Fachbereich II – Mathematik - Physik - Chemie

BEUTH HOCHSCHULE FÜR TECHNIK BERLIN
University of Applied Sciences

01/2014

Karl Michael Ortmann

Forecasting mortality patterns

Prognose von Sterblichkeitsmustern (englischsprachig)

Reports in Mathematics, Physics and Chemistry
Berichte aus der Mathematik, Physik und Chemie

ISSN (print): 2190-3913

Reports in Mathematics, Physics and Chemistry

Berichte aus der Mathematik, Physik und Chemie

The reports are freely available via the Internet:
http://www1.beuth-hochschule.de/FB_II/reports/welcome.htm

01/2014, February 2014

Forecasting mortality patterns
Prognose von Sterblichkeitsmustern (englischsprachig)

Editorial notice / Impressum

Published by / Herausgeber:
Fachbereich II
Beuth Hochschule für Technik Berlin
Luxemburger Str. 10
D-13353 Berlin
Internet: http://public.beuth-hochschule.de/FB_II/
E-Mail: fbiireports@beuth-hochschule.de

Responsibility for the content rests with the author(s) of the reports.
Die inhaltliche Verantwortung liegt bei den Autor/inn/en der Berichte.

ISSN (print): 2190-3913

Abstract
In this study we analyse outcomes of the stochastic mortality forecast model developed by Ortmann (2013). In a retrospective study we apply several backtests in order to evaluate ex-post predictive power of the model. Moreover, we evaluate plausibility of ex-ante forecasts. In addition, we conduct a comparison to other standard forecast models, such as those proposed by Cairns et al. (2009), Haberman and Renshaw (2011) as well as Lee and Carter (1992). Thereby, we discuss the plausibility criterion for forecasting mortality that requires changes in mortality by age over time to be smooth across time and birth cohorts.

Key Words: age, period and cohort effect; stochastic mortality model; mortality shock; mortality stress scenario

Zusammenfassung
In diesem Artikel untersuchen wir die Ergebnisse des Sterblichkeitsprognosemodells von Ortmann (2013). In Simulationen auf der Basis von historischen Daten wird die ex-post Vorhersagegüte des Modells analysiert. Außerdem werden konkrete ex-ante Vorhersagen auf ihre Plausibilität untersucht. Nicht zuletzt werden Vergleiche mit den aus der Literatur bekannten Standardmodellen nach Cairns und anderen (2009), Haberman und Renshaw (2011) sowie Lee und Carter (1992) angestellt. Dabei wird besonders auf das Plausibilitätskriterium eingegangen, welches verlangt, dass Sterblichkeitsveränderungen in Richtung Kalenderjahr und Geburtsjahr möglichst glatt sein sollen.

1 Introduction

Recently, Ortmann (2013) applied certain techniques of computer vision and image processing to the detection of specific patterns in mortality surfaces. In particular, he adapted the well-known Canny operator (Canny 1986) to reliably detect age, period and cohort effects in a given matrix of historic mortality data. These effects are marked as abrupt local changes in value of incremental mortality development factors, also called improvement rates in practice. They indicate the boundaries between areas of distinctly higher and lower mortality improvements.

In fact, the process proposed by Ortmann (2013) is a multistage algorithm. The procedure is useful for descriptive and exploratory data analysis in demography and actuarial science. Based on the objective detection of age, period and cohort effects, Ortmann (2013) proposed a stochastic mortality forecast model. Mortality patterns can be suitably identified and removed from the rest of the data. Subsequently, residuals and mortality effects were forecasted separately into the future. In this study, the quality of the model's forecasts is measured ex-post as well as ex-ante.

Our paper is organised as follows: we study stochastic mortality forecasts as part of a retrospective study in section two. Moreover, we apply several backtesting techniques in order to evaluate quality of ex-post forecasts. In section three we assess future mortality probabilities as well as life expectancies at birth and normal retirement age.

2 Ex-post analysis

Let x denote age, t time, $m(x,t)$ be the mortality rate relating to death between ages x and $x+1$ between time t and $t+1$. Further, let $l(x,t)$ be the corresponding number of persons alive. Both age and time shall have a well-defined range of admissible values. Then, Ortmann (2013) defined incremental mortality development factors $IF(x,t)$ by

$$IF(x,t) = \frac{m(x,t)}{m(x,t-1)} - 1 \tag{1}$$

as the basis from which to identify mortality patterns such as age, period and cohort effects by means of an adaptation of the well known Canny algorithm (Canny 1986). In particular, it can be shown (Ortmann 2013) that weighted development factors $\overline{IF}(x,t)$ defined by

$$\overline{IF}(x,t) = IF(x,t) \cdot \sqrt{\frac{l(x,t-1) \cdot m(x,t-1)}{1-m(x,t-1)}} \tag{2}$$

may be assumed to be standard normal distributed across time given the mortality rate in the proceeding period.

In order to conduct a goodness of fit analysis Ortmann (2013) took a simulated random sample of incremental development factors $IF(x,t)$ as well as a generic pattern of mortality effects $ME(x,t)$. In particular, he designed various generic period and cohort effects of different strength. He then superimposed these two matrices by adding up the two rates in each cell, i.e. superimposed rates were calculated as $SR(x,t) = ME(x,t) + IF(x,t)$. By so doing, he was able to derive quantitative criteria to objectively detect age, period and cohort effects.

We now proceed in the opposite direction: Let a matrix of weighted incremental mortality development factors $\overline{IF}(x,t)$ derived from historic data be given. We then identify mortality effects $\overline{ME}(x,t)$ to obtain residual mortality development factors $\overline{RF}(x,t)$ by subtraction: $\overline{RF}(x,t) = \overline{IF}(x,t) - \overline{ME}(x,t)$.

In essence, Ortmann (2013) suggested forecasting future mortality by a bootstrap model. He argued that, apart from specific mortality patterns which constitute a certain correlation structure, incremental mortality development factors cannot be assumed to be independent across time. The reason is that adjacent values of $IF(x,t)$ and $IF(x,t+1)$ depend on the same data, i.e. $m(x,t)$. Hence, incremental mortality development factors are correlated across time. It is necessary to preserve this type of serial correlation that is inherent in the model.

First of all, the correlation structure consisting of age, period and cohort effects was removed from the data by objectively detecting age, period and cohort effects. Thus, it is reasonable to assume that there is no systematic change in mean and variance left in the set of residual mortality development factors $\overline{RF}(x,t)$.

The bootstrapping approach applied by Ortmann (2013) consisted of resampling residual mortality development factors taken from the data set $\overline{RF}(x,t)$. Ortmann (2013) suggested building blocks for each age x_0 with reference to historically observed residual development factors for that age only, i.e. he considered the vector $\left(\overline{RF}(x_0,t_1),\ldots,\overline{RF}(x_0,t_n)\right)$ over a suitably defined time interval $[t_1,t_n]$. The time series of subsequent realisations for the same age over time within the vector $\left(\overline{RF}(x_0,t_1),\ldots,\overline{RF}(x_0,t_n)\right)$ captures all systematic correlations inherent in the data. This serial dependence across time for each age is crucial for forecasting purposes. Future mortality development factors were then generated by sampling with replacement from the specified set of historic data for each age separately. This exercise was repeated a large number of times in order to derive moments and percentiles of simulations.

Dowd et al. (2010b) developed a backtesting framework for stochastic mortality modeling. Accordingly, we distinguish two backtest approaches that differ by defining the historic time horizon, the so called lookback window.

In a rolling horizon backtest we consider a fixed lookback window of thirty years of historic data that is used to predict the following year's mortality development factor. This window is then sequentially moved forward in time to only predict one year into the future. Hence, the lookforward window always is one year.

In a fixed horizon backtest we consider a fixed lookback window of thirty years of mortality experience. We then use the bootstrap method to forecast mortality development factors over the next twenty years. Here, the lookforward window is twenty years.

2.1 Age, period and cohort effects

Ortmann (2013) took mortality rates and exposures for English & Welsh males from the human mortality database (2010) for ages 0 – 99 and fifty years of experience: 1959 – 2008. It would have certainly been feasible and possibly quite convenient for practitioners to apply a pragmatic and judgmental approach with a view to identifying mortality patterns in the data. Instead, Ortmann (2013) applied his mortality effects detection algorithm in order to objectively identify age, period and cohort effects. Having said this, relevant mortality patterns may be alternatively established on the basis of qualified judgments.

The matrix of mortality effects entailed 4 cohort effects and 23 period effects. The latter relate to older ages only. Possible explanations for period effects may be related to strong and mild winters, summer heat waves, flu epidemics, etc. The main cohort effect of mortality improvements relates to birth cohorts 1927-1936. Cutler et al. (2006) suggest that medical advances had a particular effect on the reduction of mortality from cardiovascular diseases for these particular cohorts. Ortmann (2013) modelled a linear trend over time within this effect. The same modelling approach applies to the cohort effect for people born in 1920. The main fitting results are as follows:

	$\overline{IF}(x,t)$	$\overline{RF}(x,t)$
expected value	-0.643	-0.015
standard deviation	2.653	1.717
skewness	0.339	0.104
kurtosis	3.861	0.146

Table 1: First four moments for initial weighted mortality development factors $\overline{IF}(x,t)$ and weighted residual mortality development factors $\overline{RF}(x,t)$.

Interestingly, all cohort effects combined imply a reduction in expected value from -0.643 to -0.287. Therefore, they account for a reduction of about 55% of total mortality improvements experienced over the last fifty years across all ages. Furthermore, standard deviation is reduced by about 43% to 1.505. Hence, cohort effects alone do not explain total variability of mortality development factors well enough. Moreover, the hypothesis on normality would have to be firmly rejected if we model cohort effects only.

All period effects combined reduce the mean from -0.643 to -0.322, resulting in a reduction of about 50%. Standard deviation is reduced to by about 42% to 1.538. Here again, residual mortality development factors cannot be regarded standard normal distributed. Period effects alone do not fully explain changes in mortality over time.

The model fit includes a fair number of period effects as well as cohort effects and one age effect. Whilst the mean has come to zero standard deviation on the other hand has remained bigger than one.

2.2 Rolling one-year analysis

Dowd et al. (2010a) set out a framework to evaluate the goodness of fit of stochastic models to historic mortality data. Specifically, they suggested applying a number of statistical tests, which we adapt to our purposes. In particular, we consider a backtest that involves a lookback window of thirty years to predict one year into the future. Furthermore, this window rolls forward through time to predict one year after another. This approach is in line with actuarial practice where mortality forecasts are updated each year. Notably, solvency capital requirements are also partly based on one-year value-at-risk calculations (Plat 2010). Hence, a one-year backtest analysis is particularly relevant for Solvency II purposes.

We conduct a bootstrap simulation comprising 1.000 runs. For each age x_0 and each calendar year t_0 we consider the set $\left(\overline{RF}(x_0,t_{0-30}),\ldots,\overline{RF}(x_0,t_{0-1})\right)$ or equivalently $\left(RF(x_0,t_{0-30}),\ldots,RF(x_0,t_{0-1})\right)$. Future mortality development factors are then generated by sampling with replacement from the specified set of historic data for each age and each calendar year separately.

Keilmann (1998) and Koissi et al. (2006) focussed on ex post forecast errors of mortality predictions. Lee and Miller (2001) conducted a number of tests to evaluate the ex post accuracy of mortality forecasts for the Lee Carter method (Lee Carter 1992) in particular in various graphical ways. Cairns et al. (2009) as well as Dowd et al. (2010a) suggested analysing standardised mortality residuals. Here, we use the notation which is based on actual death count versus expected death count. Mortality probabilities are given by

$$m(x,t) = m(x,t-1)\cdot\left(1+RF(x,t)\right)$$

It follows that actual deaths can be computed by

$$d(x,t) = l(x,t)\cdot m(x,t)$$

where $l_{x,t}$ is the exposure of people aged x at year t. Likewise, the number of expected deaths is given by

$$\hat{d}(x,t) = l(x,t)\cdot \hat{m}(x,t)$$

where estimated mortality is

$$\hat{m}(x,t) = m(x,t-1)\cdot\left(1+\hat{RF}(x,t)\right) .$$

Notably, exposures do not need to be estimated since the forecast horizon is one year only. Mortality residuals can thus be defined by

$$MR(x,t) = \frac{d(x,t)-\hat{d}(x,t)}{\sqrt{\hat{d}(x,t)}} = \sqrt{l(x,t)\cdot m(x,t-1)}\cdot\frac{RF(x,t)-\hat{RF}(x,t)}{\sqrt{1+\hat{RF}(x,t)}} ,$$

Figure 1 shows the pattern of standardised mortality residuals by age and forecast year. As Koissi et al. (2006) and Cairns et al. (2009) stressed the patterns should appear random. Any patterns that may be spotted in Figure 1 indicate the existence of further age, period or cohort effects that have not been removed from the initial data set of mortality development factors $\overline{IF}(x,t)$

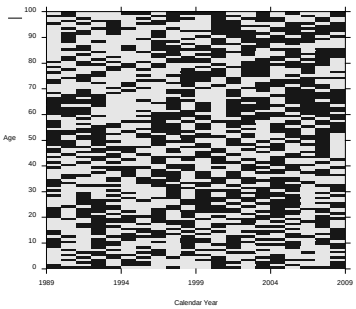

Figure 1: Plot of standardised mortality residuals $MR(x,t)$ using 1.000 bootstrap simulations. Bright areas indicate that mortality residuals are negative, and dark areas indicate positive mortality residuals.

Haberman and Renshaw (2011) contended that mortality residuals should also be plotted by age, period and cohort separately. Therefore, we formally report mean and standard deviation of mortality residuals by age group and calendar year band in Tables 2 and 3.

age / year	1989-1993	1994-1998	1999-2003	2004-2008	Total
0-9	-0.32	0.57	0.29	0.29	0.21
10-19	-0.10	0.15	0.02	0.27	0.08
20-29	0.22	0.22	-0.44	-0.07	-0.02
30-39	0.36	0.09	-0.18	0.15	0.11
40-49	-0.20	0.35	0.07	-0.09	0.03
50-59	-0.09	-0.39	0.17	-0.38	-0.17
60-69	-0.40	0.87	-0.05	-2.12	-0.43
70-79	0.77	0.43	-0.62	0.19	0.19
80-89	1.22	-0.03	-0.21	-0.57	0.10
90-99	0.38	0.37	-0.22	-0.87	-0.08
Total	0,18	0,26	-0,12	-0,32	**0,00**

Table 2: mean absolute values of mortality residuals $MR(x,t)$ by age group and calendar year band based on 1.000 simulations.

age / year	1989-1993	1994-1998	1999-2003	2004-2008	Total
0-9	1.72	1.74	1.66	1.53	1.68
10-19	1.52	1.82	1.51	1.78	1.66
20-29	1.81	1.69	1.45	1.66	1.66
30-39	1.41	1.63	1.73	1.58	1.59
40-49	1.94	2.07	1.75	1.93	1.92
50-59	2.07	2.32	2.94	2.49	2.47
60-69	2.20	2.26	2.98	2.84	2.79
70-79	2.24	2.30	2.23	3.14	2.54
80-89	2.01	1.89	2.60	2.92	2.47
90-99	1.76	1.53	1.61	2.86	2.06
Total	1.93	1.96	2.12	2.44	**2.13**

Table 3: standard deviations of mortality residuals $MR(x,t)$ using 1.000 bootstrap simulations by age group and calendar year band based on 1.000 simulations.

Overall, the standard deviation of $MR(x,t)$ turns out to be 2.13 which is lower than the standard deviation of mortality residuals associated with those mortality forecast models analysed by Cairns et al. (2009).

As part of an additional analysis motivated by Lee and Miller (2001), we evaluate forecasting performance with reference to actual mortality versus estimated. Estimated death counts $\hat{d}(x,t)$ can also be computed using mortality residuals $MR(x,t)$ by solving a quadratic equation (Koissi 2006):

$$\hat{d}(x,t) = \left(-\frac{MR(x,t)}{2} + \sqrt{d(x,t) + \frac{MR^2(x,t)}{4}} \right)^2 .$$

Estimated mortality probabilities $\hat{m}_{x,t}$ are then calculated by

$$\hat{m}(x,t) = \frac{\hat{d}(x,t)}{l(x,t)}$$

Booth et al. (2006) focussed on absolute errors in log mortality. In contrast, we compute average percentage errors and the percentage of mortality estimations that are lower than actual mortality in Tables 4 and 5.

age / year	1989-1993	1994-1998	1999-2003	2004-2008	Total
0-9	3.2%	-3.6%	-0.8%	0.6%	-0.1%
10-19	1.2%	-0.7%	0.6%	-1.5%	0.0%
20-29	-0.9%	-1.0%	2.8%	0.7%	0.2%
30-39	-1.6%	-0.3%	0.9%	-0.6%	-0.4%
40-49	0.8%	-0.9%	-0.2%	0.4%	0.0%
50-59	0.0%	0.7%	-0.1%	0.6%	0.3%
60-69	0.5%	-0.9%	0.0%	2.3%	0.5%
70-79	-0.6%	-0.4%	0.5%	-0.1%	-0.1%
80-89	-1.1%	0.1%	0.2%	0.5%	0.0%
90-99	-0.8%	-0.9%	0.5%	2.0%	0.6%
Total	-0.3%	-0.4%	0.3%	0.7%	**0.1%**

Table 4: average percentage errors relating to estimated mortality over actual mortality by age group and calendar year band based on 1.000 simulations.

age / year	1989-1993	1994-1998	1999-2003	2004-2008	Total
0-9	50.0%	36.0%	46.0%	44.0%	44.0%
10-19	46.0%	54.0%	48.0%	50.0%	49.5%
20-29	52.0%	42.0%	62.0%	54.0%	52.5%
30-39	38.0%	52.0%	54.0%	50.0%	48.5%
40-49	54.0%	42.0%	44.0%	48.0%	47.0%
50-59	52.0%	56.0%	48.0%	62.0%	54.5%
60-69	58.0%	38.0%	50.0%	74.0%	55.0%
70-79	34.0%	40.0%	58.0%	48.0%	45.0%
80-89	30.0%	48.0%	54.0%	56.0%	47.0%
90-99	40.0%	38.0%	62.0%	58.0%	49.5%
Total	45.4%	44.6%	52.6%	54.4%	**49.3%**

Table 5: percentage of estimated mortality predictions that are bigger than actual ones by age group and calendar year band based on 1.000 simulations.

The overall mean percentage error is rather small. On average, about half the time, estimated mortality was bigger than actual mortality. There is a slight bias towards overestimation especially for older ages in later calendar years.

Finally, Cairns et al. (2008) demanded that any valuable mortality forecast model should be able to produce sample paths and produce prediction intervals. Koissi et al. (2006) described in detail how bootstrap percentile intervals may be used to derive prediction intervals. Figure 2 illustrates sample paths and simulated prediction intervals for selected ages.

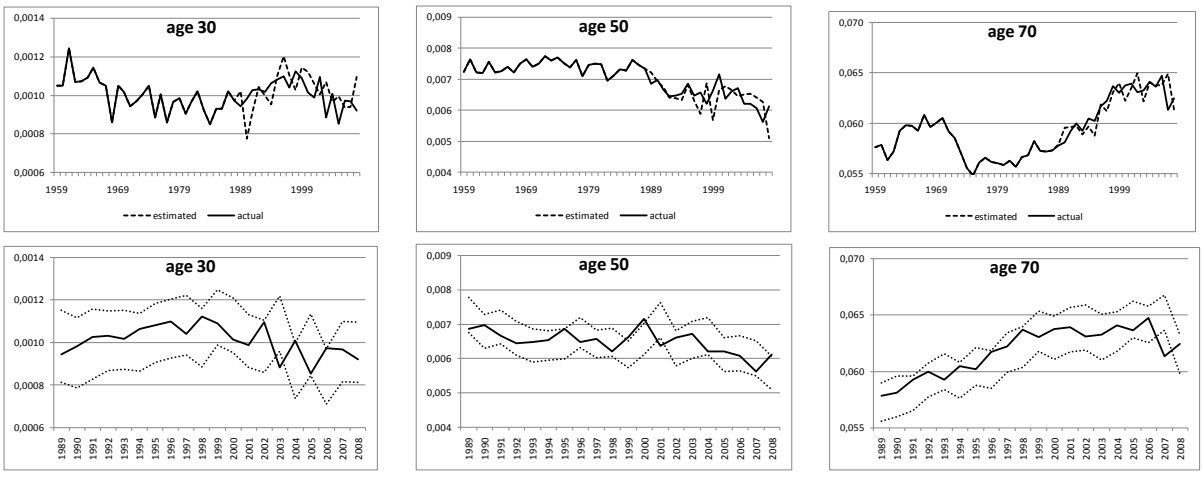

Figure 2: Actual mortality $m(x,t)$ versus estimated $\hat{m}(x,t)$. In the top row one can see sample paths for selected ages; in the bottom row there are 95%-prediction intervals based on 1.000 bootstrap simulations.

The graphs exhibit a certain robustness concerning uncertainty over time. The prediction interval width is stable for all ages across time. Moreover, these intervals are wider for older ages. The model, therefore, fulfils the demand for reasonableness of uncertainty as formulated by Cairns et al. (2011). Reasonable uncertainty of estimates and robustness are particularly relevant for solvency capital requirements.

2.3 Fixed horizon analysis

In addition to the above rolling forward lookback window approach we now consider a fixed lookback window. Thereby, we aim at predicting the next twenty years. To be precise, the lookback window is defined by the time span 1959−1988 and the lookforward window is the period from 1989−2008.

Dowd et al. (2010b) analysed the evolution of forecasts in expanding horizons, beginning with a lookforward window of one year and increasing to a span of twenty eight years. The outcome was graphically evaluated for consistency. In contrast, we are most interested in predictions for 2008, the final year of the lookforward window. In line with section 3.2 above, we evaluate sample paths and prediction intervals.

In a bootstrap approach we now simulate non-weighted residual mortality development factors $RF(x,t)$ rather than weighted development factors $\overline{RF}(x,t)$ since historic values of the latter refer to different number of lives. By so doing, we avoid estimating number of lives $\hat{l}(x,t)$ and death counts $\hat{d}(x,t)$. In this modelling approach estimated mortality probabilities are given by

$$\hat{m}(x,t) = \hat{m}(x,t-1)\cdot\left(1+RF(x,t)\right) .$$

Figure 3 illustrates sample paths as well as prediction intervals.

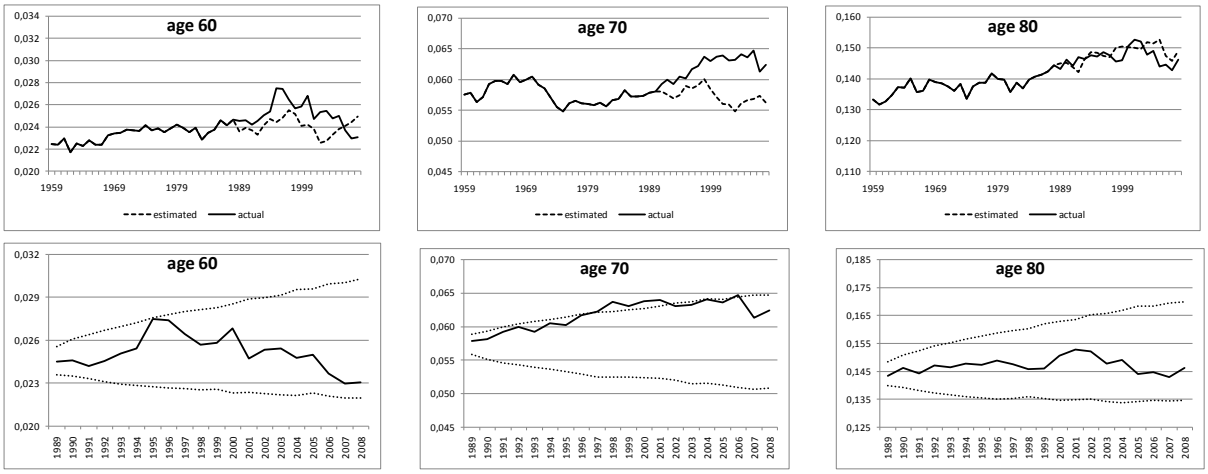

Figure 3: Actual mortality $q(x,t)$ versus estimated $\hat{q}(x,t)$ across time for selected ages. Sample paths are presented in the top row; in the bottom row there are 95%-prediction intervals based on 1.000 bootstrap simulations.

We note that predictions intervals get wider over time. This behaviour is desired: when estimated rates are further apart from the historic data on which basis they were produced, uncertainty about the estimation should necessarily increase. As age increases prediction intervals get wider. This fact may be somewhat masked in Figure 3 by the scales used for each age.

As for age 70, we note that actual mortality is partly above expectations. This type of underestimation is generally considered prudent by annuity insurers and pension providers.

Further, Dowd et al. (2010b) analysed mortality probability density forecasts. Thereby, actual probability of death was set in context to estimated probability density. Likewise we take the bootstrap model to produce cumulative density functions (CDF) as well as probability density functions (PDF) for residual mortality development factors $\hat{RF}(x,t)$ and, consequently, estimated mortality rates $\hat{m}(x,t)$ with respect to selected ages in calendar 2008.

In particular, we are interested in the percentage of mortality estimations that are smaller than actual mortality. As outlined above, overestimation of mortality can be especially dangerous to pension providers and annuity insurers since insured benefits would be undervalued. Said percentage can be interpreted as a *p*-value of a one side statistical test (Dowd 2010b). In Figure 4 we see cumulative density functions for selected ages. The dot on

each CDF curve indicates the percentage of all estimates that are smaller or equal to the actual mortality at that age.

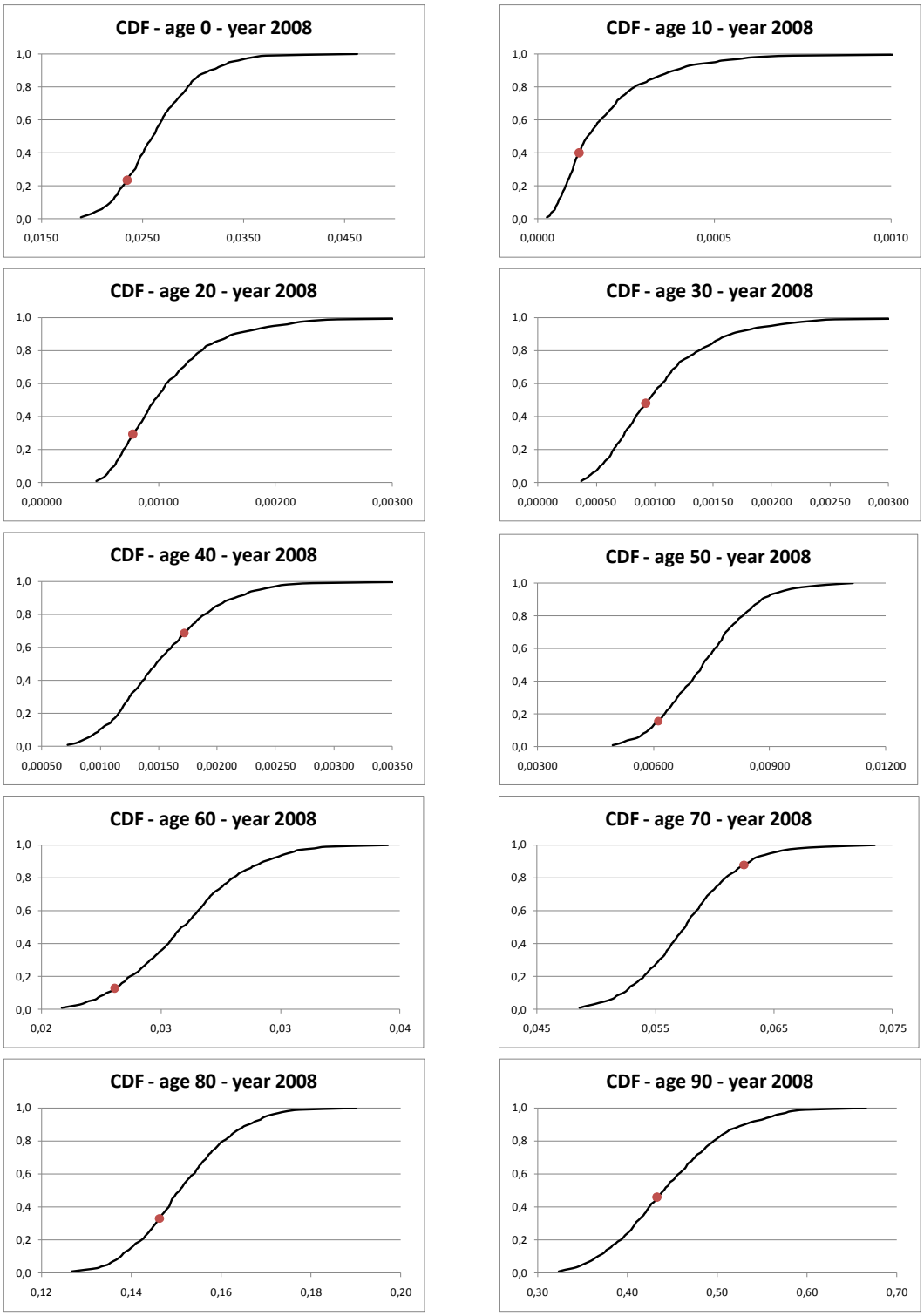

Figure 4: Cumulative density functions with respect to estimated mortality rates $\hat{m}(x,t)$. The dot marks the point on the curve whose *x*-coordinate is equal to the actual mortality in 2008 and whose *y*-coordinate marks the share of simulations that produced an estimation smaller or equal to actual mortality.

Overall, the percentage of estimates that are smaller than actual mortality is 0.45. That means, in about half the cases the share of simulated mortality estimates, that are too low, is about equal to the share of estimations that are too high. Obviously, results differ by age. Figure 5 illustrates our findings across age in calendar year 2008.

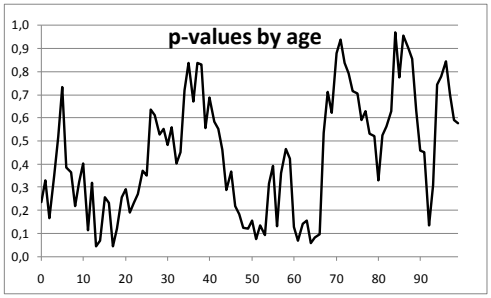

Figure 5: percentages of estimations that produced estimated mortality smaller or equal to actual mortality − by age in calendar year 2008.

Notably, we find that actual mortality for four ages only falls outside the 95% prediction interval concerning around our estimations. Overestimations of mortality are more hurtful for pension providers and annuity insurers as explained above. Dowd et al. (2010b) explored the ex-post performance of several stochastic mortality models. As a result they found that forecasts for age 65 showed a bias towards overestimation of mortality 28 years ahead across all models.

Our forecasts for residual mortality development factors also exhibit a similar bias for ages 60−66. This overestimation, however, is counterbalanced by underestimations relating to ages 68−75 and 84−88. We conclude that our stochastic mortality model provides reasonable and sound ex-post forecasts for residual mortality development factors. The undesirable one-sided bias towards overestimation of mortality can be avoided by employing our forecast model.

2.4 Expanding horizon analysis

Dowd et al. (2010b) further analysed ex-post forecasting performance of several stochastic models for expanding horizons. They fitted what they called models M1, M2B, M3B, M5, M6, M7 to LifeMetrics mortality data (Coughlan et al. 2007). The model specifications were

M1: $$\log m(x,t) = \beta_x^{(1)} + \beta_x^{(2)}\kappa_t^{(2)}$$

M2B: $$\log m(x,t) = \beta_x^{(1)} + \beta_x^{(2)}\kappa_t^{(2)} + \beta_x^{(3)}\gamma_c^{(3)}$$

M3B: $$\log m(x,t) = \beta_x^{(1)} + \frac{1}{n_\alpha}\kappa_t^{(2)} + \frac{1}{n_\alpha}\gamma_c^{(3)}$$

M5: $$\text{logit } q(x,t) = \kappa_t^{(1)} + \kappa_t^{(2)}(x-\bar{x})$$

M6: $$\text{logit } q(x,t) = \kappa_t^{(1)} + \kappa_t^{(2)}(x-\bar{x}) + \gamma_c^{(3)}$$

M7: $$\text{logit } q(x,t) = \kappa_t^{(1)} + \kappa_t^{(2)}(x-\bar{x}) + \kappa_t^{(3)}\left((x-\bar{x})^2 - \sigma_x^2\right) + \gamma_c^{(4)}$$

Next, they forecasted mortality rates from 1980 to 2008 for the age range 65 to 84. We computed implied mortality development factors that we weighted and smoothed further. Subsequently, the outcome is compared to realised mortality development factors in Figure 6a and 6b.

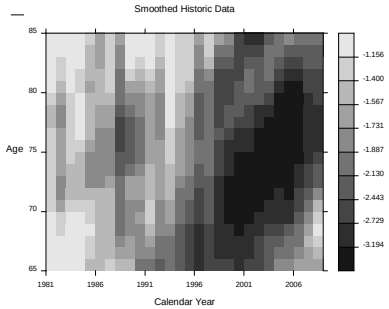

Figure 6a: smoothed historic mortality development factors for ages 65-84 from 1981-2008 versus

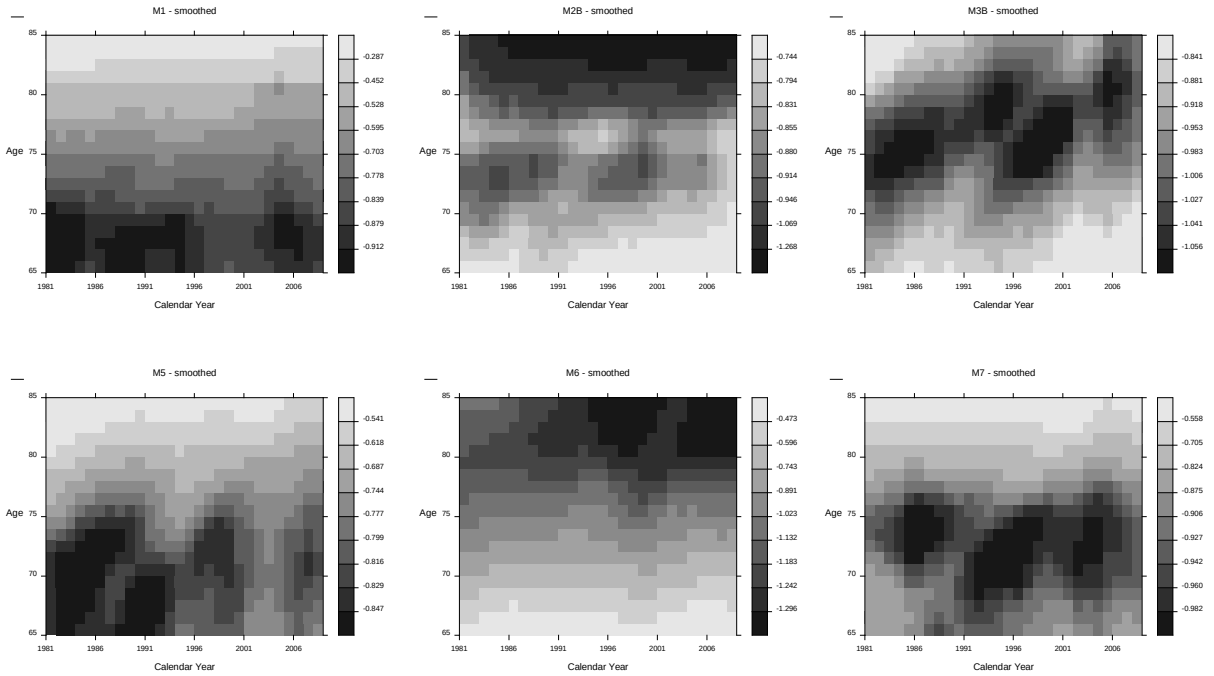

Figure 6b: smoothed mean forecasted mortality development factors for different stochastic models.

Mean forecasted incremental mortality development factors look quite different from subsequently realised ones for all six models. Considering the change in age dependent mortality over time as described above those stochastic models analysed by Dowd et al. (2010b) are incompatible with realised outcomes.

Interestingly, some of the above models incorporate the existence of cohort effects. Yet, none of these models predicted the pronounced cohort effect relating to birth years 1928-35. Ortmann (2013) stated that this effect began to emerge in 1975. Hence, it could have been picked up and continued into the future by analysing historic data up to 1979. Furthermore, none of these standard forecast models allows for the occurrence of new and additional mortality patterns.

It is natural to demand that changes in mortality should be reasonably smooth over time. However, all of these models analysed by Dowd et al. (2010b) exhibit a step change in the development of incremental mortality development factors over time at the beginning of the projection period. This insight might constitute an argument to consider these models less plausible than others.

3 Ex-ante analysis

Ortmann's (2013) idea was to separate the correlation structure as manifested by age, period and cohort effects from the data. First of all, Ortmann (2013) partitioned the data $IF(x,t)$ into mortality effects $ME(x,t)$ and residual development factors $RF(x,t)$. On the basis of said decomposition future residual mortaility development factors were forecasted future $\hat{RF}(x,t)$ by a non-parametric bootstrap model. Furthermore, Ortmann (2013) separately predicted future mortality effects $\hat{ME}(x,t)$, which were superimposed on residual mortality factors. Finally, he added up the two matrices cell by cell to derive forecasted incremental mortality development factors $\hat{IF}(x,t) = \hat{RF}(x,t) + \hat{ME}(x,t)$.

In essence, Ortmann (2013) proposed an additive APC-model. Notably, an age effects refers to a specific age over a certain period of time. A period effect relates to a calendar year and applies to a certain age band. A cohort effect refers to a year of birth and a calendar period. Age, period and cohort effects are detected on a case by case basis by means of the algorithm proposed by Ortmann (2013).

During the time period from 1979−2008 we observed 17 period effects. Hence the odds ratio is 17:13 that a period effect would happen again in any given year in the future. We further assume that if a period effect occurs in the future it will be identical to an observed historic effect. As far as future period effects are concerned we apply a non parametric bootstrap simulation based on these specifications. Figure 7 illustrates a sample pattern of future mortality effects.

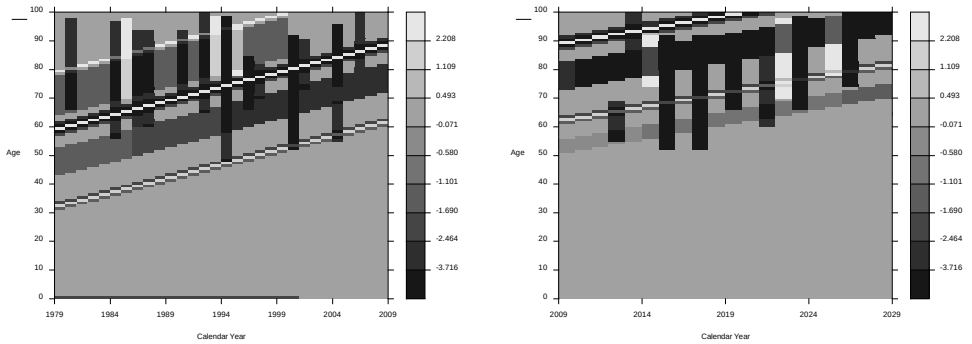

Figure 7: Historic mortality patterns $ME(x,t)$ for ages 0−99 during 1979−2008 on the left hand side and a random sample of simulated mortality effects $\hat{M}E(x,t)$ for 2009−2028 on the right.

In addition, Figure 8 shows a random sample of residual mortality development factors $\hat{R}F(x,t)$ as well as superimposed factors $\hat{I}F(x,t)$ that can be used to estimate mortality rates $\hat{m}(x,t)$.

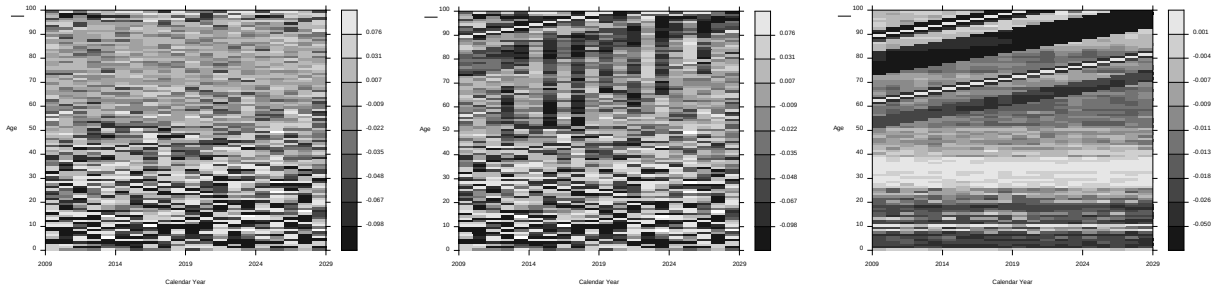

Figure 8: A random sample of simulated residual mortality development factors $\hat{R}F(x,t)$ for ages 0−99 during 2009−2028 on the left and a random sample of superimposed factors $\hat{I}F(x,t) = \hat{R}F(x,t) + \hat{M}E(x,t)$ in the middle. Mean forecasted incremental mortality development factors are displayed on the right.

The cohort effects detected earlier are continued into the future. As these effects are part of any sample they remain clearly visible when computing the mean as can be seen in Figure 8.

Haberman and Renshaw (2011) analysed several stochastic mortality models. The following predictor structures were compared:

LC: $\eta_{xt} = \alpha_x + \beta_x \kappa_t$

H1: $\eta_{xt} = \alpha_x + \beta_x \kappa_t + \iota_{t-x}$

LC2: $\eta_{xt} = \alpha_x + \beta_x^{(1)} \kappa_t^{(1)} + + \beta_x^{(2)} \kappa_t^{(2)}$

M5: $\eta_{xt} = \kappa_t^{(1)} + (x - \bar{x}) \kappa_t^{(2)}$

M6: $\eta_{xt} = \kappa_t^{(1)} + (x - \bar{x}) \kappa_t^{(2)} + \iota_{t-x}$

M7: $\eta_{xt} = \kappa_t^{(1)} + (x - \bar{x}) \kappa_t^{(2)} + b(x) \kappa_t^{(3)} + \iota_{t-x}$

In particular, Haberman and Renshaw forecasted mortality rates from 2009 to 2028 for the age range 65 to 84. Based on these projections we computed mortality development factors that we weighted and smoothed further. For plausibility purposes, we compared forecasted means in Figure 9.

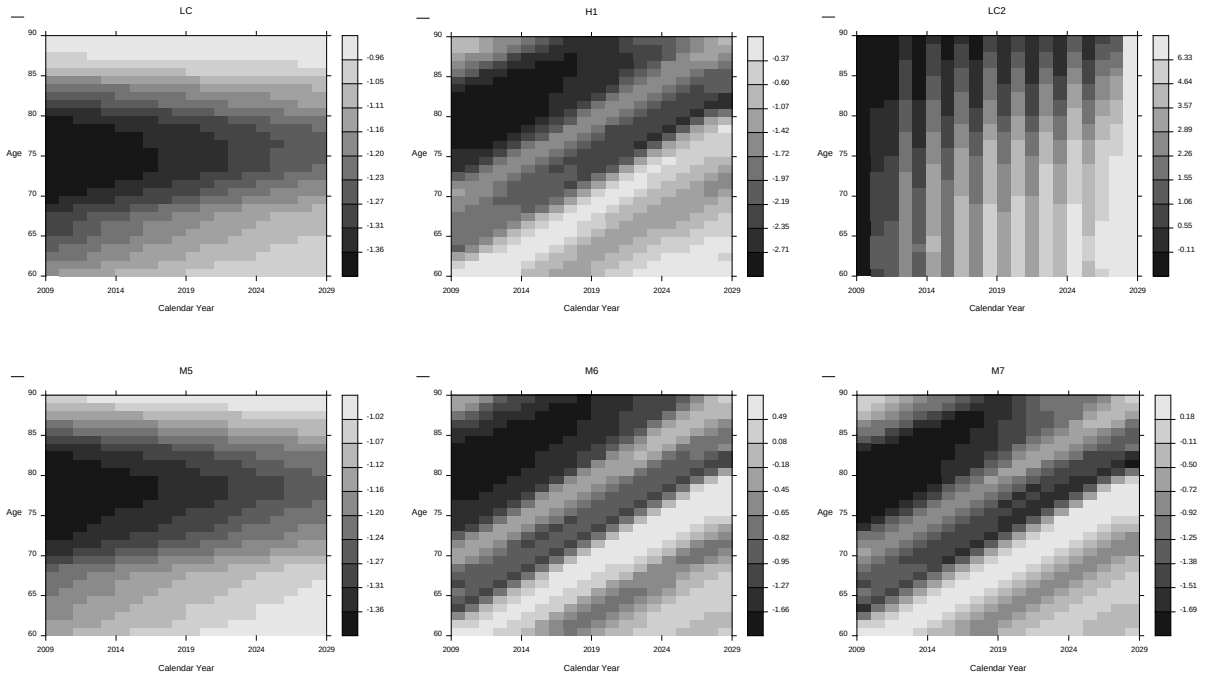

Figure 9: smoothed mean forecasted mortality development factors for different stochastic models.

Models H1, M6, and M7 exhibit a pronounced cohort effect that is continued into the future. For this reason, they seem more plausible than models LC, LC2 and M5.

3.1 Mortality forecasts

In line with section 2.3 above we now consider a lookback window spanning 30 years of historic data from 1979−2008. The lookforward window is the period from 2009−2028. We then simulate residual mortality development factors $\hat{RF}(x,t)$ as well as mortality effects $\hat{ME}(x,t)$ as described above. Based on forecasted incremental mortality development factors $\hat{IF}(x,t) = \hat{RF}(x,t) + \hat{ME}(x,t)$ we are able to iteratively compute estimated mortality

$$\hat{m}(x,t) = \hat{m}(x,t-1) \cdot \left(1 + \hat{IF}(x,t)\right), \qquad t = 2009,\ldots,2028,$$

by taking into account that $\hat{m}(x,2008) = m(x,2008)$. We then derive mortality forecasts based on a bootstrap simulation of 1.000 runs. Results are in line with those presented by Ortmann (2013).

It was not apparent in the results presented in Ortmann (2013) that 95% prediction intervals at age 80 are wider than as at age 60. More generally, prediction intervals get wider by age as can be seen in Figure 10. This feature was deemed a critical criterion by Cairns et al. (2011) as they discredited two stochastic mortality models, including the Lee-Carter model (Lee and Carter 1992), on grounds of plausibility for producing less uncertain estimates at age 85 than at age 65.

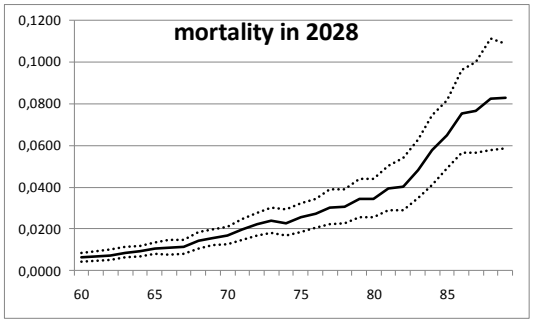

Figure 10: Forecasted mean death rate for ages 60−89 in 2028 with 95% prediction intervals based on 1.000 simulations.

3.2 Life expectancy forecasts

In this section we switch our attention to life expectancy forecasts at birth and as at normal retirement age 65. This approach is line with Cairns et al. (2011) who considered the survivor index as well as annuity prices apart from mortality rates. Haberman and Renshaw (2011) forecasted life expectancies and annuity values in particular.

We define life expectancy at birth as usual. Based on a bootstrap simulation of 1.000 runs we have computed mean life expectancies and 95%-prediction intervals over the period from 2009−2028, see Figure 11.

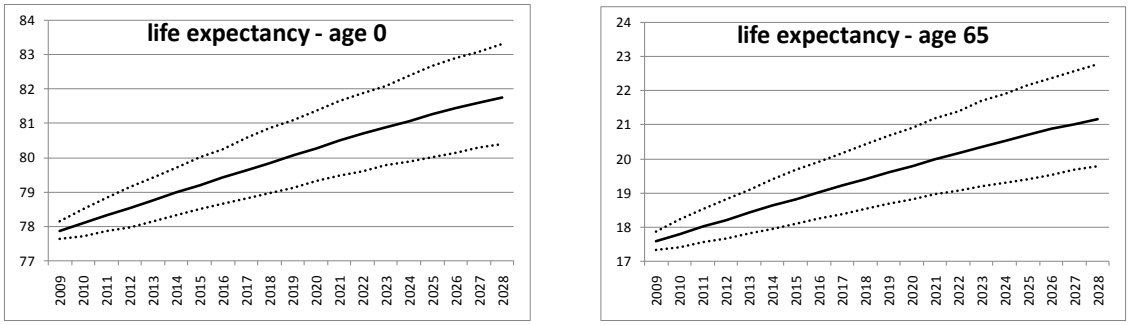

Figure 11: mean life expectancy and 95% prediction intervals at age 0 and at age 65 from 2009−2028.

It can be seen that the best estimate increase in life expectancy at birth of about 3.9 years over the next twenty years is almost fully generated by an increase of remaining life expectancy at normal retirement age which is predicted to rise by 3.6 years. Furthermore uncertainty increases as time goes by as the prediction intervals get wider over time.

Furthermore, we have derived the cumulative density function for life expectancy at birth and at age 65 in 2028. Figure 12 illustrates our findings.

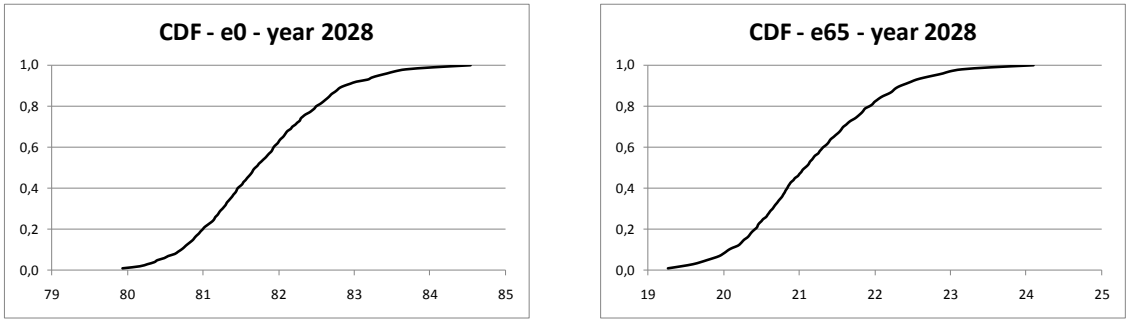

Figure 12: CDF of forecasted life expectancy at age 0 and at age 65 in 2028 based on 1.000 simulations.

Figure 12 offers an alternative visualisation to fan charts that were used to represent the uncertainty in projections of future life expectancy by Dowd et al. (2010c) and Cairns et al. (2011).

4 Summary

Existing mortality forecast models focus on probabilities of death, mortality rates, or forces of mortality respectively. We argue that one should additionally consider and analyse changes in mortality over time. This approach is in line with demanding differentiability on top of continuity when extrapolating a real valued function $f:[a,b]\to\mathbb{R}$. As a consequence, we have turned our investigations to incremental mortality development factors and discussed this plausibility criterion.

The key idea for forecasting mortality is to separate the correlation structure, as manifested by age, period and cohort effects, from the rest of the data. Notably, Ortmann (2013) presented a framework for reliably and objectively detecting mortality effects by drawing on established techniques in computer vision and image processing. Furthermore, he forecasted mortality effects as well as residual development factors separately into the future based on non-parametric block bootstrapping.

We evaluated the model's performance ex-post as well as ex-ante by adapting and enhancing a number of backtesting methods that were recently developed. In contrast to some other stochastic forecast models that tend to overestimate mortality, the new methodology does not exhibit this undesirable bias.

5 References

[1] Booth, H., Hyndman, R. J., Tickle, L., and de Jong, P. (2006): Lee-Carter mortality forecasting: a multi-country comparison of variants and extensions, Demographic Research, 15: 289–310.

[2] Cairns, A. J. G., Blake, D., and Dowd, K. (2008): Modelling and management of mortality risk: A review, Scandinavian Actuarial Journal, 2–3, 79–113.

[3] Cairns, A. J. G., Blake, D., Dowd, K., Coughlan, G. D., Epstein, D., Ong, A. and 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, 1–35.

[4] Cairns, A. J. G., Blake, D., Dowd, K., Coughlan, G. D., Epstein, D., Khalaf-Allah, M. (2011): Mortality density forecasts: An analysis of six stochastic mortality models, Insurance: Mathematics and Economics, 48, 355–367.

[5] Canny, J. F.(1986): A computational approach to edge detection, IEEE Transaction on Pattern Analysis and Machine Intelligence 8, 679–698.

[6] Coughlan, G., Epstein, D., Sinha, A., Honig, P. (2007): q-Forwards: Derivatives for Transferring Longevity and Mortality Risks. Available at www.jpmorgan.com/lifemetrics.

[7] Cutler, D., Deaton, A., Lleras-Muney, A. (2006): The Determinants of Mortality, Journal of Economic Perspectives, 20, 97–120.

[8] Dowd, K., Cairns, A. J. G., Blake, D., Coughlan, G. D., Epstein, D., Khalaf-Allah, M. (2010a): Evaluating the goodness of fit of stochastic mortality models, Insurance: Mathematics and Economics, 47, 255–265.

[9] Dowd, K., Cairns, A. J. G., Blake, D., Coughlan, G. D., Epstein, D., Khalaf-Allah, M. (2010b): Backtesting stochastic mortality models: An ex-post evaluation of multi-period-ahead density forecasts, North American Actuarial Journal, 14, 281–298.

[10] Dowd, K., Blake, D., Cairns, A. J. G. (2010c): Facing up to uncertain life expectancy: The longevity fan charts, 47, 67–78.

[11] Haberman, S., Renshaw, A. E. (2011): A comparative study of parametric mortality projection models. Insurance: Mathematics and Economics, 48, 35–45.

[12] Human Mortality Database (2010): University of California , Berkeley (USA), and Max Planck Institute for Demographic Research (Germany), available at http://www.mortality.org.

[13] Keilmann, N. (1998): How accurate are the United States world population projections? Population and Development Review, 24, 15 − 41.

[14] Koissi, M.-C., Shapiro, A. F., Högnäs, G. (2006): Evaluating and Extending the Lee-Carter Model for Mortality Forecasting: Bootstrap Confidence Interval, Insurance: Mathematics and Economics 38: 1–20.

[15] Lee, R. D., Carter, L. R. (1992): Modeling and Forecasting U. S. Mortality, Journal of the American Statistical Association, 87, 659–671.

[16] Lee, R. D., Miller, T. (2001): Evaluating the Performance of the Lee-Carter Method for Forecasting Mortality, Demography, 38, 537–549.

[17] Ortmann, K. M. (2013): Computerised detection and stochastic forecast of age, period and cohort effects, Zeitschrift für die Gesamte Versicherungswissenschaft, Band 102, Heft 5, 577-595

[18] Plat, R. (2010): One-year value-at-risk for mortality and longevity, The Pensions Institute, Discussion Paper PI-1015, ISSN 1367-580X