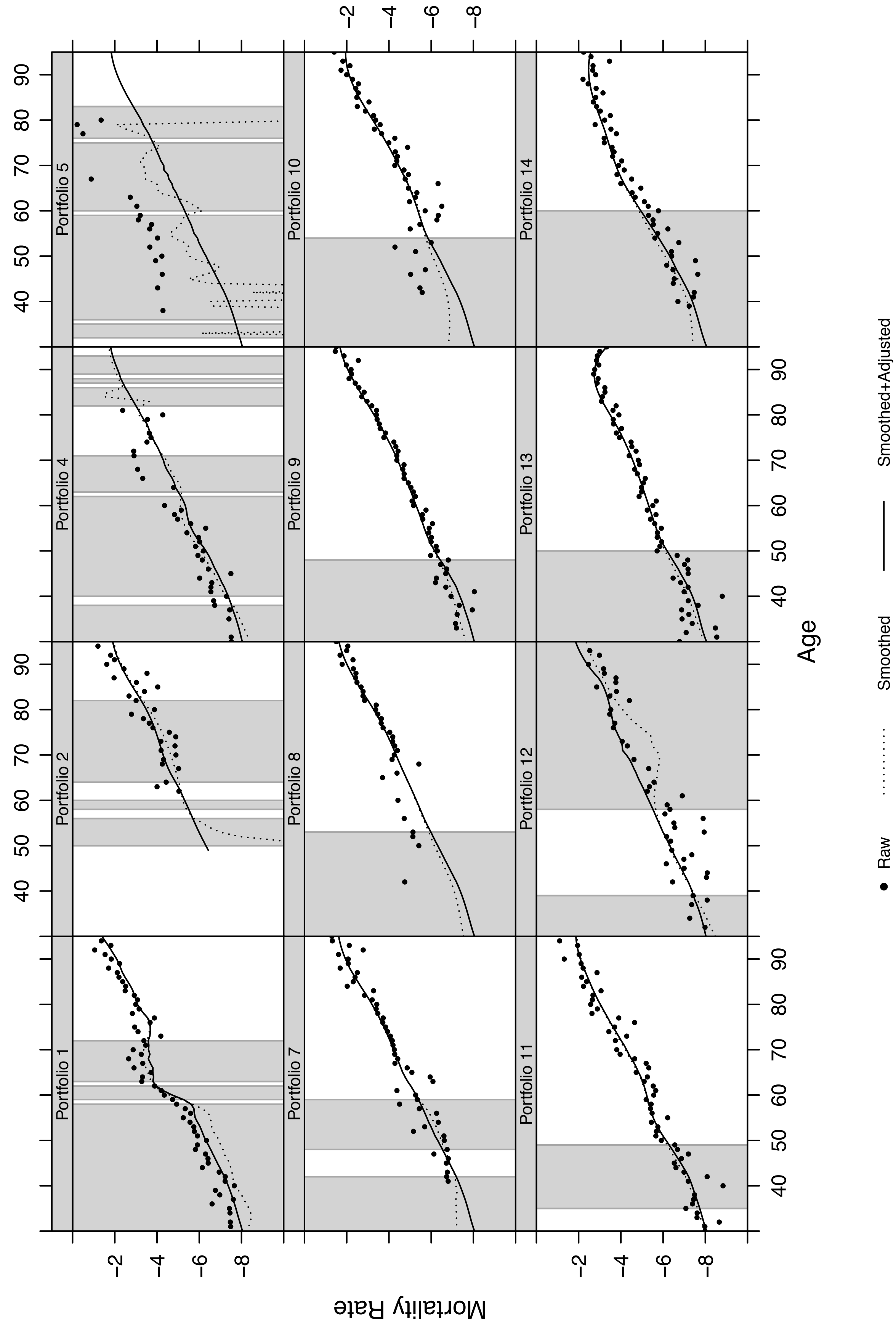

FIGURE 2.2 – The next-period mortality rate based on the graduation of mortality (dashed line) as well as its credibility based adjustment (solid line) based on the period 2007 − 2010. Both predictions are compared to the observed mortality rate over the period 2011 (black dots).

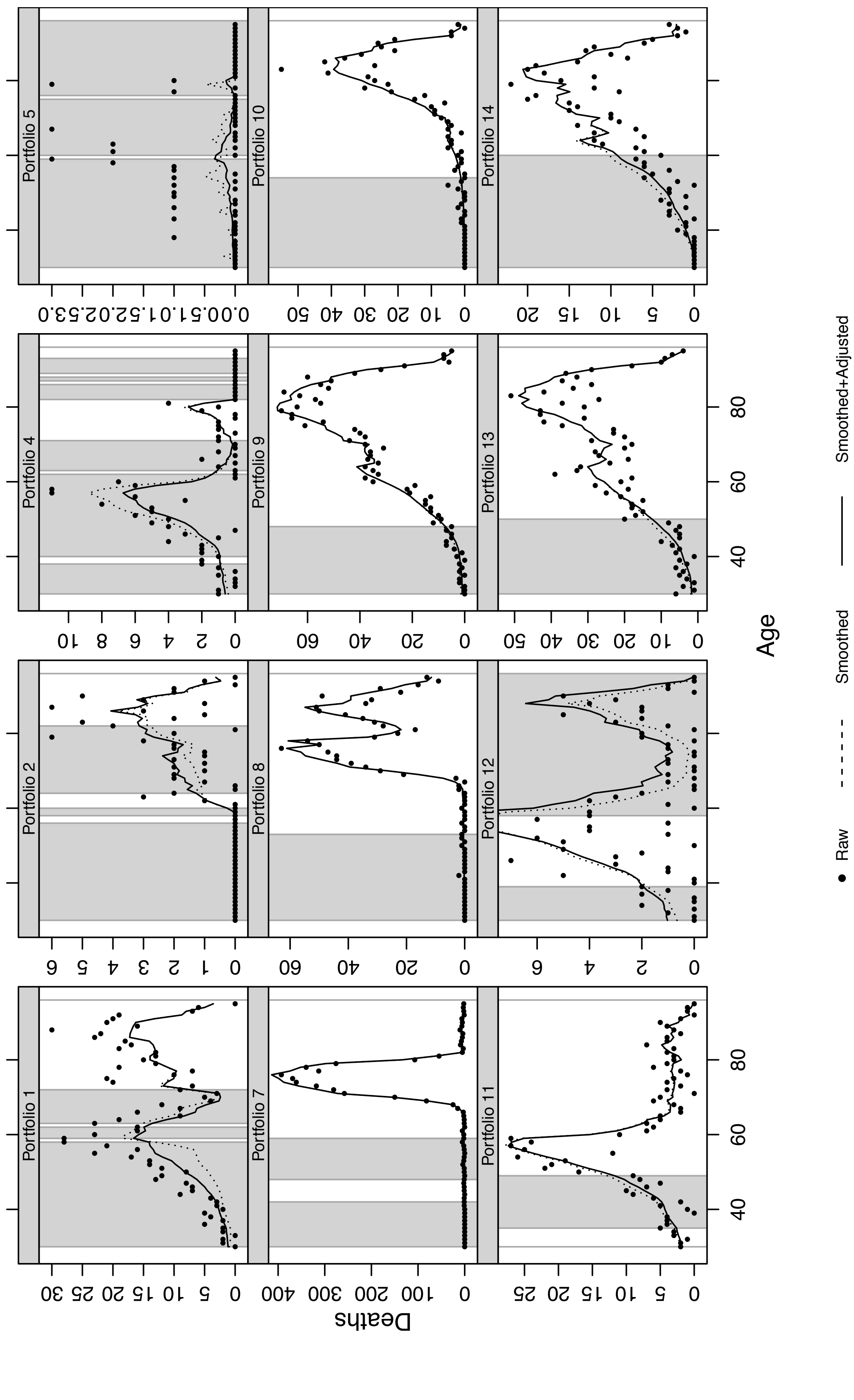

FIGURE 2.3 – The next-period deaths prediction based on the graduation of mortality (dashed line) as well as its credibility based adjustment (solid line) based on the period 2007 − 2010. Both predictions are compared to the observed deaths over the year 2011 (black dots).

2.4.4 Proximity Between the Observations and the Model

Besides the visual inspection of the proposed adjustments and in order to understand the impact of the latter, we will use some known statistics to quantify the proximity between the observations and the outputs of the two curves considered in Figure 2.2 and Figure 2.3. We assess the overall deviation with the observed mortality by comparing criteria measuring the distance between the observations and the models with the χ^2, i.e. Forfar et al. (1988b), the mean average percentage error (MAPE), see for instance Felipe et al. (2002) ; as well as the standardized mortality ratio (SMR), i.e. the ratio of deaths observed to those predicted. The quantities summarizing the proximity between the observations and the model, for each portfolio i at calendar year $t = 2011$, are described as follows :

(i) The χ_i^2 allows to measure the quality of the fit of the model. It writes,

$$\chi_i^2 = \sum_{x=x_1}^{x_n} \frac{\left(D_x^i - E_x^i\,\widehat{q}_x^i\right)^2}{E_x^i\,\widehat{q}_x^i\left(1-\widehat{q}_x^i\right)}.$$

(ii) The MAPE is the average of the absolute values of the deviations from the observations,

$$\text{MAPE}^i = \frac{\sum_{x=x_1}^{x_n}\left|\left(D_x^i/E_x^i - \widehat{q}_x^i\right)/\left(D_x^i/E_x^i\right)\right|}{\sum_{x=x_1}^{x_n} D_x^i} \times 100.$$

(iii) The SMR is computed as the ratio between the observed and fitted number of deaths in each portfolio

$$\text{SMR}^i = \frac{\sum_{x=x_1}^{x_n} D_x^i}{\sum_{x=x_1}^{x_n} E_x^i\,\widehat{q}_x^i}.$$

Hence, if $\text{SMR}^i > 1$, the fitted deaths are under-estimated and vice-versa if $\text{SMR} < 1$. Note that we can consider the SMR_i as a global criterion which does not take the age structure into account, compared to the χ_i^2 and the MAPE_i for instance.

Tableau 2.3 and Tableau 2.4 summarize the above mentioned quantities giving the overall deviation between the observations and the adjustment analysis for the portfolios P1 to P14 (except 3 and 6 which do not contain observations of year 2011) obtained by the smoothing approach together with the credibility adjustment procedure. When looking at criteria and quantities which take the age structure of the error into account, the credibility approach has an important benefit compared to the sole graduated curve. The quality of the fit increases, sometimes drastically, i.e. portfolio P1, in terms of having the minimum χ_i^2 and MAPE_i values, i.e. the last panels of Tableau 2.3. Also, the credibility adjustment exhibits the highest p-value for the likelihood ratio test. Even when we consider a global indicator of the quality of the fit such as the SMR_i which does not take the age structure into account, the proposed procedure seems to perform better than

Table 2.3 – Tests and quantities summarizing the deviation between the observations and the model

		Hardy-Panjer	Poisson-Gamma	Makeham-Credibility	Smoothed	Smoothed+Adj.
χ^2	Portfolio 1	1901.240	1928.680	259.400	357.870	193.967
MAPE (%)		102.660	102.000	32.870	3.018	2.349
SMR		1.737	1.756	1.126	1.487	1.385
χ^2	Portfolio 2	34.890	33.640	30.940	37.612	31.166
MAPE (%)		48.030	49.120	53.990	20.119	20.842
SMR		1.037	1.002	0.905	1.102	0.948
χ^2	Portfolio 4	130.120	132.890	79.321	58.615	51.515
MAPE (%)		95.390	92.490	44.880	14.006	13.078
SMR		0.826	0.853	1.405	0.984	1.168
χ^2	Portfolio 5	473.680	573.940	348.180	NA	370.401
MAPE (%)		85.660	88.040	90.420	59.296	56.038
SMR		2.857	3.424	5.021	3.513	5.534
χ^2	Portfolio 7	221.640	223.560	195.000	77.997	72.795
MAPE (%)		135.390	135.710	37.250	0.534	0.509
SMR		0.846	0.844	0.823	0.922	0.922
χ^2	Portfolio 8	2575.630	2583.900	2414.250	66.033	61.174
MAPE (%)		323.780	324.610	263.210	1.100	1.122
SMR		0.232	0.231	0.243	0.928	0.930

TABLE 2.4 – Tests and quantities summarizing the deviation between the observations and the model - continued

		Hardy-Panjer	Poisson-Gamma	Makeham-Credibility	Smoothed	Smoothed+Adj.
χ^2	Portfolio 9	1572.530	1573.970	1502.870	57.461	53.735
MAPE (%)		368.080	368.290	125.640	0.764	0.755
SMR		0.423	0.423	0.419	0.932	0.940
χ^2	Portfolio 10	115.820	116.470	97.880	83.790	72.448
MAPE (%)		89.680	91.030	46.140	3.356	3.530
SMR		0.871	0.862	0.960	0.948	0.950
χ^2	Portfolio 11	415.320	417.530	76.480	55.888	55.127
MAPE (%)		152.870	151.690	46.970	5.934	5.548
SMR		0.829	0.837	1.018	0.918	0.964
χ^2	Portfolio 12	130.050	129.230	90.740	88.836	76.459
MAPE (%)		110.540	107.220	95.270	36.577	33.344
SMR		0.598	0.619	0.543	0.669	0.539
χ^2	Portfolio 13	351.560	351.360	263.550	94.570	89.428
MAPE (%)		180.910	180.610	54.620	1.765	1.608
SMR		0.839	0.840	0.832	0.914	0.930
χ^2	Portfolio 14	227.860	227.950	85.920	59.317	50.885
MAPE (%)		159.740	160.600	53.530	5.659	4.852
SMR		0.792	0.788	0.939	0.827	0.860

the graduated curve. However, notice that the impact of adjustment in smaller when the portfolios are quiet big. This is already remarkable when one operated visual checks as mentioned earlier.

2.4.5 Comparaison with Classical Approaches

We first note that the Hardy-Panjer and Poisson-Gamma approaches produce relatively similar graduations as the tests suggest sensibly similar outputs. However, we notice some differences with the Makeham credibility model which displays more favorable results. This is already outlined in Salhi et al. (2016) and this may be explained by the age-specific adjustment but also thanks to the structural feature added by the Makeham parametric model. Taking into account the age and thus the structure of the portfolio increases the goodness of fit of as well as the predictive performance of the constructed mortality. Regarding the local likelihood approach we notice that the force of mortality adjust to more complex mortality structures and thus offer a better fit for portfolio with sparse information. However, for very small portfolios the smoothed mortality rates fail to properly predict the next-year deaths compared to the aforementioned approaches. The credibility-based revision at the age-level globally enhanced the predictive ability of the graduated mortality. Specifically, when it comes to the tests that are sensitive to the age structure, we notice that the credibility-based adjustment offers an outstanding fit as the tests are favorable compared to the Hardy-Panjer, Poisson-Gamma and the Makeham-based approaches. More importantly, the main advantage of our method over these approaches is its ability to adjust to more complex mortality structures and thus offer a better fit for portfolio with sparse information. e.g. the next-year MAPE, for Portfolio 2 (small), goes from 3.10% to 2.34%.

2.5 Concluding Remarks

In this chapter we proposed a methodology to adjust the graduated mortality table that uses an adaptive smoothing procedure based on the local likelihood. The adjustment is based on the credibility weighting technique of the smoothed curves and a reference. Out approach takes into account the age specific heterogeneity that may arise in real world datasets. We thus consider updating the mortality for each age based on the upcoming past information from the same age but also the neighboring ages. The inclusion of the neighboring ages is crucial as the particular smoothing procedure used in this chapter add a dependency between the single ages. Based on classical results on the inference of the smoothing procedure we derived the closed form formulas needed to adjust the mortality.

The proposed methodology is shown to outperform compared to the classical credibility approaches that does not take into account the age structure of the portfolio. This is in line with the recent work in this field as mentioned by Salhi et al. (2016). Even when the age structure is accounted for the methodology developed in this chapter has an important benefit. This is mainly due to the underlying curve built using an adaptive procedure compared to the parametric model considered in Salhi et al. (2016).

We should note that the model proposed can be investigated in order to quantify mathematically the errors induced in the assessment of the next-period mortality curve. This amend to consider the uncertainty stemming from the estimation of the different variables used in the updating procedure. There are also several practical issues we do not address here such as the impact on the pricing of life insurance contracts. These are open questions that we openly acknowledge and leave for future research.

Conclusion

Dans ce mémoire nous avons proposé deux méthodologies de construction de table de mortalité pour des portefeuilles de petite taille. Ces deux méthodologies reposent sur une application de la théorie de la crédibilité tenant en compte la fiabilité des données mais aussi leur quantité. Etant de natures différentes, l'une se basant sur une approche paramétrique et l'autre une construction semi-paramétrique, l'inclusion de la crédibilité réussi à améliorer le pouvoir prédictif des tables construites par les deux méthodes. En effet, la procédure qu'on propose emprunte l'information nécessaire à la construction des probabilités de décès des autres portefeuilles qui ont une information plus abondante mais aussi plus fiable.

Comme nous l'avons constaté dans les deux chapitres 1 et 2 la prise en compte de l'information provenant des différents portefeuilles permet de pallier au manque d'observation pour d'autres portefeuilles. Par conséquent, un modèle paramétrique permet alors un ajustement plus fidèle aux données brutes. Il s'avère, néanmoins, que l'utilisation d'un modèle semi-paramétrique (de type lissage par maximisation de la vraisemblance locale) offre ajustement plus proche des décès observés et offre une meilleure prédiction selon la conclusion du chapitre 2 basée sur une évaluation quantitative de la pertinence des deux approches. En effet, les divergences entre ces deux méthodes sont appréhendées sur la fiabilité de l'ajustement de la mortalité passée mais aussi sur l'erreur de prédiction de la mortalité future. La méthode de lissage utilisant la vraisemblance locale complétée d'un ajustement par crédibilité réussit à produire des estimations futures de décès plus proche de la réalité. Il est donc clair que cette approche est la mieux adaptée pour répondre à la problématique initialement posée qui est la construction d'une table *best estimate* pour le provision du risque de mortalité à la maille portefeuille ou contrat.

Enfin, la méthodologie proposée dans ce mémoire peut être étendue à d'autres risques comme liés au maintien en arrêt de travail, à la longévité... etc. Néanmoins, l'application à la longévité nécessite une réadaptation de la théorie pour prendre en compte l'aspect temporel de ce risque lié principalement au risque de tendance.

Annexes

Annexe A

Annexe du chapitre 1

A.1 Tests and quantities summarizing the deviation between the observations and the models for the year 2010

Tables A.1 and A.2 present the tests and quantities summarizing the overall deviation between the observations and the credibility analysis for the male population obtained by the Hardy-Panjer, Poisson-Gamma and the Makeham credibility approaches with the two baselines mortality considered for the year 2010.

A.2 Tests and quantities summarizing the deviation between the observations and the models for the year 2011

Tables A.3 and A.4 present the tests and quantities summarizing the overall deviation between the observations and the credibility analysis for the male population obtained by the Hardy-Panjer, Poisson-Gamma and the Makeham credibility approaches with the two baselines mortality considered for the year 2011.

A.3 Fitted probabilities of death in the log scale for the year 2010

Figure A.1 displays the fitted probabilities of death in the log scale for the male population for the year 2010.

A.4 Fitted probabilities of death in the log scale for the year 2011

Figure A.2 displays the fitted probabilities of death in the log scale for the male population for the year 2011.

A.5 Fitted number of deaths for the year 2010

Figure A.3 displays the fitted number of deaths for the male population for the year 2010.

A.6 Fitted number of deaths for the year 2010

Figure A.4 displays the fitted number of deaths for the male population for the year 2010.

A.7 Standardized residuals for the year 2010

Figure A.5 displays the standardized residuals for the male population for the year 2010.

A.8 Standardized residuals for the year 2011

Figure A.6 displays the standardized residuals for the male population for the year 2011.

Table A.1 – Tests and quantities summarizing the deviation between the observations and the model, calendar year 2010, male population.

			INSEE			IA2103		
			Hardy-Panjer	Poisson-Gamma	Makeham-Credibility	Hardy-Panjer	Poisson-Gamma	Makeham-Credibility
Portfolio 1	Standardized	> 2	60	60	35	46	46	15
	residuals	> 3	48	48	28	32	32	5
	χ^2		5481.86	5542.82	3569.97	1705.25	1747.25	208.81
	MAPE (%)		233.22	230.94	373.89	117.01	115.42	42.35
	Likelihood	ξ^{LR}	946.98	947.72	443.16	463.48	468.46	88.03
	ratio test	p-value	0	0	0	0	0	0.0364
		SMR	1.1792	1.1919	0.5265	1.7629	1.7957	1.0532
	SMR test	ξ^{SMR}	4.0379	4.2939	12.1893	13.0352	13.4202	1.2845
		p-value	0	0	0	0	0	0.0995
Portfolio 2	Standardized	> 2	9	11	0	1	1	0
	residuals	> 3	2	1	0	0	0	0
	χ^2		102.84	101.50	29.62	41.54	40.41	30.75
	MAPE (%)		108.16	116.37	48.80	47.18	48.09	54.70
	Likelihood	ξ^{LR}	90.3	94.99	33.8	36.43	36.77	33.35
	ratio test	p-value	3e-04	1e-04	0.9517	0.908	0.901	0.9573
		SMR	0.6421	0.6014	0.8764	1.0149	0.9868	0.8567
	SMR test	ξ^{SMR}	3.6844	4.2907	0.9805	0.074	0.0207	1.1681
		p-value	1e-04	0	0.1634	0.4705	0.4918	0.1214
Portfolio 3	Standardized	> 2	34	32	11	7	7	4
	residuals	> 3	9	9	5	0	0	0
	χ^2		416.19	420.04	161.66	110.28	110.89	64.16
	MAPE (%)		156.33	154.14	76.78	64.99	64.67	45.48
	Likelihood	ξ^{LR}	239.44	236.76	115.84	91.13	90.85	38.51
	ratio test	p-value	0	0	1e-04	0.0219	0.023	0.9973
		SMR	0.5361	0.5451	0.8955	0.8989	0.9052	1.1212
	SMR test	ξ^{SMR}	7.0465	6.8379	1.0892	1.049	0.9746	1.1174
		p-value	0	0	0.138	0.1471	0.1649	0.1319
Portfolio 4	Standardized	> 2	20	19	15	8	5	2
	residuals	> 3	2	1	3	0	0	0
	χ^2		183.96	181.13	199.86	83.98	83.32	41.51
	MAPE (%)		201.49	196.70	189.51	92.75	90.01	44.33
	Likelihood	ξ^{LR}	212.87	208	174.37	101.22	98.75	36.28
	ratio test	p-value	0	0	0	0	1e-04	0.9406
		SMR	0.3590	0.3665	0.4332	0.6161	0.6326	1.0677
	SMR test	ξ^{SMR}	11.537	11.2597	8.408	4.9251	4.6315	0.5798
		p-value	0	0	0	0	0	0.2810
Portfolio 5	Standardized	> 2	8	9	8	8	10	13
	residuals	> 3	8	8	7	8	8	6
	χ^2		368.00	470.94	205.33	259.26	366.90	209.05
	MAPE (%)		72.14	78.00	67.45	79.85	85.30	82.04
	Likelihood	ξ^{LR}	63.85	63.4	59.53	52.94	55.85	43.15
	ratio test	p-value	0.1069	0.1141	0.1930	0.3992	0.2977	0.7746
		SMR	1.4167	1.7941	1.2442	2.1557	2.9553	3.1797
	SMR test	ξ^{SMR}	1.6446	2.6956	1.0308	3.4572	4.6617	4.9234
		p-value	0.05	0.0035	0.1513	3e-04	0	0
Portfolio 6	Standardized	> 2	62	62	56	50	50	24
	residuals	> 3	61	61	50	44	44	7
	χ^2		7615.50	7615.40	1538.42	1364.75	1364.60	256.21
	MAPE (%)		2558.24	2558.59	652.41	631.01	631.13	145.45
	Likelihood	ξ^{LR}	7417.14	7417.88	1707.66	1575.04	1575.17	272.40
	ratio test	p-value	0	0	0	0	0	0
		SMR	0.5444	0.5443	0.9802	0.9337	0.9335	0.9829
	SMR test	ξ^{SMR}	42.5496	42.56	1.2532	4.3697	4.3813	1.0796
		p-value	0	0	0.1051	0	0	0.1402

Table A.2 – Tests and quantities summarizing the deviation between the observations and the model, calendar year 2010, male population.

			INSEE			IA2103		
			Hardy-Panjer	Poisson-Gamma	Makeham-Credibility	Hardy-Panjer	Poisson-Gamma	Makeham-Credibility
Portfolio 8	Standardized	> 2	62	62	5	26	25	4
	residuals	> 3	59	59	1	6	6	0
	χ^2		2962.00	2967.84	115.90	274.12	275.25	85.69
	MAPE	(%)	837.99	839.75	67.40	116.81	117.16	23.93
	Likelihood	ξ^{LR}	2455.66	2462.18	130.73	247.92	248.89	55.14
	ratio test	p-value	0	0	0	0	0	0.8273
		SMR	0.5638	0.5627	0.9673	0.8972	0.8953	0.9811
	SMR test	ξ^{SMR}	37.8094	37.9526	1.9899	6.6063	6.7345	1.1323
		p-value	0	0	0.0233	0	0	0.1287
Portfolio 9	Standardized	> 2	63	63	40	45	45	12
	residuals	> 3	59	59	30	38	38	3
	χ^2		5759.31	5759.28	591.79	741.90	742.08	147.99
	MAPE	(%)	754.36	754.31	192.26	198.83	198.97	22.01
	Likelihood	ξ^{LR}	3443.71	3443.52	427.46	502.06	502.38	77.09
	ratio test	p-value	0	0	0	0	0	0.1653
		SMR	0.5262	0.5262	0.9084	0.8627	0.8622	0.9078
	SMR test	ξ^{SMR}	41.6671	41.6629	5.6716	8.7994	8.8355	5.7137
		p-value	0	0	0	0	0	0
Portfolio 10	Standardized	> 2	48	48	1	6	7	3
	residuals	> 3	33	33	0	1	1	1
	χ^2		669.50	672.46	80.63	121.74	122.90	86.65
	MAPE	(%)	504.88	509.44	75.75	110.69	112.55	55.72
	Likelihood	ξ^{LR}	631.46	636.74	82.38	114.75	116.28	48.65
	ratio test	p-value	0	0	0.0839	2e-04	1e-04	0.9461
		SMR	0.5336	0.5292	0.8434	0.8352	0.8263	0.91
	SMR test	ξ^{SMR}	16.4396	16.6765	4.1025	4.344	4.6133	2.2303
		p-value	0	0	0	0	0	0.0129
Portfolio 11	Standardized	> 2	43	43	23	33	33	2
	residuals	> 3	37	37	20	17	17	1
	χ^2		1387.49	1391.13	695.98	380.60	383.26	74.55
	MAPE	(%)	257.02	255.43	464.55	125.94	124.91	46.19
	Likelihood	ξ^{LR}	429.18	426.67	338.9	161.53	161.07	39.42
	ratio test	p-value	0	0	0	0	0	0.9949
		SMR	0.5373	0.5407	0.4887	0.9009	0.9094	1.092
	SMR test	ξ^{SMR}	14.8749	14.7085	14.254	2.2628	2.0519	1.8578
		p-value	0	0	0	0.0118	0.0201	0.0316
Portfolio 12	Standardized	> 2	33	33	25	17	18	4
	residuals	> 3	17	17	20	3	3	0
	χ^2		588.18	592.25	449.62	161.43	164.16	91.65
	MAPE	(%)	241.27	236.71	514.74	111.99	108.62	89.85
	Likelihood	ξ^{LR}	274.67	270.12	329.89	122.89	120.3	96.49
	ratio test	p-value	0	0	0	0	1e-04	0.0085
		SMR	0.4877	0.4971	0.3291	0.7957	0.8243	0.7125
	SMR test	ξ^{SMR}	10.8436	10.5217	15.3598	3.1391	2.6305	4.7701
		p-value	0	0	0	8e-04	0.0043	0
Portfolio 13	Standardized	> 2	55	55	41	27	27	19
	residuals	> 3	44	44	30	16	16	11
	χ^2		2162.97	2162.79	761.52	331.68	331.72	252.49
	MAPE	(%)	478.97	478.32	200.96	136.82	136.60	46.85
	Likelihood	ξ^{LR}	1360.75	1359.18	469.14	241.26	241.02	136.5
	ratio test	p-value	0	0	0	0	0	0
		SMR	0.5378	0.5385	0.9215	0.8966	0.8979	0.8966
	SMR test	ξ^{SMR}	24.9715	24.9134	2.9868	4.0137	3.9601	4.0113
		p-value	0	0	0.0014	0	0	0

Table A.3 – Tests and quantities summarizing the deviation between the observations and the model, calendar year 2011, male population.

			INSEE			IA2103		
			Hardy-Panjer	Poisson-Gamma	Makeham-Credibility	Hardy-Panjer	Poisson-Gamma	Makeham-Credibility
Portfolio 1	Standardized	> 2	58	57	56	48	48	13
	residuals	> 3	52	52	48	35	35	5
	χ^2		5574.22	5621.44	3126.56	1901.24	1928.68	259.40
	MAPE (%)		201.74	200.25	178.37	102.66	102.00	32.87
	Likelihood	ξ^{LR}	1027.03	1027.4	806.13	524.68	528.16	106.43
	ratio test	p-value	0	0	0	0	0	0.0012
		SMR	1.1124	1.1216	1.2011	1.7371	1.7557	1.1256
	SMR test	ξ^{SMR}	2.935	3.1588	4.9944	14.192	14.441	3.2557
		p-value	0.0017	8e-04	0	0	0	6e-04
Portfolio 2	Standardized	> 2	4	4	0	3	2	1
	residuals	> 3	2	1	0	0	0	0
	χ^2		77.89	78.43	29.07	34.89	33.64	30.94
	MAPE (%)		114.12	124.28	52.48	48.03	49.12	53.99
	Likelihood	ξ^{LR}	66.67	71.07	29.92	28.72	28.99	28.16
	ratio test	p-value	0.0385	0.0169	0.9811	0.9877	0.9864	0.9901
		SMR	0.6545	0.6097	0.8668	1.0371	1.0016	0.905
	SMR test	ξ^{SMR}	3.7113	4.3984	1.1388	0.2609	0.0268	0.7649
		p-value	1e-04	0	0.1274	0.3971	0.5107	0.2222
Portfolio 4	Standardized	> 2	20	20	15	13	12	6
	residuals	> 3	4	4	3	4	4	1
	χ^2		250.57	250.57	1026.72	130.12	132.89	79.00
	MAPE (%)		202.16	196.53	226.84	95.39	92.49	44.88
	Likelihood	ξ^{LR}	173.25	168.79	172.37	90.66	89.04	51.08
	ratio test	p-value	0	0	0	7e-04	0.0011	0.51
		SMR	0.4852	0.498	0.5443	0.826	0.8534	1.4047
	SMR test	ξ^{SMR}	8.9106	8.5491	6.3742	2.1049	1.7255	3.4889
		p-value	0	0	0	0.0177	0.0422	2e-04
Portfolio 5	Standardized	> 2	8	8	8	10	12	17
	residuals	> 3	8	8	6	8	8	12
	χ^2		706.87	851.26	262.78	473.68	573.94	348.18
	MAPE (%)		77.15	80.93	77.56	85.66	88.04	90.42
	Likelihood	ξ^{LR}	64.56	65.02	52.53	56.7	58.91	50.61
	ratio test	p-value	0.1133	0.1061	0.4534	0.3041	0.2374	0.5288
		SMR	1.714	2.0544	1.8163	2.857	3.4243	5.0206
	SMR test	ξ^{SMR}	2.4494	3.1986	2.6942	4.4512	5.0828	6.2982
		p-value	0.0072	7e-04	0.0035	0	0	0
Portfolio 7	Standardized	> 2	55	55	15	21	21	11
	residuals	> 3	45	45	8	3	3	9
	χ^2		1593.26	1597.79	236.83	221.64	223.56	195.00
	MAPE (%)		620.28	621.10	95.95	135.39	135.71	37.25
	Likelihood	ξ^{LR}	1448.73	1452.33	201.91	227.8	229.08	118.01
	ratio test	p-value	0	0	0	0	0	1e-04
		SMR	0.5775	0.5768	0.811	0.8455	0.844	0.8229
	SMR test	ξ^{SMR}	35.3923	35.4792	2.74	10.1297	10.2409	11.8209
		p-value	0	0	0	0	0	0
Portfolio 8	Standardized	> 2	65	65	37	50	50	29
	residuals	> 3	63	63	29	29	29	29
	χ^2		4987.77	5002.11	2485.39	2575.63	2583.90	2414.25
	MAPE (%)		788.87	790.77	292.03	323.78	324.61	263.21
	Likelihood	ξ^{LR}	4970.14	4984.29	1891.04	2059.63	2066.49	1765.46
	ratio test	p-value	0	0	0	0	0	0
		SMR	0.1483	0.148	0.2404	0.2315	0.2311	0.2431
	SMR test	ξ^{SMR}	82.2167	82.3412	56.2334	58.1115	58.2101	55.6816
		p-value	0	0	0	0	0	0

Table A.4 – Tests and quantities summarizing the deviation between the observations and the model, calendar year 2011, male population.

			INSEE			IA2103		
			Hardy-Panjer	Poisson-Gamma	Makeham-Credibility	Hardy-Panjer	Poisson-Gamma	Makeham-Credibility
Portfolio 10	Standardized	> 2	50	50	3	5	5	5
	residuals	> 3	34	33	1	1	1	1
	χ^2		635.65	638.50	74.42	115.82	116.47	97.88
	MAPE	(%)	408.56	412.51	51.54	89.68	91.03	46.14
	Likelihood	ξ^{LR}	613.89	619.28	73.18	112.54	113.73	47.99
	ratio test	p-value	0	0	0.2542	3e-04	2e-04	0.9535
		SMR	0.5666	0.5617	0.9229	0.8708	0.8623	0.9596
	SMR test	ξ^{SMR}	15.1709	15.4268	1.9491	3.4132	3.6619	0.9826
		p-value	0	0	0.0256	3e-04	1e-04	0.1629
Portfolio 11	Standardized	> 2	43	43	24	35	35	4
	residuals	> 3	41	41	22	17	19	0
	χ^2		1379.61	1382.88	926.73	415.32	417.53	76.48
	MAPE	(%)	299.83	297.80	555.88	152.87 151.69	46.97	
	Likelihood	ξ^{LR}	511.89	508.51	429.79	214.06	213.07	52.92
	ratio test	p-value	0	0	0	0	0	0.8779
		SMR	0.4927	0.4961	0.4443	0.8291	0.8369	1.0183
	SMR test	ξ^{SMR}	16.7405	16.5554	15.6387	4.0301	3.8212	0.3648
		p-value	0	0	0	0	1e-04	0.3576
Portfolio 12	Standardized	> 2	35	35	21	10	11	4
	residuals	> 3	16	15	16	1	1	0
	χ^2		470.25	471.58	263.73	130.05	129.23	90.74
	MAPE	(%)	231.00	226.18	470.92	110.54	107.22	95.27
	Likelihood	ξ^{LR}	317.04	310.88	337.99	144.41	140.05	114.53
	ratio test	p-value	0	0	0	0	0	2e-04
		SMR	0.3668	0.3745	0.2039	0.5981	0.6188	0.5426
	SMR test	ξ^{SMR}	12.9497	12.6324	17.505	6.0626	5.624	7.3459
		p-value	0	1e-04	0	0	0	0
Portfolio 13	Standardized	> 2	56	56	39	28	28	23
	residuals	> 3	49	50	29	19	19	10
	χ^2		2058.43	2057.75	678.98	351.56	351.36	263.55
	MAPE	(%)	589.24	588.36	245.40	180.91	180.61	54.62
	Likelihood	ξ^{LR}	1316.24	1314.45	414.88	237.69	237.35	141.71
	ratio test	p-value	0	0	0	0	0	0
		SMR	0.5092	0.5099	0.8679	0.8392	0.8404	0.8316
	SMR test	ξ^{SMR}	27.2355	27.1712	5.2064	6.4792	6.4261	6.8303
		p-value	0	0	0	0	0	0
Portfolio 14	Standardized	> 2	48	48	21	24	24	7
	residuals	> 3	36	36	5	5	6	0
	χ^2		862.31	862.27	248.72	227.86	227.95	85.92
	MAPE	(%)	445.95	445.66	135.88	159.74	160.60	53.53
	Likelihood	ξ^{LR}	709.24	708.85	186.3	204.38	205.17	57.2
	ratio test	p-value	0	0	0	0	0	0.7717
		SMR	.5019	0.5022	0.9239	0.7916	0.7879	0.9385
	SMR test	ξ^{SMR}	16.5241	16.5063	1.6821	5.1598	5.2678	1.3381
		p-value	0	0	0.0463	0	0	0.0904

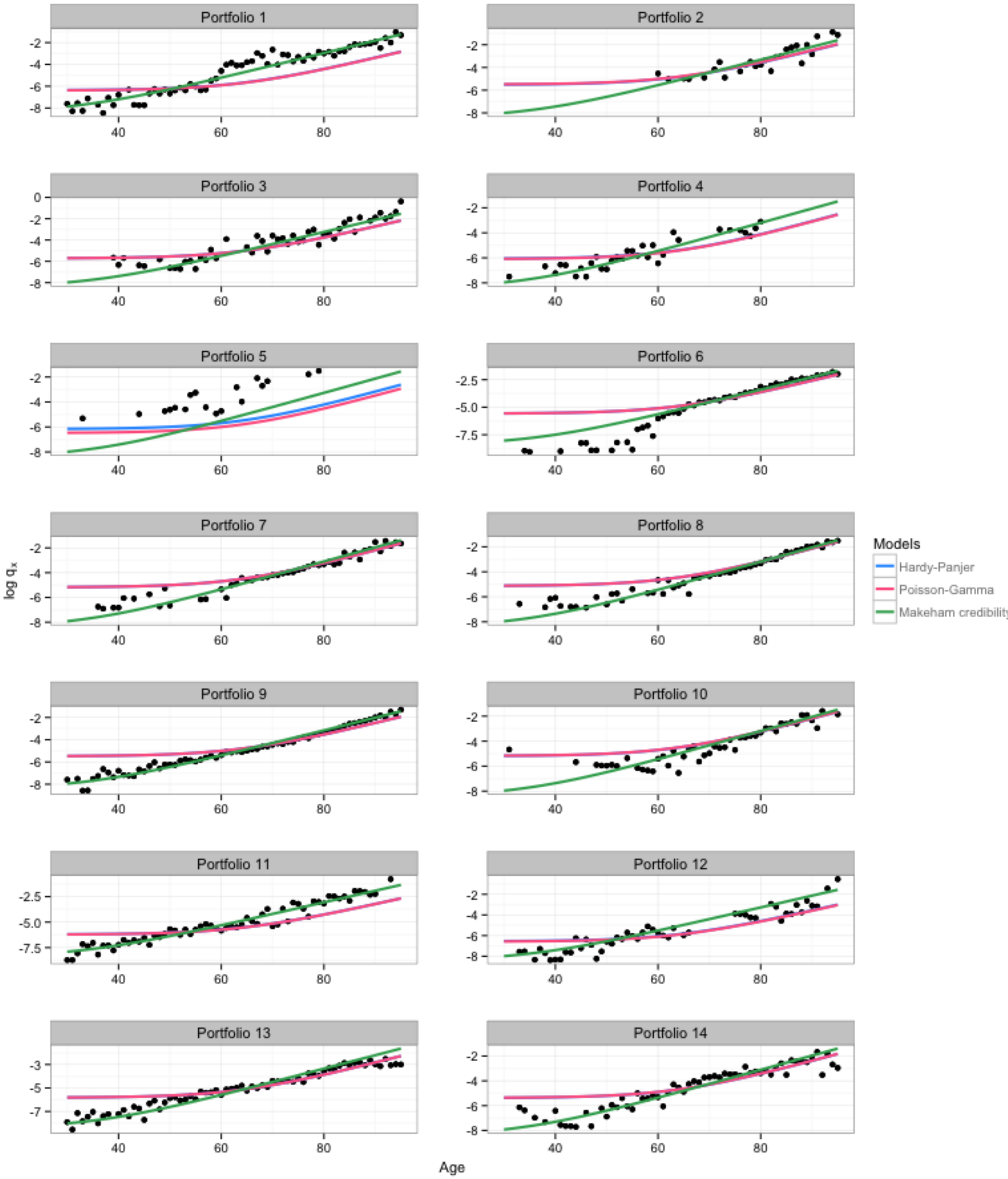

FIGURE A.1 – Fitted probability of death, log scale, for the year 2010, male population

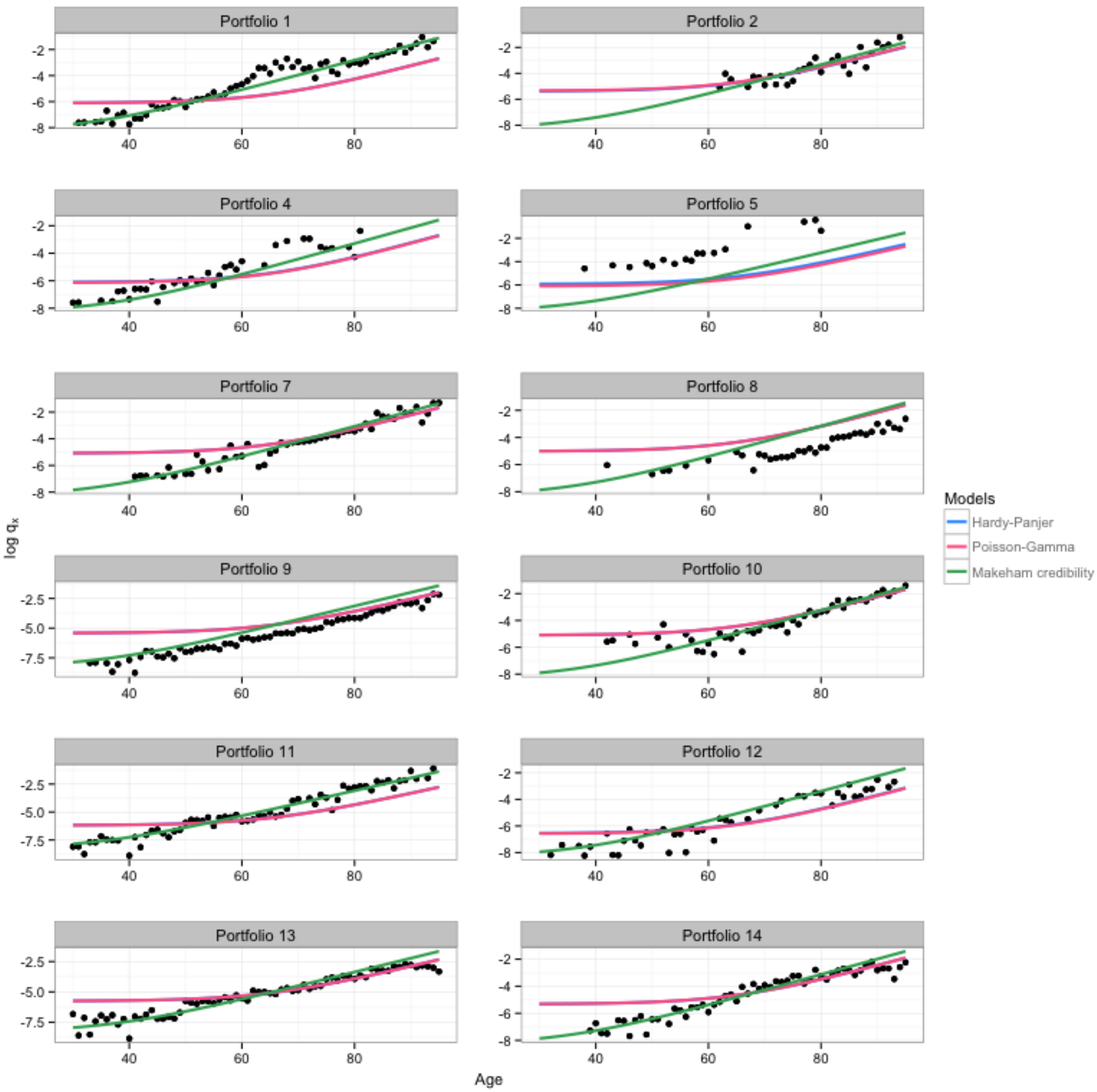

FIGURE A.2 – Fitted probability of death, log scale, for the year 2011, male population

FIGURE A.3 – Fitted number of deaths for the year 2010, male population

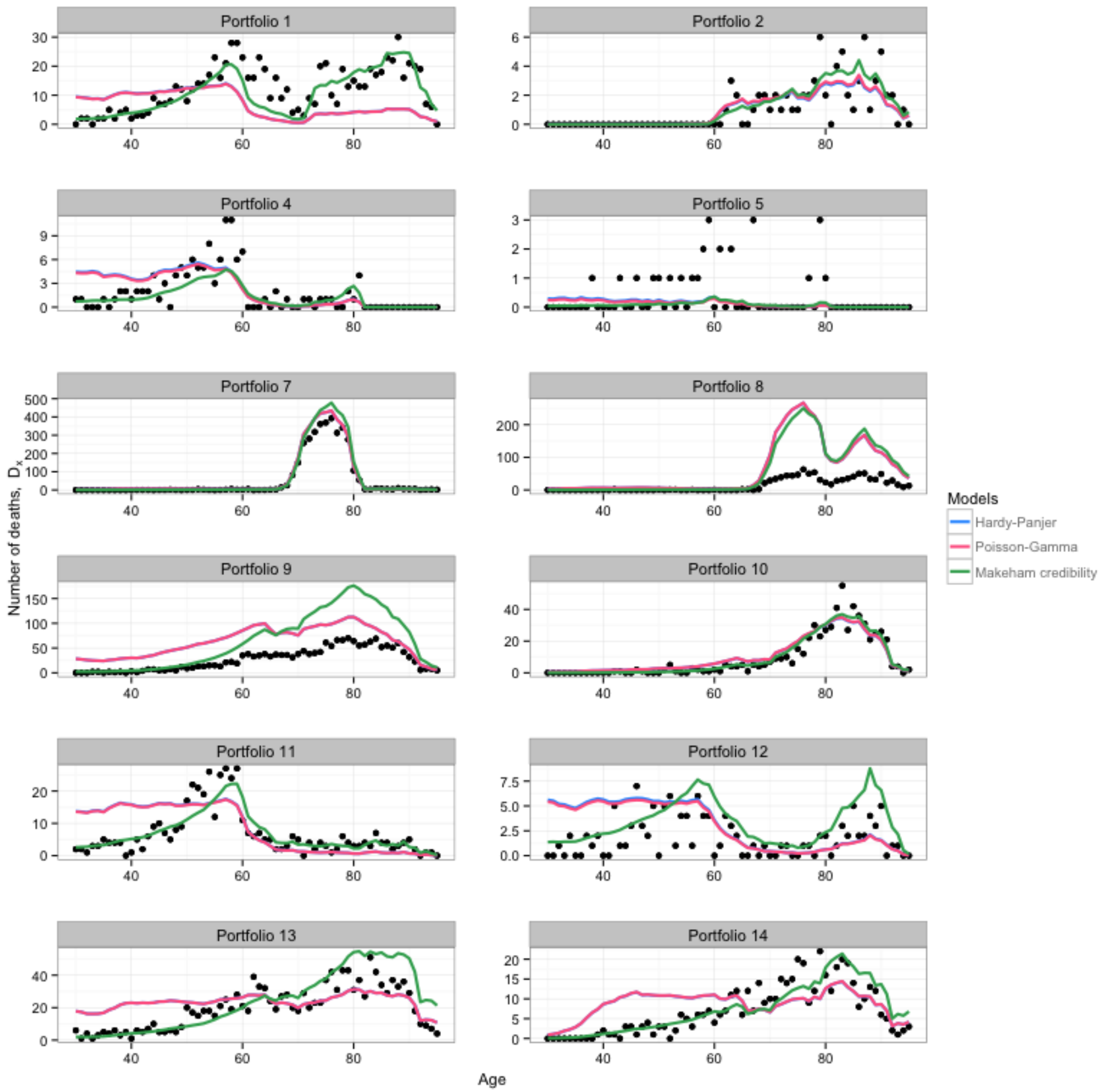

FIGURE A.4 – Fitted number of deaths for the year 2011, male population

FIGURE A.5 – Standardized residuals, calendar year 2010, male population

FIGURE A.6 – Standardized residuals, calendar year 2011, male population

Bibliographie

Aalen, O., "Nonparametric inference for a family of counting processes," *The Annals of Statistics* (1978), 701–726.

Anderson, J. and A. Senthilselvan, "Smooth estimates for the hazard function," *Journal of the Royal Statistical Society. Series B (Methodological)* (1980), 322–327.

Bagkavos, D. and P. Patil, "Variable bandwidths for nonparametric hazard rate estimation," *Communications in Statistics–Theory and Methods* 38 (2009), 1055–1078.

Barrieu, P., H. Bensusan, N. El Karoui, C. Hillairet, S. Loisel, C. Ravanelli and Y. Salhi, "Understanding, modelling and managing longevity risk : key issues and main challenges," *Scandinavian actuarial journal* 2012 (2012), 203–231.

Blanpain, N. and O. Chardon, "Projections de populations 2007-2060 pour la France métropolitaine : méthode et principaux résultats," Série des Documents de Travail de la direction des statistiques Démographiques et Sociales F1008, Institut National de la Statistique et des Études Économiques, 2010.

Bongaarts, J., "Long-range trends in adult mortality : Models and projection methods," *Demography* 42 (2005), 23–49.

Breslow, N., C. Chiang, J. Hoem, P. Jagers, N. Keiding, L. Kish, K. Manton and J. Tukey, "The natural variability of vital-rates and associated statistics-discussion," *Biometrics* 42 (1986), 712–732.

Bühlmann, H. and A. Gisler, *A course in credibility theory and its applications* (Springer, 2005).

Cleveland, W. S., "Robust locally weighted regression and smoothing scatterplots," *Journal of the American Statistical Association* 74 (1979), 829–836.

Code des assurances (Journal Officiel de la République Française, 2017).

Copas, J. and S. Haberman, "Non-parametric graduation using kernel methods," *Journal of the Institute of Actuaries* 110 (1983), 135–156.

Cox, D., "Regression Models and Life-Tables," *Journal of the Royal Statistical Society. Series B (Methodological)* 34 (1972), 187–220.

Currie, I., "Fitting models of mortality with generalized linear and non-linear models.," Technical Report, 2013.

Debón, A., F. Montes and R. Sala, "A comparison of nonparametric methods in the graduation of mortality : Application to Data from the Valencia Region (Spain)," *International statistical Review* 74 (2006), 215–233.

Delwarde, A., D. Kachkhidze, L. Olie and M. Denuit, “Modèles linéaires et additifs géneralisés, maximum de vraisemblance local et méthodes relationelles en assurance sur la vie,” *Bulletin Français d’Actuariat* 6 (2004), 77–102.

Fan, J. and I. Gijbels, “Data-driven bandwidth selection in local polynomial fitting : variable bandwidth and spatial adaptation,” *Journal of the Royal Statistical Society. Series B (Methodological)* (1995), 371–394.

———, *Local polynomial modelling and its applications : monographs on statistics and applied probability 66*, volume 66 (CRC Press, 1996).

Felipe, A., M. Guillén and A. Pérez-Marín, “Recent mortality trends in the Spanish population,” *British Actuarial Journal* 8 (2002), 757–786.

Forfar, D., J. McCutcheon and A. Wilkie, “On graduation by mathematical formula,” *Journal of the Institute of Actuaries* 115 (1988a), 1–149.

———, “On graduation by mathematical formula,” *Journal of the Institute of Actuaries* 115 (1988b), 1–459.

Gámiz, M. L., E. Mammen, M. D. M. Miranda and J. P. Nielsen, “Double one-sided cross-validation of local linear hazards,” *Journal of the Royal Statistical Society : Series B (Statistical Methodology)* 78 (2016), 755–779.

Gavrilova, N. and L. Gavrilov, “Ageing and Longevity : Mortality Laws and Mortality Forecasts for Ageing Populations,” *Demografie* 53 (2011), 109–128.

Gompertz, B., “On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies,” *Philosophical transactions of the Royal Society of London* 115 (1825), 513–583.

Gram, J., *Om Rækkeudviklinger, bestemte ved Hjælp af de mindste Kvadraters Methode*, Ph.D. thesis, Andr. Fred. Høst & Son (1879).

———, “Ueber die Entwickelung reeller Functionen in Reihen mittelst der Methode der kleinsten Quadrate.,” *Journal für die reine und angewandte Mathematik* 94 (1883), 41–73.

Gray, R. J., “Some diagnostic methods for Cox regression models through hazard smoothing,” *Biometrics* (1990), 93–102.

Gustafsson, J., J. Nielsen, P. Pritchard and D. Roberts, “Quantifying operational risk guided by kernel smoothing and continuous credibility,” *The ICFAI Journal of Financial Risk Management* 3 (2006), 23–47.

Gustafsson, J. K. A., J. Nielsen, P. Pritchard and D. Roberts, “Quantifying operational risk guided by kernel smoothing and continuous credibility-A practioner’s view,” *Journal of Operational Risk* 1 (2009), 43–55.

Hardy, M. and H. Panjer, “A credibility approach to mortality risk,” *ASTIN Bulletin* 28 (1998), 269–283.

Hougaard, P., “Life table methods for heterogeneous populations : distributions describing the heterogeneity,” *Biometrika* 71 (1984), 75–83.

Hyndman, R. J., H. Booth and F. Yasmeen, “Coherent mortality forecasting : the product-ratio method with functional time series models,” *Demography* 50 (2013), 261–283.

Jiang, J. and K. Doksum, “On local polynomial estimation of hazard rates and their derivatives under random censoring,” *Lecture Notes-Monograph Series* (2003), 463–481.

Jones, M. C., J. S. Marron and S. J. Sheather, “A brief survey of bandwidth selection for density estimation,” *Journal of the American Statistical Association* 91 (1996), 401–407.

Kaplan, E. L. and P. Meier, “Nonparametric estimation from incomplete observations,” *Journal of the American statistical association* 53 (1958), 457–481.

Keyfitz, N., “The limits of population forecasting,” *Population and development review* (1981), 579–593.

Lee, R. D. and L. R. Carter, “Modeling and forecasting US mortality,” *Journal of the American statistical association* 87 (1992), 659–671.

Lejeune, M., “Estimation non-paramétrique par noyaux : régression polynomiale mobile,” *Revue de Statistique Appliqu ee* 23 (1985), 43–67.

Liddell, F. D. K., “Simple exact analysis of the standardised mortality ratio,” *Journal of Epidemiology and Community Health* 38 (1984), 85–88.

Loader, C., *Local regression and likelihood* (Springer Science & Business Media, 2006).

Lourdes, C., “The Bühlmann-Straub Model with the premium calculated according to the variance principle,” *Insurance : Mathematics & Economics* 8 (1989), 3–10.

Makeham, W. M., “On the law of mortality,” *Journal of the Institute of Actuaries* 13 (1867), 325–358.

———, “On the Further Development of Gompertz’s Law,” *Journal of the Institute of Actuaries* 28 (1890), 316–332.

Mammen, E., M. D. Martínez Miranda, J. P. Nielsen and S. Sperlich, “Do-validation for kernel density estimation,” *Journal of the American Statistical Association* 106 (2011), 651–660.

Musset, A.-S., *Risque de longévité : la référence à la tendance de la population nationale est-elle justifiée ?*, Master’s thesis, Mémoire d’Actuaire, ISFA, ISFA (2016).

Nelson, W., “Theory and applications of hazard plotting for censored failure data,” *Technometrics* 14 (1972), 945–966.

Ngai, A. and M. Sherris, “Longevity risk management for life and variable annuities : The effectiveness of static hedging using longevity bonds and derivatives,” *Insurance : Mathematics and Economics* 49 (2011), 100–114.

Nielsen, J. and B. Sandqvist, “Credibility weighted hazard estimation,” *Astin Bulletin* 30 (2000), 405–418.

———, “Proportional hazard estimation adjusted by continuous credibility,” *ASTIN BULLETIN* 35 (2005), 239–258.

Nielsen, J. P., C. Tanggaard and M. Jones, “Local linear density estimation for filtered survival data, with bias correction,” *Statistics* 43 (2009), 167–186.

Pérez, M. L. G., L. Janys, M. D. M. Miranda and J. P. Nielsen, "Bandwidth selection in marker dependent kernel hazard estimation," *Computational Statistics & Data Analysis* 68 (2013), 155–169.

Plat, R., "Stochastic portfolio specific mortality and the quantification of mortality basis risk," *Insurance : Mathematics and Economics* 45 (2009), 123–132.

Ramlau-Hansen, H., "The choice of a kernel function in the graduation of counting process intensities," *Scandinavian Actuarial Journal* 1983 (1983), 165–182.

Renshaw, A., S. Haberman and P. Hatzopoulos, "The modelling of recent mortality trends in United Kingdom male assured lives," *British Actuarial Journal* 2 (1996), 449–477.

Salhi, Y., P.-E. Thérond and J. Tomas, "A credibility approach of the Makeham mortality law," *European Actuarial Journal* 6 (2016), 61–96.

Solvabilité 2, Directive 2009/138/CE (Journal Officiel de l'Union Européenne, 2009).

Thatcher, A. R., "The long-term pattern of adult mortality and the highest attained age," *Journal of the Royal Statistical Society : Series A (Statistics in Society)* 162 (1999), 5–43.

Tibshirani, R. and T. Hastie, "Local likelihood estimation," *Journal of the American Statistical Association* 82 (1987), 559–567.

Tomas, J., "A local likelihood approach to univariate graduation of mortality," *Bulletin Français d'Actuariat* 11 (2011), 105–153.

Tomas, J. and F. Planchet, "Construction et validation des références de mortalité de place," Technical report II1291-11 v1.3, Institut des Actuaires, 2013.

———, "Constructing entity specific projected mortality table : adjustment to a reference," *European Actuarial Journal* 4 (2014), 247–279.

Vaupel, J., K. G. Manton and E. Stallard, "The impact of heterogeneity in individual frailty on the dynamics of mortality," *Demography* 16 (1979), 439–454.

Wand, M. P. and M. C. Jones, *Kernel smoothing* (Crc Press, 1994).

Wang, W., *Proportional Hazards Regression Model with Unknown Link Function and Applications to Longitudinal Time-to-event Data*, Ph.D. thesis, University of California, Davis (2001).

———, "Proportional hazards regression models with unknown link function and time-dependent covariates," *Statistica Sinica* (2004), 885–905.

Weibull, W., "Wide applicability," *Journal of applied mechanics* 103 (1951), 293–297.

Zhu, Z. and Z. Li, "Logistic Regression for Insured Mortality Experience Studies," *SCOR Inform* (2013).