Annals of Actuarial Science (2020), **14**, pp. 420–444
doi:10.1017/S1748499520000081

PAPER

Mortality data reliability in an internal model

Fabrice Balland[1], Alexandre Boumezoued[2]*, Laurent Devineau[2], Marine Habart[1] and Tom Popa[1]

[1]GIE AXA, 21 Avenue Matignon, 75008 Paris, France; and [2]Milliman, 14 Avenue de la Grande Armée, 75017 Paris, France
*Corresponding author. Email: alexandre.boumezoued@milliman.com

(Received 03 January 2019; revised 02 September 2019; accepted 03 April 2020)

Abstract

In this paper, we discuss the impact of some mortality data anomalies on an internal model capturing longevity risk in the Solvency 2 framework. In particular, we are concerned with abnormal cohort effects such as those for generations 1919 and 1920, for which the period tables provided by the Human Mortality Database show particularly low and high mortality rates, respectively. To provide corrected tables for the three countries of interest here (France, Italy and West Germany), we use the approach developed by Boumezoued for countries for which the method applies (France and Italy) and provide an extension of the method for West Germany as monthly fertility histories are not sufficient to cover the generations of interest. These mortality tables are crucial inputs to stochastic mortality models forecasting future scenarios, from which the extreme 0.5% longevity improvement can be extracted, allowing for the calculation of the solvency capital requirement. More precisely, to assess the impact of such anomalies in the Solvency II framework, we use a simplified internal model based on three usual stochastic models to project mortality rates in the future combined with a closure table methodology for older ages. Correcting this bias obviously improves the data quality of the mortality inputs, which is of paramount importance today, and slightly decreases the capital requirement. Overall, the longevity risk assessment remains stable, as well as the selection of the stochastic mortality model. As a collateral gain of this data quality improvement, the more regular estimated parameters allow for new insights and a refined assessment regarding longevity risk.

Keywords: Mortality; Longevity; Cohort effect; Solvency 2; Internal model

1. Introduction

In this paper, our aim is to discuss and assess the impact of anomalies in national mortality tables provided by the Human Mortality Database[1] (HMD) on the internal model of a typical insurer which captures longevity risk in the Solvency 2 framework[2]. Of special interest in this paper are some abnormal cohort effects observed on period tables in which some diagonals show special patterns. Particular attention is devoted to the 1919–1920 effect: the 1919 diagonal shows particularly

[1] *Human Mortality Database. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at* www.mortality.org *or* www.humanmortality.de. It is worth mentioning that at the time of writing, the Human Mortality Database released an update on February 2018, including in particular a revision of exposure calculation based on monthly birth counts. We refer the reader to the Version 6 of the HMD Methods Protocol for more details. This Version 6 of HMD has not corrected mortality rates for West Germany for cohorts born before 1946 because HFD monthly fertility records for West Germany are only provided starting at year 1946.

[2]The main contribution of this paper from our point of view is to provide a real case study based on a realistic industry model. The counterpart is that we are limited on developing some technical details to preserve confidentiality. The reader has to be aware of this limitation.

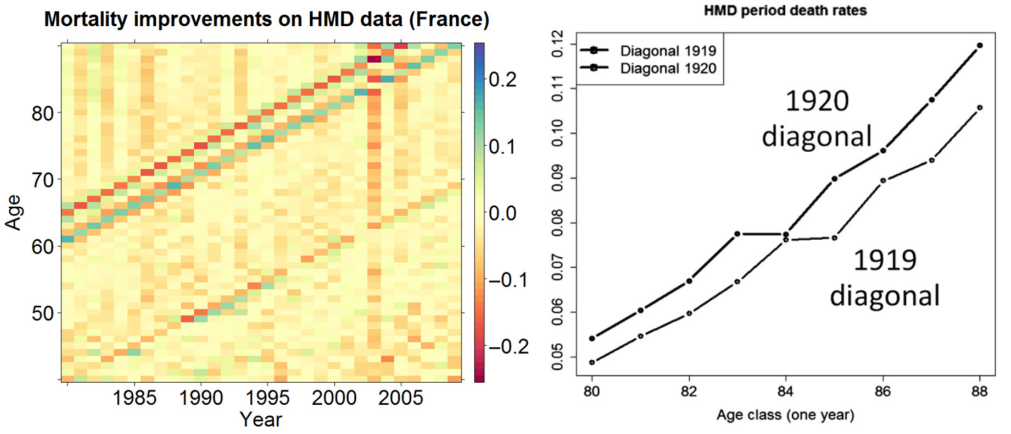

Figure 1. Abnormal cohort effects: example of the period table for France.

low mortality rates, whereas the 1920 generation shows particularly high ones. This effect can be easily observed using matrices of mortality improvements by age and time, on which clear diagonal effects appear, or by comparing the mortality rates of the generations considered, see the example of France in Figure 1.

Such results are counter intuitive regarding natural demographic insight: some populations born more recently live less longer on average (see again Figure 1), as it is indeed observed that the "younger" generation 1920 shows higher mortality rates in the period table.

In the actuarial literature, few references have been focusing on mortality data reliability. The awareness about such anomalies has recently emerged from the work by Richards (2008), Cairns et al. (2016) and Boumezoued (2016) (see also Boumezoued & Devineau, 2017). To our knowledge, the first conjecture about the potential causes of these isolated cohort effects was by Richards (2008). He focused on the 1919 birth cohort for England and Wales for which he suggested the possibility of errors due to erratic number of births. The methodology of the Office for National Statistics has been studied by Cairns et al. (2016) in several directions, including assumptions related to fertility. Then, Boumezoued (2016) underlined the universality of such anomalies (isolated cohort effects) by highlighting them in a variety of countries from the HMD. By studying the HMD methodology to construct period tables, he proposed to source its fertility counterpart, the Human Fertility Database (HFD), and to apply the philosophy of the work by Cairns et al. (2016) to produce adjusted period mortality tables for a set of countries.

In this paper, we discuss the impact of some mortality data anomalies on the longevity risk module of an internal model in the framework of Solvency 2, and in particular on trend risk which can be defined as the risk that the trend driving the future longevity evolution may experience unexpected changes. The three countries of interest in this paper are France, Italy and West Germany, due to their business exposure, and as they have been identified as embedding anomalies in their HMD mortality data. By applying and extending the correction methodology as described in Boumezoued (2016), we are able to measure the discrepancy between original and adjusted period mortality tables. These discrepancies are crucial in the way the insurance market measures and manages longevity risk, especially in the present context.

Among life risks modules, one of the most important risks threatening insurers is longevity risk, which is the risk that insured people may on average survive longer than expected. The longevity exposure is the result of long-term commitments (up to 60 years) with high uncertainties in the midterm (10–20 years). Several decades and therefore several generations of contracts might be necessary before detecting any risk deviation. There is no consensus on the future evolution of longevity, either in terms of model or market price. Many interconnected factors are to

https://doi.org/10.1017/S1748499520000081 Published online by Cambridge University Press

be considered in the future longevity trend, as, for example, technological improvements, socioeconomic trends, political systems, and demographic structures. While some experts consider that the human being has already reached its maximum age, others dream of immortality thanks to transhumanism. Facing with such uncertainty, preparing insurance companies for ageing issues and its business consequences is fundamental.

As a reminder, the Solvency II Directive became fully applicable on1 January 2016[t]. Solvency II reviews the prudential regime for insurance and reinsurance undertakings in the European Union in a harmonised prudential framework. The risk profile of each individual insurance company is further considered, with the aim to promote comparability, transparency and competitiveness. The Solvency II Directive provides two ways of measuring risk: insurance and reinsurance companies can use either the standard formula or their own internal model, which enables them to assess their own risk more accurately. In the framework of the standard formula, a risk classification is provided, which includes risk modules (such as market risks, life risks, and non-life risks) and risk submodules (such as mortality, longevity, lapse, and catastrophic for life risks module). For each submodule, a stress test or a closed formula is used to determine a capital charge. The capital requirement is then calculated using a bottom-up approach, by aggregating submodules and then modules based on a correlation matrix. In comparison, the development of an internal model allows the company to refine its own risk assessment. In particular regarding life risks, the company has the possibility to use external data which reflects its risk in terms of countries considered, such as general population tables as provided by the HMD for more than 30 countries and regions worldwide. It is the purpose of the present paper to identify, adjust and test the impact of anomalies in national general population tables as key inputs in the internal model.

It is worth mentioning that the longevity shocks calibration for the standard formula has resulted from the comparison of retrospective and prospective analyses, see the Committee of European Insurance and Pensions Supervisors (CEIOPS') Advice for Level 2 Implementing Measures on Solvency II, especially the Annex B "Longevity risk calibration analysis". More specifically, two analyses have been conducted: the first one focusing on past historical mortality improvements and the second one generating shocks in the future. It has been concluded that the direct use of historical data led to a higher longevity stress compared to that derived from the forecasts of a stochastic model. In the same spirit, a prospective and retrospective analysis has been derived by Boumezoued (2016) on both crude and adjusted mortality data; the analysis of adjusted mortality tables showed that, not only the mortality levels for several cohorts were highly over/underestimated, but also that the volatility of mortality improvement rates over the last 30 years was overestimated in original HMD mortality tables, with a large difference for many countries. In addition, it has been shown how the volatility levels reproduced by classical stochastic mortality models now closely match historical mortality improvements in adjusted tables, although it was not the case on crude data as pointed out by several studies as cited before.

The previous considerations illustrate the potential impact of such mortality data adjustment on several steps of internal modelling process. To assess the business impact of such a complex risk, we use in this paper a simplified internal model specification. Market practices rely on two main classes of stochastic mortality models: the first class regroups models derived from – Carter approach, while the second class of models is derived from Cairns, Blake and Dowd approach. Usually calibrated on national mortality tables, such models have proven their ability to capture historical mortality behaviours. However, a core issue is their sensitivity to possible singularities or anomalies in the underlying mortality data.

In this paper, a simplified prototype of internal model is used and embeds the following features: a set of mortality models is fitted based on national population mortality data, and the "best" model is selected based on statistical and qualitative criteria. The selected mortality model is then used to draw simulated forecasts, and a prospective mortality table corresponding to the 99.5th percentile in terms of improvements is built. This allows to provide the longevity shock at the core of the Solvency II economic capital calculation, here based on an annuity product portfolio.

https://doi.org/10.1017/S1748499520000081 Published online by Cambridge University Press

The paper is organised as follows. Section 2 deals with the mortality data anomalies and focuses on the methodologies to detect and adjust them, as well as on the key characteristics of adjusted mortality tables. This includes the use of the method developed by Boumezoued (2016) to adjust period tables for Italy and France, and the novel approach we propose in order to update the West Germany mortality table, for which fertility histories are not sufficient to cover the 1919–1920 generations of interest. In Section 3, we introduce the stylised internal model which is used to calculate the solvency capital requirement (SCR) associated with longevity risk, and the impact of the mortality data anomalies on the internal model in the light of several analyses; in particular, we discuss the way it impacts the fit of stochastic mortality models, the future forecasts, the model selection process, the calculation of the SCR, as well as the stability of longevity risk assessment. The paper ends with some concluding remarks in Section 4.

2. Mortality Data Reliability Issues

In this section, we focus on the identification and correction of several anomalies in mortality tables obtained from the HMD for our three countries of interest: France, Italy and West Germany. After highlighting the recent awareness on such anomalies in the literature as well as the key observations on the crude datasets, we detail the adjustment methodologies we use and we describe the main characteristics of adjusted mortality tables.

2.1 On the recent awareness about anomalies in mortality tables

As mentioned in the Introduction, few actuarial work has been focusing on mortality data reliability, but the awareness about such anomalies has recently emerged from the work by Richards (2008), Cairns et al. (2016) and Boumezoued (2016). The last author put into evidence the universality of such anomalies (isolated cohort effects) by highlighting them in a variety of countries from the HMD. By studying the HMD methodology to construct period tables, he proposed to source its fertility counterpart, the HFD, and to apply the philosophy of the work by Cairns et al. (2016) to produce adjusted period mortality tables for a set of countries.

In this paper, we apply and extend the adjustment methodology as described in Boumezoued (2016). This way, we are able to measure the discrepancy between original and adjusted period mortality tables, which are crucial inputs for the measurement and management of longevity risk in the insurance market.

In order to present the up-to-date awareness on such abnormal cohort effects, we summarise below the main conclusions from the work of Boumezoued (2016):

1. While comparing period and cohort mortality tables, anomalies in period tables ones have been highlighted in the form of isolated cohort effects for several countries available in HMD: the period mortality rates for specific generations appear surprisingly low or high compared to the others.
2. The HMD methodology to construct mortality estimates has been identified to embed a strong assumption of uniform distribution of births that is specific to the computation of period mortality tables, which shows that we are facing a universal reliability issue which is shared by most countries.
3. To perform an automatic correction procedure, it is proposed to rely on the HFD, which is considered as the perfect counterpart of the HMD in terms of fertility.
4. The analysis of adjusted mortality tables for several countries led to the conclusion that isolated "cohort effects" in original mortality tables (often for years of birth around 1915, 1920, 1940 and 1945) are in fact universal anomalies that disappear in adjusted tables. Further analysis of adjusted mortality tables shows that, not only the mortality levels for several cohorts are highly over/underestimated, but also that the volatility of mortality improvement rates over the last 30 years was overestimated in original HMD mortality tables, with a large difference for many countries.

https://doi.org/10.1017/S1748499520000081 Published online by Cambridge University Press

In Section 2.2, we detail the correction methodology based on fertility data which is used in the present paper to adjust mortality tables for France and Italy. In Section 2.3, we motivate the need for an extension of the methodology, and we propose a novel approach to produce an adjusted table for West Germany. Finally, Section 2.4 details the main characteristics of the adjusted mortality tables and the key discrepancies with the original data.

As a final remark, let us emphasise that two kinds of mortality tables are provided by the HMD: *period* and *cohort*. Anomalies in terms of abnormal cohort effects have been identified in *period* mortality tables, which are naturally designed to study the dynamics of mortality from one year to the next, and therefore systematically used in practice to calibrate stochastic mortality models. For a review of the definition and construction methods for period and cohort tables, see Boumezoued (2016).

2.2 The correction method based on fertility data

In practice, period mortality tables for 1-year time periods and 1-year age classes are produced based on annual population estimates derived from census, as well as number of deaths combining the information of death reports and population counts. Note that in most countries, census may be performed out of the beginning of the year, or at intervals greater than 1 year, which leads the HMD to perform several adjustments, see Wilmoth et al. (2017) for more details. In our framework, however, we assume that input data provided by the HMD as annual population counts and number of deaths in Lexis triangles are accurate.

Two ingredients are at the core of death rate computation for a given period: the number of deaths in this period, which we assumed to be reliable, divided by the so-called "exposure-to-risk" which represents the "quantity" of individuals at risk of death. In the so-called period mortality tables, the exposure to risk takes the mathematical form of an integral over a 1-year time period and a 1-year age class, see Boumezoued (2016) for more details on the formalism. In this setting, the period mortality rate writes

$$m(x,t) = \frac{D(x,t)}{E(x,t)}$$

where $D(x,t)$ is the number of deaths and $E(x,t)$ is the total time lived in the year $[t, t+1)$ at age between x and $x+1$. The mortality rate is the core object to be modelled by stochastic mortality models, as it is also the case for the death probability $q(x,t)$, more interpretable as the probability for individuals aged x at time t to die in the next year, and linked to the mortality rates through the following equation:

$$q(x,t) = 1 - \exp\left(-m(x,t)\right)$$

In the standard computation practice, the exposure-to-risk component $E(x,t)$ is usually approximated based on annual population estimates, as it is done by reference providers as national institutes or the HMD. In other words, the standard practice works with the average between the population at the beginning of the year and that at the end of the year, which is taken as a proxy for the integral over the year in continuous time. This computation amounts to assume that births are uniformly distributed in the year, as well as from one cohort to the next.

Although for several years where population flows are quite stable the assumption may seem reasonable (although not perfect), the approximation error can be huge in situations in which population numbers are fluctuating in the year, that is particularly the case when births are erratic due to demographic shocks (e.g. before and after wars and pandemics). Such specific patterns create a "convexity effect", and in this context, it is no longer possible to approximate the exposure to risk by a standard average.

To correct for such approximation errors for countries available in the HMD, the approach proposed by Boumezoued (2016) relies on monthly fertility records available from the HFD. Such

https://doi.org/10.1017/S1748499520000081 Published online by Cambridge University Press

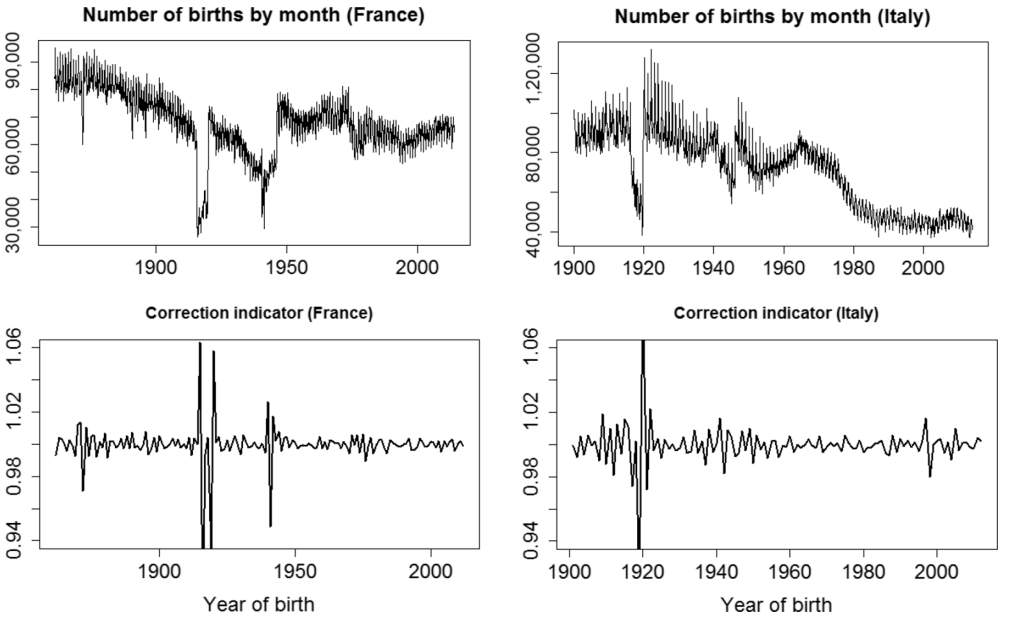

Figure 2. Number of births by months and correction indicator for France and Italy.

monthly population estimates are used to construct a "data quality indicator" assessing the reliability of estimates in the (period) mortality table for each generation. The indicator takes the form of a ratio which measures the deviation between an annual and a monthly approximation of the (annual) exposure to risk. For each year of birth, this ratio is then used to adjust the estimates along each diagonal of the mortality table (that is following each cohort), producing an adjusted table which does not present the initial anomalies.

In this latter work, adjusted period tables have been provided for France, Switzerland, Finland, Sweden and Austria, countries for which complete and deep fertility histories were available at time of publication. Meanwhile, the monthly fertility data for Italy has been officially released by the HFD, allowing us to use the same methodology for this country as well. Note also that this is the case of Iceland, which will be added to the set of explanatory countries for West Germany, see the next subsection c). In the following, we present the results in terms of correction indicator for France and Italy; the specific treatment of West Germany, for which the monthly fertility history is not deep enough, is detailed in Section 2.3. The key characteristics of adjusted period tables are then described in Section 2.4.

The number of births by months as well as the correction indicator computed for France and Italy using the method by Boumezoued (2016) is depicted in Figure 2. This figure shows how shocks in birth patterns impact the exposure-to-risk approximation as it is performed in period tables provided by the HMD. The virtue of the correction indicator appears from these graphs: for a given year of birth, a value greater (respectively, lower) than one indicates that the HMD period mortality rates for the generation are overestimated (respectively underestimated). In addition, the ratio measures the magnitude of the discrepancy between the annual exposure-to-risk approximation and its (better) monthly counterpart. It should be noted that although the dynamics of the number of births over the months differs between France and Italy, the 1919–1920 effect shows some universality: the period mortality rates for generation 1919 are underestimated, whereas those for generation 1920 are overestimated. Moreover, in each case the order of magnitude of the error is $\pm 6\%$, see again Figure 2; this may represent a level of particular attention for practitioners using national mortality data.

https://doi.org/10.1017/S1748499520000081 Published online by Cambridge University Press

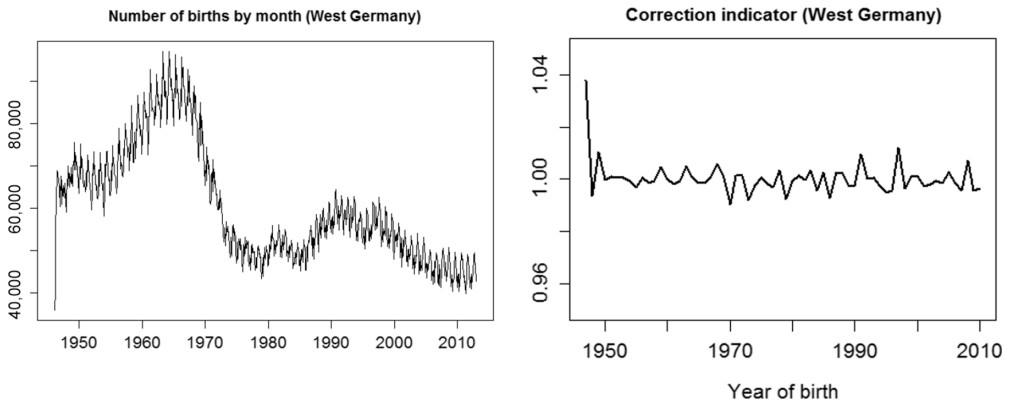

Figure 3. Number of births by months and correction indicator for West Germany.

2.3 An extension of the correction method for West Germany

As mentioned in the previous part, in its present form, the adjustment methodology can only be applied to years of birth for which the number of births by months is provided by the HFD. Unfortunately, this is not the case for West Germany for which the HFD monthly fertility records are provided starting at year 1946, which consequently limits the set of years of birth for the correction ratio (see Figure 3).

Therefore, there is a need to extend the previous methodology to enable producing an output adjusted table, especially for the typical 1919–1920 phenomenon which, again, lies outside of the historical period.

2.3.1 Regression with stepwise selection process

The idea now is to try to reconstruct the fertility history for West Germany, by looking at those other countries for which the birth series are available. Although number of births show clear different patterns from one country to another (see Figure 2, as well as Figure 9 in Boumezoued, 2016), the correction indicator, as a quantity without dimension (ratio of population estimates), shows quite similar patterns among the several countries.

Note that the approach developed here is intended to provide a first idea of possible adjustment without fertility data by month; the development of a full methodology to adjust population exposures in the absence of births-by-month data is left for further research. Let us also clarify that the results in terms of SCR will be provided for each country (in fact, each insurance entity within each country), therefore one has to keep in mind that a larger uncertainty remains concerning the calculation based on the Germany West adjusted data.

To infer the correction indicator for West Germany starting from that of other countries, we propose to use a multiple regression approach, while performing an optimal selection procedure. In order to properly detail the regression and prediction methodology, let us introduce some notations. We denote by $I_C(t)$ the correction indicator available for country C and year of birth t. We emphasise here that t is annual and take its values in the space of all possible years of birth. The set of countries for which reasonable monthly fertility series are now available is set as

$$\mathcal{S} = \{\text{Finland, France, Iceland, Italy, Sweden, Switzerland}\}$$

This set is selected in order to satisfy the two following criteria.

https://doi.org/10.1017/S1748499520000081 Published online by Cambridge University Press

- Data on births by month are officially released by the HFD (i.e. not stated as "preliminary release"), to avoid any major data issue on birth series.
- The monthly fertility series allow constructing correction indicators for years of birth before 1914, which covers the 1919–1920 effect with margin.

The aim now is to find an optimal set of regressors $\mathcal{S}^* \subset \mathcal{S}$, and coefficients $\{\alpha_C,\ C \in \mathcal{S}^*\}$ and μ such that

$$I_{\text{West Germany}}(t) = \mu + \sum_{C \in \mathcal{S}^*} \alpha_C\, I_C(t)$$

for each t such as the indicator for West Germany, $I_{\text{West Germany}}(t)$ can be constructed, that is here for each year of birth t between 1947 and 2010 (2010 being chosen as an upper bound for all countries in order to avoid depending on recent revisions of demographic data).

2.3.2 Results of the regression and prediction of the correction indicator

The method carried out is a step-by-step comparison of the models obtained as a combination of regressors from the set $\mathcal{S}$. At this stage, a statistical criterion has to be specified in order to compare models; we tested the classical criteria Bayesian Information Criterion (BIC) and adjusted R-square, both providing some penalisation of the fitting ability of the model by the number of parameters. The results in terms of "best" model selected and associated parameter values are presented in the table below (rounded at two significant numbers).

Interestingly, the results show strong similarities between the BIC and the R-squared criteria. In particular, in each case the "best" model includes Finland and France as predictors of West Germany, although the adjusted R-squared criterion also includes Italy. Another interesting feature is that the three countries of interest here, which have been identified to embed strong 1919–1920 anomalies, are linked with each other (with the adjusted R-squared criterion), France and Italy being key predictors for West Germany. In terms of parameters, significant weight is given to France correction ratio as an explanatory variable in each case.

At this stage, empirical expert judgement has to be used to define the final choice between both criteria; in this paper, we choose the results given by the R-squared adjustments having in mind the three following arguments.

- It is preferred to use more series to increase stability, therefore to include Italy as well, as we consider it as a good input to explain fertility shocks in West Germany.
- The output adjusted table for West Germany according to the R-squared criterion, see the next section, shows graphically even less isolated cohort trends than those produced according to the BIC criterion.
- On the whole, both criteria (BIC or R-squared) provide similar results in terms of residuals, correction indicator series and output adjusted tables.

2.3.3 Prediction of the correction indicator

For birth years outside the range of the correction indicator availability for West Germany, that is for t before 1946, the *true* indicator $I(t)$ can be predicted by its estimate $\hat{I}(t)$ according to the regression formula

$$\hat{I}^{\text{West Germany}}(t) = \hat{\mu} + \sum_{C \in \mathcal{S}^*} \hat{\alpha}_C I_C(t)$$

The estimated parameters are given in Figure 4 and the reconstructed correction indicator is depicted in Figure 5, see the left panel for the full range of cohorts. The key fact which appears from the observation of the reconstructed correction ratio is the following: although almost no spikes were observed in the available period for regression (plain line), the use of external data

https://doi.org/10.1017/S1748499520000081 Published online by Cambridge University Press

	BIC criterion	Adjusted R-squared
$\mathcal{S}^*$	{Finland, France}	{Finland, France, Italy}
$\hat{\mu}$	–0.19	–0.14
$\hat{\alpha}_{\text{Finland}}$	0.38	0.21
$\hat{\alpha}_{\text{France}}$	0.81	0.63
$\hat{\alpha}_{\text{Iceland}}$	-	-
$\hat{\alpha}_{\text{Italy}}$	-	0.29
$\hat{\alpha}_{\text{Sweden}}$	-	-
$\hat{\alpha}_{\text{Switzerland}}$	-	-

Figure 4. Results of the linear regression.

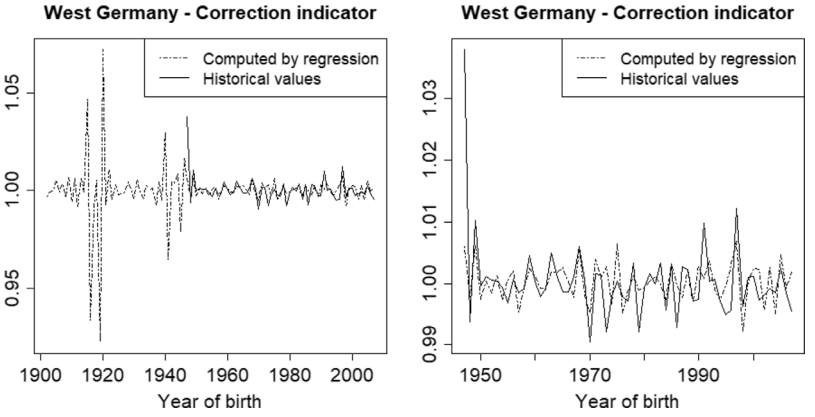

Figure 5. Completion by regression of the correction indicator for West Germany.

allows us to reproduce the generational fertility shocks at the core times in history (including around 1915 and around 1920) – this feature captured here lies in the universality of the abnormal cohort countries in the selected regressors, but also the precise monthly timing of the war which seems to be consistent between the different countries selected.

Note that Figure 5 also provides on the right panel a more detailed analysis of the fitted correction indicator, compared to the original indicator as calculated for cohorts for which births-by-month data are available. This figure shows that the method works for many years, with, however, a very large discrepancy for year 1947, as well as noticeable discrepancies for, for example, years between 1970 and 1975. The correction efficiency of the regression approach presented here will also be evaluated in the next section in which we analyse the main characteristics of the adjusted mortality tables produced.

Remark: It is worth mentioning that although this extrapolation over another time period may induce some instability issues in several regression contexts due to non-stationarity effects, we rely here on the fact that the selected regressors are "mirror" countries which show similar error patterns (due to external and internal shocks), which are thus likely to reproduce the 1919–1920 effect (in between the Spanish flu and the end of the First World War). The illustrative results detailed in the next part indeed show the efficiency of the method to adjust the 1919–1920 effects, while providing an underlying interpretation of the correction indicator dynamics based on "explanatory" countries. At this stage, we argue that significant theoretical work remains to be done in this direction, which is out of scope of the present paper.

https://doi.org/10.1017/S1748499520000081 Published online by Cambridge University Press

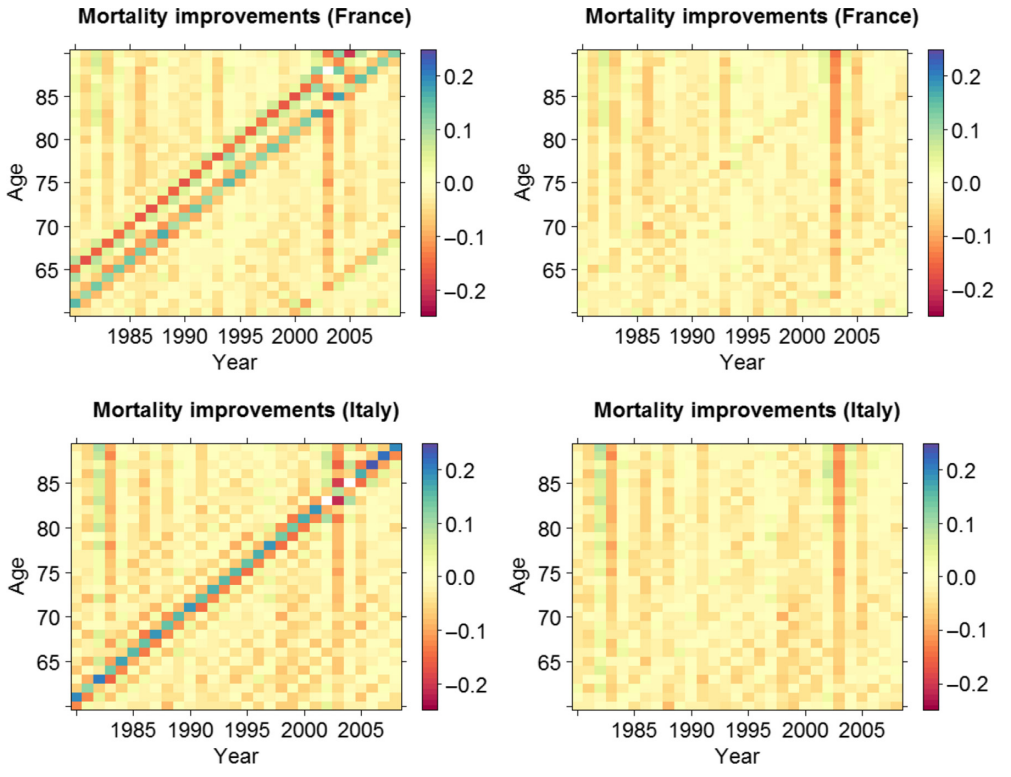

Figure 6. Crude and adjusted mortality improvements for France and Italy.

2.4 Analysis of adjusted mortality tables

Let us denote by $I(b)$ the correction indicator for year of birth b. Then the adjusted period mortality rates are constructed as

$$\tilde{m}(x,t) = \frac{m\,(x,t)}{I(t-x)}$$

Then matrices of crude and adjusted mortality improvement rates $r(x,t)$ or $\tilde{r}(x,t)$ can be analysed and compared; mortality improvement rates are here defined as

$$r\,(x,t) = \frac{m\,(x,t+1)}{m\,(x,t)} - 1 \text{ and } \tilde{r}\,(x,t) = \frac{\tilde{m}\,(x,t+1)}{\tilde{m}\,(x,t)} - 1.$$

The data used to specify the crude death rates $m(x,t)$ by single year and 1-year age bands were downloaded from the HMD on 1 September 2015 for France (civil population), Italy (civil population) and West Germany (total population).

2.4.1 Adjusted mortality tables for France and Italy

The matrices of crude and adjusted mortality improvement rates by age and time are depicted in Figure 6. The matrices are centred on the 1919–1920 effects of interest, and the colour scale is fixed. The correction of these effects for France and Italy, and also the correction for France of the additional anomalies around years of birth 1915 and 1940 can clearly be observed. Note that for all the three countries, we use both male and female adjusted tables, although we represent that related to the total population as a purpose of illustration.

https://doi.org/10.1017/S1748499520000081 Published online by Cambridge University Press

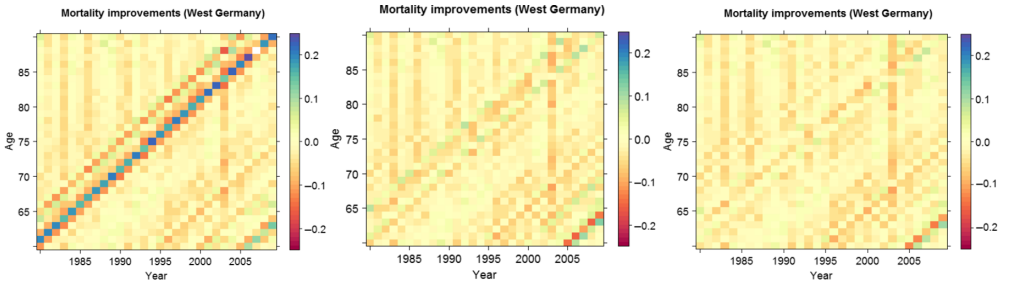

Figure 7. From left to right: crude mortality improvements for West Germany, and the adjusted versions according to the BIC and R-squared criteria.

2.4.2 Adjusted mortality tables for West Germany

As stated in Section 2.3, the efficiency of the regression approach to reconstruct the correction indicator for West Germany is to be assessed by comparing the crude and adjusted period mortality table. The results of the method are depicted on improvement rate matrices in Figure 7 for the two criteria selected (BIC or R-squared); from left to right: crude data, updated data with BIC selection, updated data with adjusted R-squared selection.

The correction process appears to be efficient concerning the 1919–1920 effects, which are the generations of interest in this paper. It should be noted that the 1915 effect is also clearly removed, but additionally that the correction specific to year of birth 1945 correction is not fully performed. Although in our context, this is a second-order issue as our main interest lies in the proper correction of the huge 1915 and 1920 anomalies, developing an improved correction method for such generations is left for further research.

Recall that the retained criterion is R-squared for the reasons detailed in Section 2.3, including in particular the fact that the adjusted table shows even less isolated cohort trends.

2.4.3 Adjustement diagnosis plots

In this part, we complete the analysis of the correction efficiency by studying the regularity of the mortality surface after the adjustment process. To do so, we rely on a similar quantity as the so-called empirical concavity function as introduced and studied in Cairns et al. (2016); our formulation is homogeneous to the discrete version of the second-order derivative in time of the mortality rate and writes

$$C(x,t) = \left| \frac{m\,(x,t+1) - 2m(x,t) + m\,(x,t-1)}{2} \right|$$

The concavity function for both crude and adjusted tables are depicted in Figure 8. Recall that classical mortality surfaces for medium-sized countries show relatively small concavity function values, as it is observed for generations before 1915 in both tables. The graphs show that the concavity function is a key tool to detect anomalies in crude tables, as already highlighted by Cairns et al. (2016). They also show that the adjusted table encompasses concavity function magnitudes for cohorts 1915–1920 in the same order as the previous generations, although still high values can be observed for the highest ages in these cohorts. This unsmoothness could be expected due to the apparent period effect for year 2003 which corresponds to the European heat wave. Finally, this figure shows some room for improvement of the method, especially regarding the 1939–1945 cohorts for which the adjustment is not well performed (as already discussed for Figure 7).

https://doi.org/10.1017/S1748499520000081 Published online by Cambridge University Press

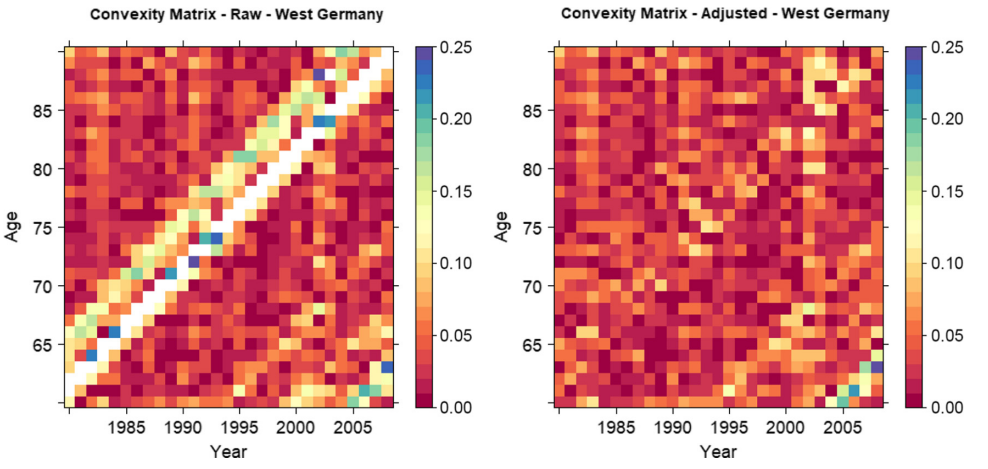

Figure 8. Convexity matrix for the raw and adjusted West Germany mortality table.

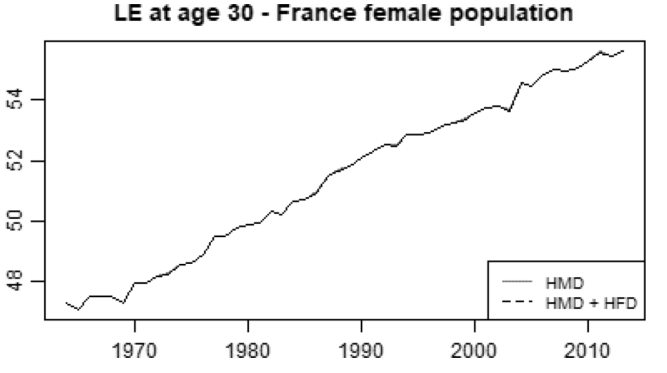

Figure 9. Historical period life expectancy at age 30, truncated at age 95, before and after HFD correction.

2.4.4 Key characteristics of adjusted historical mortality tables

To focus on historical data without introducing any external effect, analyses of the following section are conducted on mortality data embedded in the historical data tables only. As such, life expectancy and mortality indicators are restricted to ages 30 and 95.

While the method is applied on the entire mortality table, only specific generations are corrected and the main table characteristics are preserved, as shown for the period life expectancy of each year in Figure 9.

The following section highlights that these artificial cohort effects are identifiable on basic quantities usually derived on crude mortality data, as the force of mortality and the death functions.

2.4.5 Period-based forces of mortality

The analysis of the forces of mortality of French woman in Figure 10 reveals the transitory and specific volatility of generations born between 1918 and 1923 in the period-based mortality tables of year 1964–2013.

2.4.6 Death curves

In mirror to the force of mortality, the next two graphs in Figure 11 present the death curves of France female population over the same four periods between years 1964 and 2009. It is

https://doi.org/10.1017/S1748499520000081 Published online by Cambridge University Press

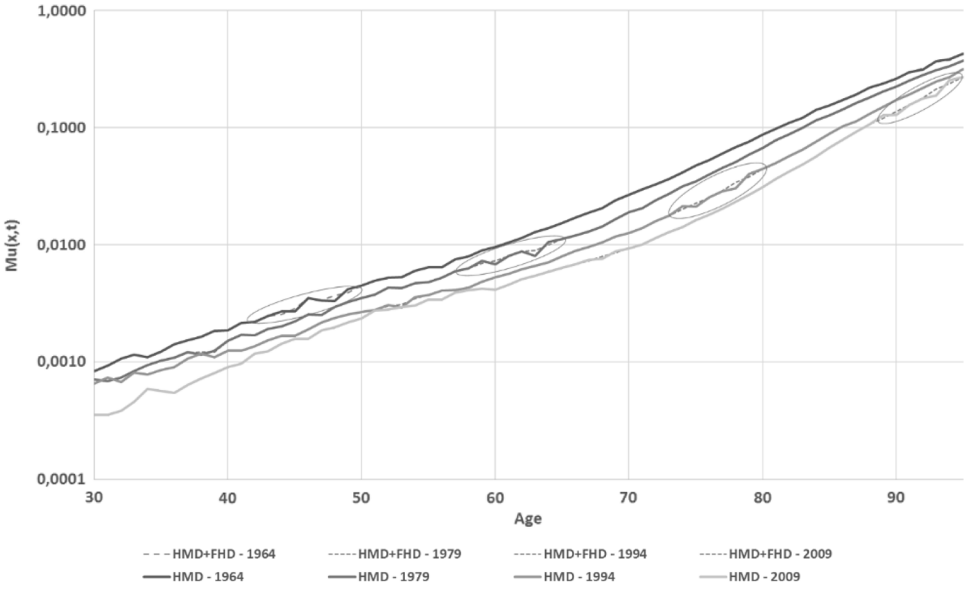

Figure 10. Comparative period force of mortality before and after HFD adjustment, France – Women population.

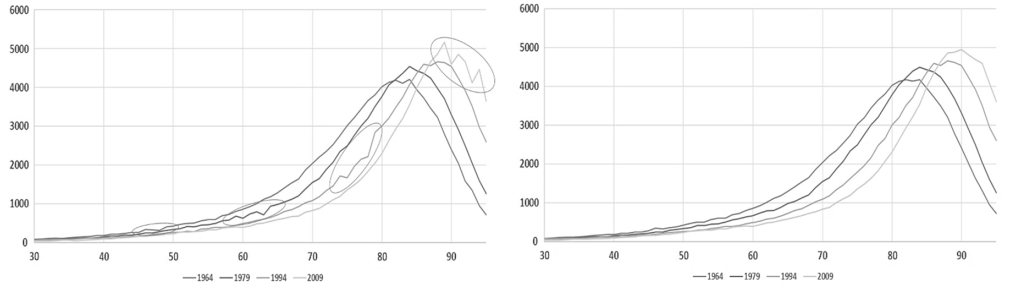

Figure 11. Period-based death functions, before (left) and after (right) HFD correction.

noticeable how adjusted tables increased the regularity of the death curves. Results are similar for all countries and gender in our study.

In summary to the past section, analytical data show that the correction added to HMD data effectively removes artificial cohort effects around generations 1918–1923, while keeping main mortality characteristics.

3. The Impact of Mortality Data Anomalies on an Internal Model

This section details the impact of retreated national population exposure on several classical mortality models. In a first step, we present the evolution of the fit of each mortality model used (Lee–Carter, Age–Period–Cohort, and Plat) and in a second step the impact on projected life expectancy based on both input sets (HMD versus HMD + HFD). Finally, we discuss how the adjusted mortality data influence the selection of the best actuarial model, the computation of the SCR, as well as on the stability of the longevity risk assessment.

3.1 Internal model specification

The simplified prototype of internal model used for this analysis works as follows: a set of mortality models is fitted based on national population mortality data, and the "best" model is selected

https://doi.org/10.1017/S1748499520000081 Published online by Cambridge University Press

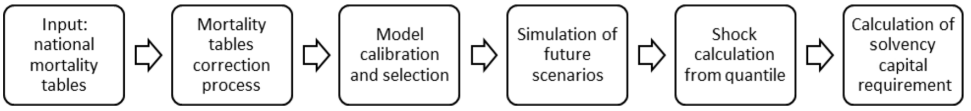

Figure 12. The mortality table adjustment process within the internal model workflow.

based on both statistical and qualitative criteria. The selected mortality model is then used to draw simulated forecasts, and a prospective mortality table corresponding to the 99.5th percentile in terms of improvements is built. This allows to provide the longevity shock at the core of the Solvency II economic capital calculation, here based on an annuity product portfolio. Longevity risk can be divided into two different risk sources: a level component and a trend component, on which this article focuses on.

The internal model process workflow is illustrated in Figure 12: the usual workflow is represented, with the additional step of mortality tables adjustment. Indeed, this correction process is directly performed on national population mortality tables and is then used in the model calibration and selection processes.

The simplified prototype of internal model, employed here to assess longevity risk, relies on several classical stochastic mortality models, which we briefly describe in the following.

The famous model by Lee and Carter (1992), also referred to as M1, decomposes mortality as a static age-structure ($\beta_x^{(1)}$), a general level driven by a stochastic process in time ($\kappa_t^{(2)}$) and an age-specific sensitivity ($\beta_x^{(2)}$) to this general level as follows:

$$\log\left(m(x,t)\right) = \beta_x^{(1)} + \beta_x^{(2)}\kappa_t^{(2)}$$

In order to account for possible cohort effects, while relying on a reasonable parametrisation, one can also consider the age–period–cohort model (M3) as proposed in Currie (2004), which is a special case of Renshaw & Haberman (2006), where the age, period and cohort components influence mortality independently in the following form:

$$\log\left(m(x,t)\right) = \beta_x^{(1)} + \kappa_t^{(2)} + \gamma_{t-x}^{(3)}$$

with γ_c the additional cohort-related factor which allows to adjust mortality rates for a "generation" (or diagonal) which originates from year $c = t - x$.

As an alternative, the model by Plat (2011) combines the Cairns-Blake-Dowd model from Cairns et al. (2006), with some features of the Lee–Carter model to produce a model that is suitable for full age ranges and captures year-of-birth effects, here simplified into

$$\log\left(m(x,t)\right) = \beta_x^{(1)} + \kappa_t^{(1)} + \kappa_t^{(2)}(x-\overline{x}) + \gamma_{ct-x}^{(1)}$$

where $\overline{x}$ denotes the mean age over the range of ages used in the calibration. This formulation can be seen as an extension of the age–period–cohort (M3) model where the joint age and period effect

$$\kappa_t^{(2)}(x-\overline{x})$$

have been added.

To select the most adequate modelling to one set of historical mortality data, each model is calibrated independently in a first phase. Calibration results are compared in a second phase, putting emphasis on a set of necessary or wishful properties, for example

- quality of fit (BIC criterion) of historical mortality tables;
- randomness of standardised residuals;
- stability of model parameters to the age band and historical length of mortality data used for calibration;
- back test of projected mortality improvements against historical improvements.

https://doi.org/10.1017/S1748499520000081 Published online by Cambridge University Press

For further discussion regarding interesting properties and tests for model selection, the reader is referred to, for example, Cairns et al. (2009). Calibration results will be discussed in the next section.

Let us recall that the data used to specify the crude death rates $m(x, t)$ by single year and one-year age bands were downloaded from the HMD on 1 September 2015 for France (civil population), Italy (civil population) and West Germany (total population). As for calibration purposes, the datasets are restricted to the most recent decades of historic, retaining an age band characteristic for a standard insured population. Mortality at higher ages than the maximum of the age band is commonly assessed based on a table closure technique. References and discussions on table closure methodologies can be found in Quashie and Denuit (2005).

3.2 Model fitting

For the purpose of illustration, we restrict our graphical analysis to France female population only, as results are very similar for Italy and West Germany for both genders. Any specific outcome for Italy or Germany will be mentioned. We are considering the population aged from 30 to 95 and years from 1984 to 2009 (Italy), 2011 (Germany) or 2013 (France), which corresponds to HMD availability at the time of the study.

3.2.1 M1 model

The M1 model solely captures age-related and period-related mortality effects. As the retreated exposures mostly correct cohort-based effects, one could expect limited impacts on the M1 model components. As a matter of fact, differences of fit results on separate model components are very moderate, as illustrated below for the French female population. The results are depicted in Figure 13.

3.2.2 M3 model

In addition to age-related (beta) and period-related (kappa) mortality effects, the M3 model also captures specific cohort effects between generations (gamma). We observe that the gamma component shows an improved stability on the specific generations related to shocks in birth patterns around the two World Wars, whereas the beta component (age-related) and kappa component (period-related) show no sensible changes. Changes in the separate model components are restricted to specific generations of the cohort-effect component only.

Female and male populations of each country in our study show similar results to the French female population illustrated in Figure 14.

3.2.3 Plat model

Fitting results for the Plat model are depicted in Figure 15, for which similar conclusions as for the M3 model hold for the common parameters, the main changes being observed for the cohort component. As for the additional $\kappa_t^{(2)}$ parameter, this shows a relative stability from crude to adjusted data.

3.3 Impact on future life expectancy

In this section, we propose to highlight potential changes on life expectancy through two complementary approaches. On the one hand, projected period mortality tables are compared with and without exposure retreatment on the basis of residual period life expectancy at age 30 and truncated at age 95. On the other hand, cohort mortality tables are compared with and without

https://doi.org/10.1017/S1748499520000081 Published online by Cambridge University Press

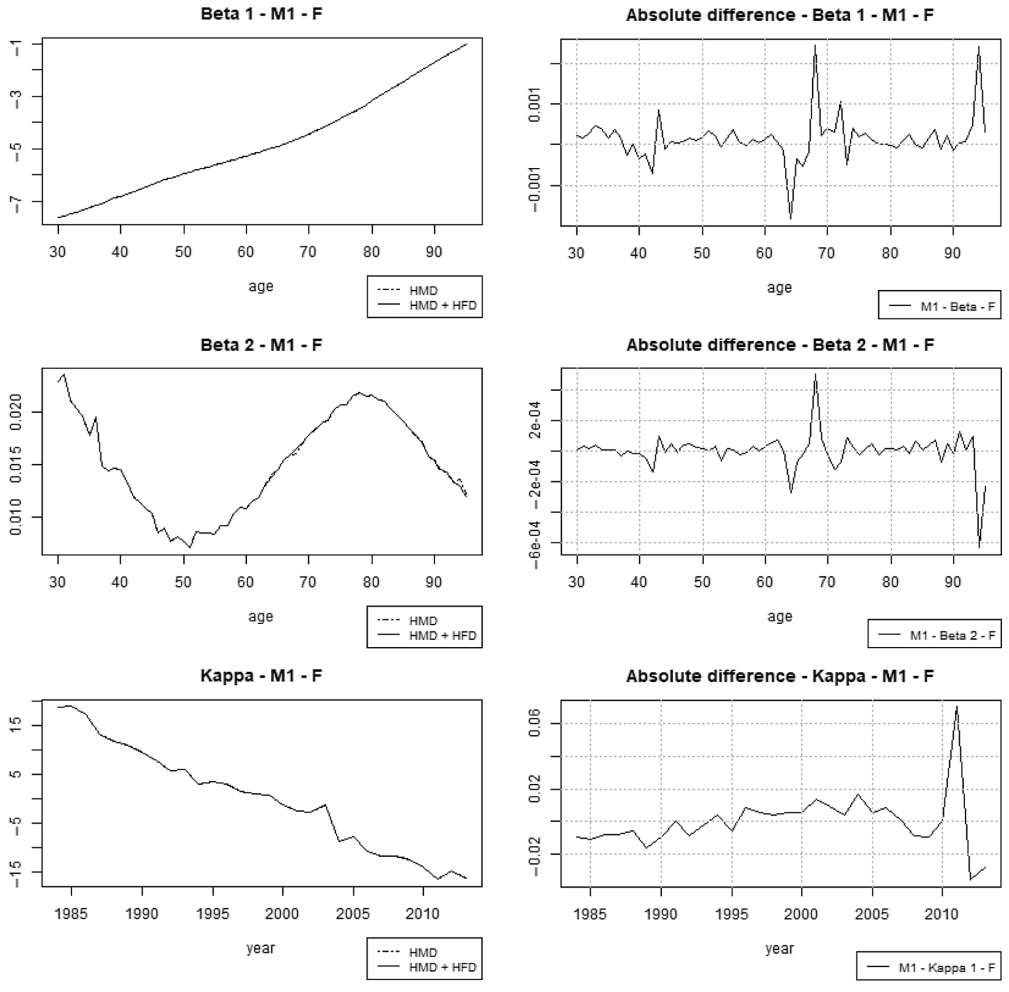

Figure 13. Lee–Carter model (M1) components fitting.

exposure retreatment for all active generations being of age 30 to 95 years in 2015. The figures depicted in this section present the outcome of the models stochastic projection and differences between life expectancies for the median (baseline), 0.5th percentile (mortality shock) and 99.5th percentile (longevity shock) scenarios.

3.3.1 M1 model

As it was expected based on the reduced impact on the fit of M1 model component, life expectancy analysis shows very limited change between both sets of population's exposure data, as depicted in Figure 16. For both genders and for each country in our study, period-based life expectancy between age 30 and 95 as well as live generation life expectancy does not vary more than half a month.

Let us recall that the LEs for non-observed cohorts are obtained by using the projections of each stochastic mortality models; in particular, models including a cohort component rely on a time series forecast of this cohort component.

https://doi.org/10.1017/S1748499520000081 Published online by Cambridge University Press

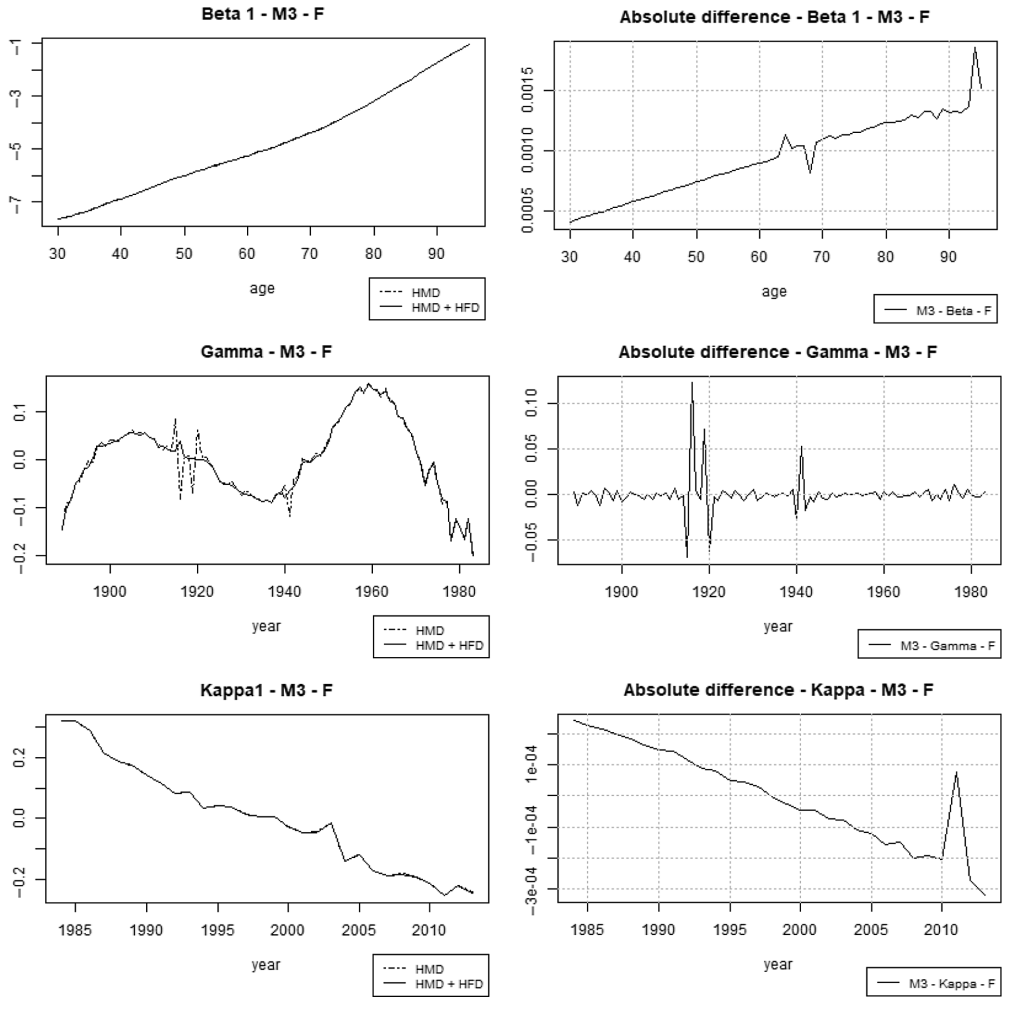

Figure 14. Age–period–cohort model (M3) components fitting.

3.3.2 M3 model

Detailing the impact on the M3 model calibration, we observe in Figure 17 that the projected life expectancy (period-based) is very stable. Cohort life expectancy shows limited variations, with highest impact located during World War II.

Note that for operational constraints to limit the use of the internal model, the Plat model forecasts were not considered in this section.

3.3.3 Conclusion

In a nutshell, the study of the model fits on historical tables shows that M1 model is almost not impacted by the "cohort effect" retreatment, due to the model not capturing the "cohort effect". M3 and Plat model outputs are not significantly impacted by the corrected exposure tables as both period and cohort-based life expectancy keep the same level. However, cohort life expectancy shows evidence of change in the inter-generational structure, and the volatility is reduced in the model calibrated with HMD embedding HFD correction (see again Figure 14).

https://doi.org/10.1017/S1748499520000081 Published online by Cambridge University Press

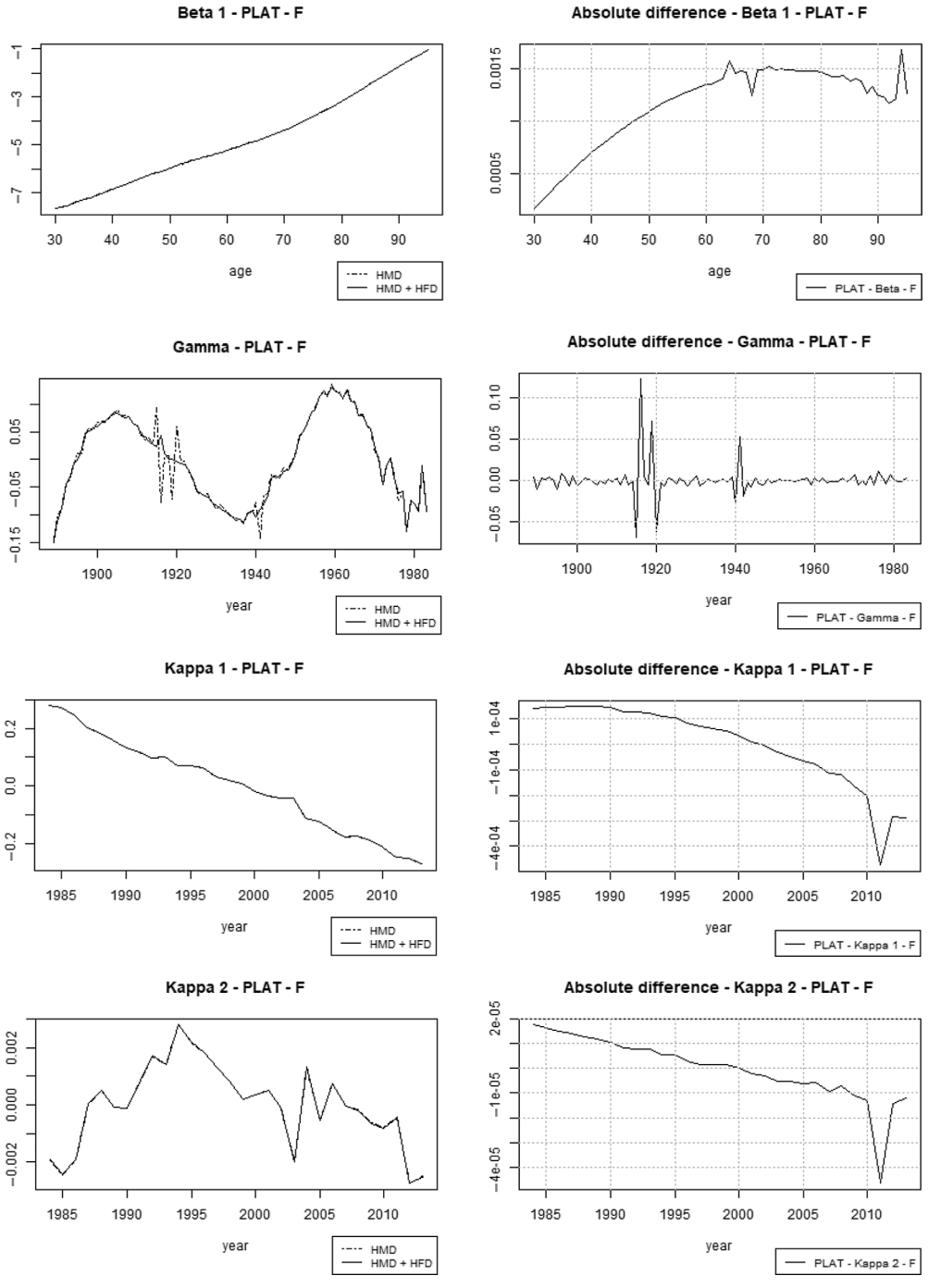

Figure 15. Plat model components fitting.

3.4 Model selection process

As far as the model selection process is concerned, in all cases of our study, the improved BIC criterion as depicted in Figure 18 shows that the retreated exposure allows the Lee–Carter (M1) model to better capture mortality dynamics embedded in HMD table. BIC criterion does not show

https://doi.org/10.1017/S1748499520000081 Published online by Cambridge University Press

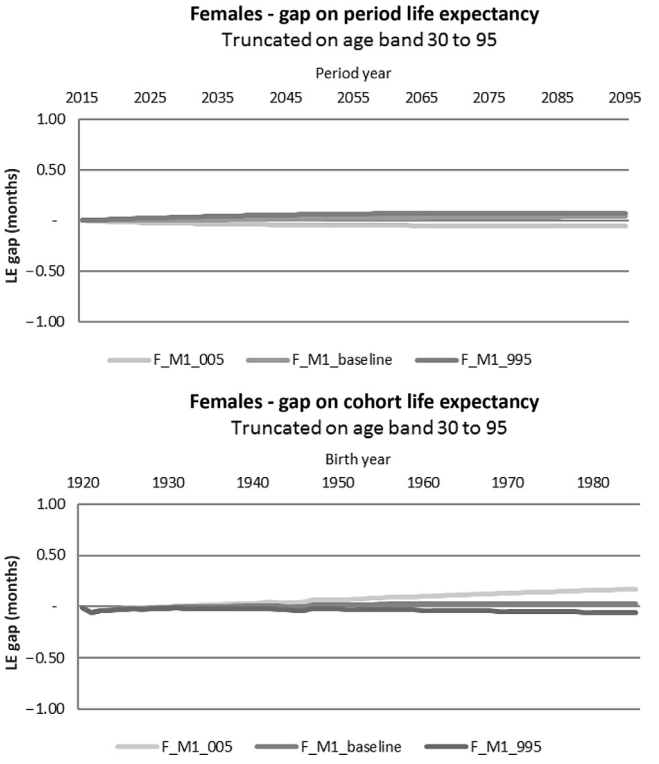

Figure 16. Period and cohort life expectancy as of 2015 for French females, M1 model.

a significant change for the M3 model nor the Plat model. This seems to be explained by the fact that the cohort parameter captures the irregularities of the original mortality tables, leading to relatively unchanged residuals between the originals and the adjusted mortality tables.

The selected model based on a BIC criterion is still the (simplified) Plat model in each configuration. Finally, for several populations as for Italian female table, retreated exposures allow for M1 to achieve a higher BIC level than the M3 model.

As highlighted in the former section, the retreatment of the exposure reduces fluctuations of the component related to the cohort effect. Therefore, models that do not take "cohort effect" into account as the Lee–Carter model improves their efficiency to fit historical mortality data; models that capture "cohort effect" also benefit from an improved stability (less fluctuation) in the cohort parameter estimate.

3.5 Solvency capital requirement

The trend component of the longevity SCR (called "SCR" in the following paragraphs) is here given as a change of future mortality improvements. The model described previously gives a full distribution of mortality improvement rate by gender, age and year. Note that the approach taken here does not consider a run-off distribution in order to remain compliant with the Solvency 2 framework, nor corresponds to taking the percentile each year, which woud be indeed too conservative. Also the methodology is based on a key indicator[3] that allows to select the 0.5th percentile of the distribution of longevity improvements, denoted $IR^{SCR}_{x,t_0\to t}$. To obtain the final mortality table for SCR calculations, the improvements of the best estimate assumptions

[3] Note again that due to confidentiality purposes, we are limited on developing some technical details.

https://doi.org/10.1017/S1748499520000081 Published online by Cambridge University Press

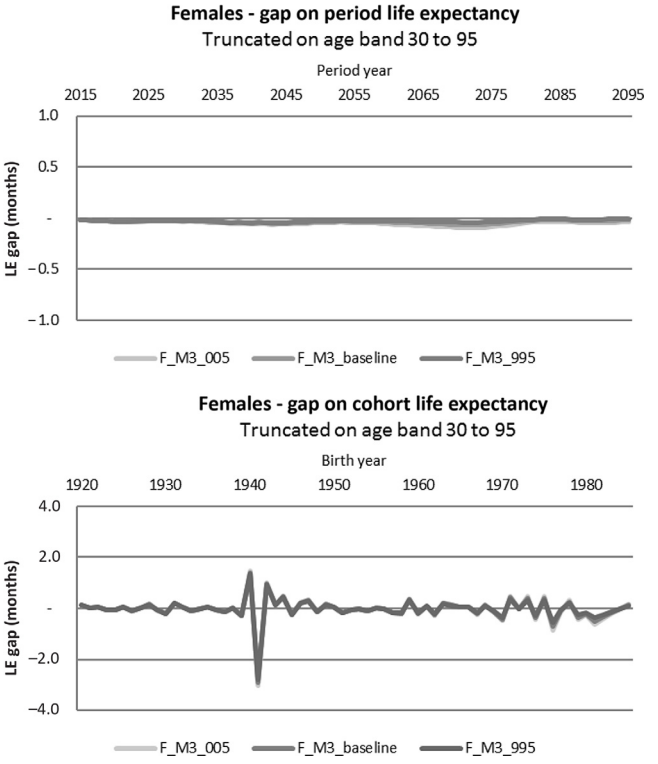

Figure 17. Period and cohort life expectancy as of 2015 for French females, M3 model.

$$IR^{BE}_{x,t_0\to t} = \frac{q^{BE}_{x,t} - q^{BE}_{x,t_0}}{q^{BE}_{x,t_0}},\ for\ t > t_0$$

are replaced by the ones obtained from the model mentioned above for all years from the valuation year t_0 onwards. The initial mortality rates of the trend SCR component for year t_0 do not change from the best estimate mortality rates for this same year. For future years $t > t_0$, the mortality rates are defined as follows for the best estimate ($q^{BE}_{x,t}$) and the SCR ($q^{SCR}_{x,t}$), respectively:

$$q^{BE}_{x,t} = q^{BE}_{x,t_0} \cdot \left(1 + IR^{BE}_{x,t_0\to t}\right)$$

$$q^{SCR}_{x,t} = q^{BE}_{x,t_0} \cdot \left(1 + IR^{SCR}_{x,t_0\to t}\right)$$

We depict in Figure 19 the impact on mortality rate scenarios of the change of data source (HMD data with HFD adjustment). As the improvements distribution is changed with the data adjustment, we observe in Figure 19 a slight impact on mortality rates due to lower cohort effects, as stated previously during the calibration of mortality models. Note that both the baseline scenario ("Base") and the SCR scenario are changed in this figure.

As stated previously, there is less volatility in mortality assumptions when using the HFD adjustment. The level of mortality rates is, however, not impacted. The best estimate used are already without this excess of volatility; therefore, in the following analysis, we will only focus on the impact of SCR from changing the HMD data with HFD-adjusted data while keeping the same best estimate assumptions.

This decrease of captured volatility in the SCR scenario can be seen for various different cohorts. This can be illustrated with the impact of longevity trend SCR on the evolution of the

https://doi.org/10.1017/S1748499520000081 Published online by Cambridge University Press

France Female	BIC criterion		
	HMD	HMD + HFD	% diff.
M1	27852.71	24176.08	–13%
M3	25166.31	25145.91	0%
PLAT	23567.71	23548.09	0%

France Male	BIC criterion		
	HMD	HMD + HFD	% diff.
M1	34279.57	30656.00	–11%
M3	27075.86	27048.32	0%
PLAT	24994.46	24968.62	0%

Italy Female	BIC criterion		
	HMD	HMD + HFD	% diff.
M1	25629.59	21994.51	–14%
M3	22118.70	22164.95	0%
PLAT	19994.71	20044.77	0%

Italy Male	BIC criterion		
	HMD	HMD + HFD	% diff.
M1	31831.51	28074.98	–12%
M3	24512.84	24571.27	0%
PLAT	21816.25	21879.49	0%

Germany Female	BIC criterion		
	HMD	HMD + HFD	% diff.
M1	32288.22	25752.47	–20%
M3	24082.62	24478.00	2%
PLAT	22413.18	22809.65	2%

Germany Male	BIC criterion		
	HMD	HMD + HFD	% diff.
M1	31901.24	27637.65	–13%
M3	23329.90	23774.33	2%
PLAT	22634.69	23079.54	2%

Figure 18. BIC criteria for model calibration.

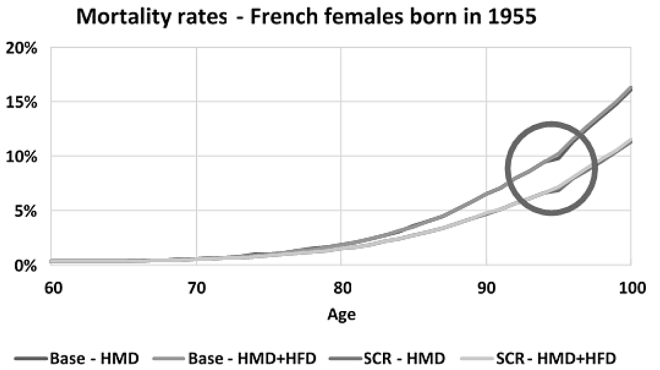

Figure 19. Baseline and longevity trend SCR future mortality for French females born in 1955.

cohort life expectancy. The cohort life expectancy including future mortality improvements is defined as follows for $A \in \{BE,\ SCR\}$:

$$e_{x,t}^{\text{Cohort}}(A) = \sum_{k=1}^{\infty} \prod_{i=0}^{k-1} \left(1 - q_{x+i,t+i}^{A}\right)$$

The impact on cohort life expectancy from replacing the best estimate improvements by the SCR improvements is defined as follows:

$$IE_{x,t} = \frac{e_{x,t}^{\text{Cohort}}(SCR) - e_{x,t}^{\text{Cohort}}(BE)}{e_{x,t}^{\text{Cohort}}(BE)}$$

The impact is illustrated in Figure 20 for French females in 2015 born between 1925 and 1975, when having on one hand the best estimate assumptions with the improvement remaining the same, and on the other hand mortality rates using longevity SCR improvements. For the latter, we

https://doi.org/10.1017/S1748499520000081 Published online by Cambridge University Press

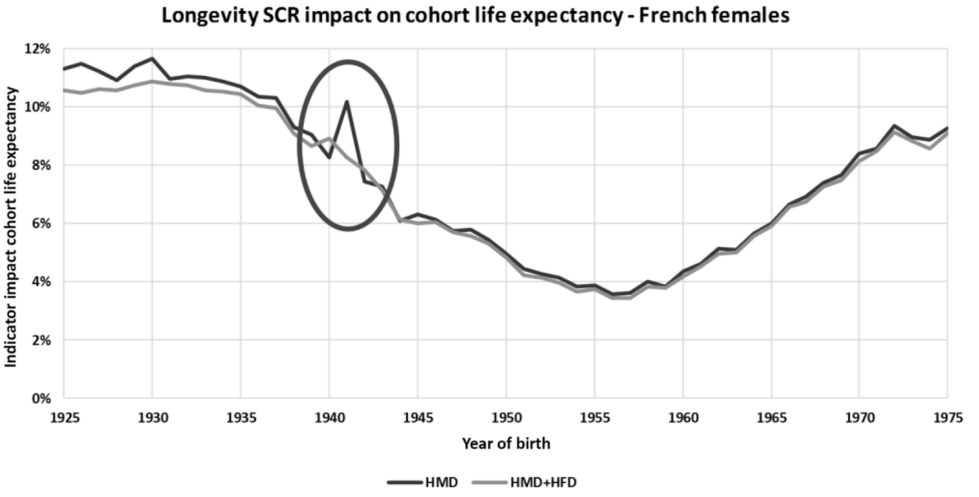

Figure 20. Longevity SCR impact on cohort life expectancy in 2015 for French females born between 1925 and 1975.

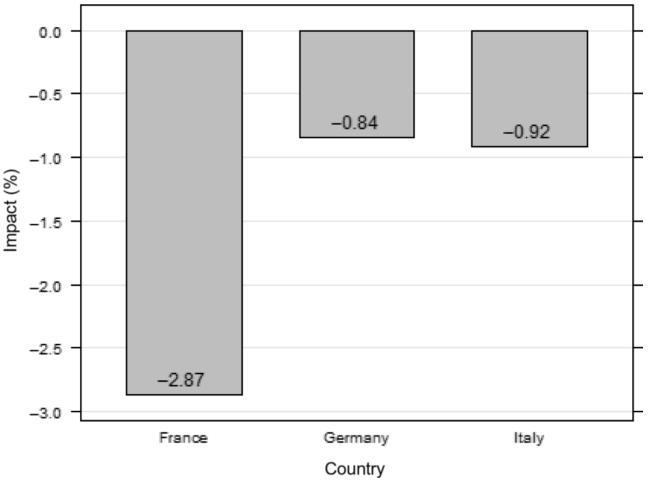

Figure 21. Impact on longevity trend SCR.

perform the calculations with the data from HMD only, and with the HMD data adjusted with the HFD data.

With the change of data, we observe a lower volatility on the indicator $IE_{x,t}$ for some cohorts, notably females being born in the early 1940s during World War II, that is, being between 70 and 75 in 2015. There is a small difference for older people (those being born between the World Wars) and almost no difference for those being born after 1945.

In order to compute the final impact depicted in Figure 21, the portfolios used for the study are originated from the three countries France, West Germany and Italy, including various products. For France, the portfolio is made of a savings and retirement business made of individual and group products, with on the in-force an average age around 65/70 for the annuity phase. For Germany, these are individual investment and savings products, with an average age of policyholders around 47. For Italy, the products are investment and savings products (62%) and long-term care products (38%) exposed to longevity risk. The savings and retirement businesses for these three countries are a mix of general account, with various guaranteed rates,

https://doi.org/10.1017/S1748499520000081 Published online by Cambridge University Press

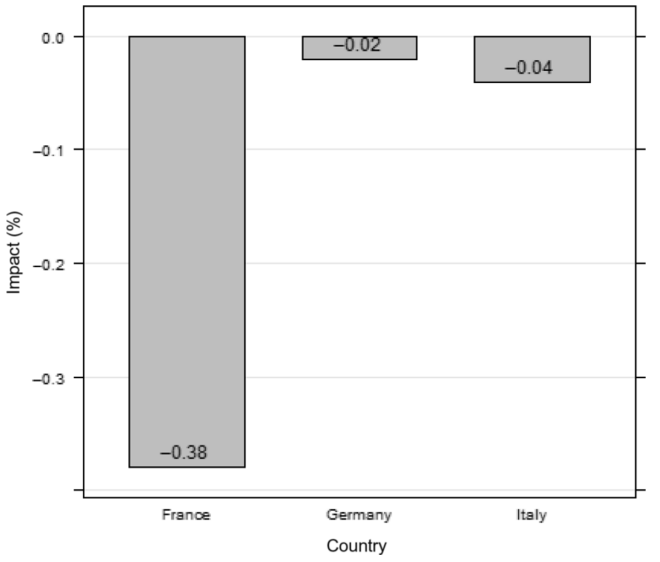

Figure 22. Impact on life SCR.

and unit-linked products for the accumulation phase. Long-term care are a pure general account products.

Using the full internal model methodology and the aggregation with other risks, with on the one hand the HMD data and on the other hand the HMD data adjusted with HFD, we observe that the adjustment leads to a small reduction of longevity SCR (Figure 21).

This negative impact is consistent with the results found previously on indicator $IE_{x,t}$: the removal of non-necessary volatility on several cohorts leads to a decrease of the longevity SCR. This is even clearer for the French portfolio, when compared to the German and Italian portfolios, as the French portfolio contains more annuitants close to ages 70–75 at which the reduction of non-necessary volatility is the strongest.

Looking now at the total life SCR of each of these entities with the other risks and the diversification between these risks calculated with an internal model methodology, the impacts on total life SCR are negative but almost non-significant (Figure 22) even as longevity SCR roughly represents 20% life SCR pre-diversification.

Even if the quantitative impacts are very moderate, the treatment of HMD data with the HFD has a positive effect on the data quality and therefore on the longevity risk assessment due to a lower volatility in future mortality assumptions.

3.6 Stability of the longevity risk assessment

The stability from a year to another of the longevity risk assessment can be roughly estimated by looking at the evolution of the impact of longevity trend SCR mortality improvements on the cohort life expectancy of the portfolio, that is, the difference for a policyholder at age x in year t:

$$e_{x,t}^{\text{Cohort}}\,(\text{SCR}) - e_{x,t}^{\text{Cohort}}\,(\text{BE})$$

and its evolution between t and $t+1$. This indicator is calculated on the whole portfolio mix. There again, the best estimate does not change but SCR mortality improvements change as we consider the improvements calibrated with HMD data, and also the improvements calibrated with HMD retreated with HFD data.

https://doi.org/10.1017/S1748499520000081 Published online by Cambridge University Press

Evolution with SCR scenario	France	Germany	Italy
HMD data	−2.31%	0.40%	−1.01%
HMD + HFD data	0.26%	1.35%	0.06%

Figure 23. Evolution of cohort life expectancy gap between baseline and longevity SCR trend between 2015 and 2016.

As shown in Figure 23, the retreatment brings a more stable longevity risk assessment between 2015 and 2016 for France and Italy but not for Germany, notably due to women. Overall, the volatility of longevity SCR shock over time decreases.

4. Concluding Remarks

To illustrate it with a caricature, even if an internal model is the most robust one from a technical standpoint and is perfectly adapted to the inherent risks of a company, the "garbage in, garbage out" effect has to be avoided. In a global context of increasing regulatory needs, particularly on the data quality, using such a methodology to improve the data of external providers, as it is the case for the HMD, is key. This database is regarded as a reference on the insurance market to have national mortality data on a large scope of countries, with similar demographic methodologies used from one country to another. Implementing the described methodology to correct data anomalies can therefore help the insurance market.

Risks are consequently better captured, assessed and monitored. The conclusion of the study and the practical application show that the longevity assessment remains stable as the global impact is non-significant on the tested portfolios but indicators are improved; with a better data quality the volatility coming from these exceptional cohort effects is lowered, here slightly decreasing the capital requirement.

This kind of database cleansing is obviously welcomed and can help demographers and actuaries to better monitor longevity and mortality risks. In the future, it is important to continue tracking and correcting other kinds of errors that can be included in the current mortality tables.

Acknowledgements. The authors are grateful to the anonymous referee for valuable comments and advice. The authors also thank Amal Elfassihi from Milliman Paris for his help in improving the first version of the paper.

References

Boumezoued, **A.** (2020). Improving HMD mortality estimates with HFD fertility data. *North American Actuarial Journal*, 1–25.

Boumezoued, **A.** & **Devineau**, **L.** (2017). Enjeux de fiabilité dans la construction des tables de mortalité nationales. *L'Actuariel, janvier 2017*.

Cairns, **A.J.G.**, **Blake**, **D.** & **Dowd**, **K.** (2006). A two-factor model for stochastic mortality with parameter uncertainty: theory and calibration. *Journal of Risk and Insurance*, **73**(4), 687–718.

Cairns, **A.J.G.**, **Blake**, **D.**, **Dowd**, **K.**, **Coughlan**, **G.D.**, **Epstein**, **D.**, **Ong**, **A.** & **Balevich**, **I.** (2009). A quantitative comparison of stochastic mortality models using data from England and Wales and the United States. *North American Actuarial Journal*, **13**(1), 1–35.

Cairns, **A.J.G.**, **Blake**, **D.**, **Dowd**, **K.** & **Kessler**, **A.R.** (2016). Phantoms never die: living with unreliable population data. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, **179**(4), 975–1005.

CEIOPS' Advice for Level 2 Implementing Measures on Solvency II - Annex B Longevity risk calibration analysis. CEIOPS-DOC-42/09.

Currie, **I.D.**, **Durban**, **M.** & **Eilers**, **P.H.C.** (2004). Smoothing and forecasting mortality rates. *Statistical modelling*, **4**(4), 279–298. https://doi.org/10.1191/1471082X04st080oa

Human Fertility Database. Max Planck Institute for Demographic Research (Germany) and Vienna Institute of Demography (Austria).). Available online at the address www.humanfertility.org [data downloaded on Jul-2016].

Human Mortality Database. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available online at the address www.mortality.org or www.humanmortality.de [data downloaded on 1-Sep-2015].

https://doi.org/10.1017/S1748499520000081 Published online by Cambridge University Press

Lee, **R.D.** & **Carter**, **L.R.** (1992). Modeling and forecasting US mortality. *Journal of the American Statistical Association*, **87**(419), 659–671.

Plat, **R.** (2011). One-year value-at-risk for longevity and mortality. *Insurance: Mathematics and Economics*, **49**(3), 462–470.

Quashie, **A.** & **Denuit**, **M.** (2005). *Modèles d'extrapolation de la mortalité aux grands âges*. Institut des Sciences Actuarielles et Institut de Statistique, Université Catholique de Louvain.

Renshaw, **A.E.** & **Haberman**, **S.** (2006). A cohort-based extension to the Lee–Carter model for mortality reduction factors. *Insurance: Mathematics and Economics*, **38**(3), 556–570.

Richards, **S.J.** (2008). Detecting year-of-birth mortality patterns with limited data. *Journal of the Royal Statistical Society, Series A*, **171**(1), 279–298.

Wilmoth, **J.R.**, **Andreev**, **K.**, **Jdanov**, **D.** & **Glei**, **D.A.** (2017). *Methods Protocol for the Human Mortality Database*. University of California, Berkeley, and Max Planck Institute for Demographic Research, Rostock. Available online at the address http://mortality.org [accessed 27-Nov-2017].

Cite this article: Balland F, Boumezoued A, Devineau L, Habart M and Popa T (2020). Mortality data reliability in an internal model. *Annals of Actuarial Science* **14**, 420–444. https://doi.org/10.1017/S1748499520000081