



Pensions
Institute

DISCUSSION PAPER PI-0908

On Stochastic Mortality Modeling

Richard Plat

July 2009

ISSN 1367-580X

The Pensions Institute
Cass Business School
City University
106 Bunhill Row London
EC1Y 8TZ
UNITED KINGDOM

<http://www.pensions-institute.org/>

ON STOCHASTIC MORTALITY MODELING

RICHARD PLAT^a

University of Amsterdam, Eureko / Achmea Holding and Netspar

This version: July 1, 2009¹

Abstract

The last decennium a vast literature on stochastic mortality models has been developed. All well known models have nice features but also disadvantages. In this paper a stochastic mortality model is proposed that aims at combining the nice features from existing models, while eliminating the disadvantages. More specifically, the model fits historical data very well, is applicable to a full age range, captures the cohort effect, has a non-trivial (but not too complex) correlation structure and has no robustness problems, while the structure of the model remains relatively simple. Also, the paper describes how to incorporate parameter uncertainty in the model. Furthermore, a risk neutral version of the model is given, that can be used for pricing.

JEL classification: G22; G23; J11

Subject classification: IM10; IE43; IB10

Keywords: stochastic mortality models, longevity risk, pricing, Solvency 2, Monte Carlo simulation

1. Introduction

In recent years there has been an increasing amount of attention of the insurance industry for the quantification of the risks that insurers are exposed to. Important drivers of this development are the increasing internal focus on risk measurement and risk management and the introduction of Solvency 2 (expected to be implemented around 2012).

Solvency 2 will lead to a change in the regulatory required solvency capital for insurers. At this moment this capital requirement is a fixed percentage of the mathematical reserve or the risk capital. Under Solvency 2 the so-called Solvency Capital Requirement (SCR) will be risk-based, and market values of assets and liabilities will be the basis for these calculations.

¹ The author likes to thank Antoon Pelsser, Peter Boswijk, Marwa Khalaf-Allah, David Epstein, Andrew Cairns, Steven Haberman and an anonymous referee.

^a University of Amsterdam, Dept. of Quantitative Economics, Roetersstraat 11, 1018 WB Amsterdam, The Netherlands, e-mail: H.J.Plat@uva.nl, tel (0031)20-525 4364

Also for pension funds, a new solvency framework will be developed, either as part of Solvency 2 or as a separate project (usually named IORP 2). It is expected that the general principles will be similar as Solvency 2, meaning market valuation of assets and liabilities and risk-based solvency requirements.

Important risks to be quantified are mortality and longevity risk. Not only is this an important risk for most (life) insurers and pension funds, the resulting solvency margin will also be part of the fair value reserve. Reason for this is that it is becoming best practice for the quantification of the Market Value Margin to apply a Cost of Capital rate to the solvency capital necessary to cover for unhedgeable risks, such as mortality and longevity risks.

There is a vast literature on stochastic modeling of mortality rates. Often used models are for example those of Lee and Carter (1992), Renshaw and Haberman (2006), Cairns et al (2006a), Currie et al (2004) and Currie (2006). For an extensive review we refer to section 2.

All well known models have nice features but also disadvantages. In this paper a mortality model is proposed that aims at combining the nice features from existing models, while eliminating the disadvantages of existing models. More specifically, the model fits historical data very well, is applicable to a full age range, captures the cohort effect, has a non-trivial (but not too complex) correlation structure, has no robustness problems and can take into account parameter risk, while the structure of the model remains relatively simple.

The remainder of the paper is organized as follows. First, in section 2 the existing literature review is extensively reviewed, focusing on stochastic mortality models and the criteria for them. In section 3 a new mortality model is proposed. Section 4 describes the fitting procedure of the model and gives results of the fitting process for mortality of different countries. Section 5 shows simulation results of mortality rates and the results of a robustness test. In section 6 a risk neutral version of the model is given, which can be used for pricing. Section 7 describes a possible method to account for parameter risk for the proposed mortality model. Conclusions are given in section 8.

2. Literature review: criteria and models

Due to the increasing focus on risk management and measurement for insurers and pension funds, the literature on stochastic mortality models has developed rapidly during the last decennium. In this section an overview of current literature on stochastic mortality models and criteria for them is given.

2.1 Criteria for stochastic mortality models

It is important to consider whether a specific stochastic mortality model is a good model or not. Therefore, Cairns et al (2008a) defined criteria against which a model can be assessed:

- 1) Mortality rates should be positive.

- 2) The model should be consistent with historical data.
- 3) Long-term dynamics under the model should be biologically reasonable.
- 4) Parameter estimates and model forecasts should be robust relative to the period of data and range of ages employed.
- 5) Forecast levels of uncertainty and central trajectories should be plausible and consistent with historical trends and variability in mortality data.
- 6) The model should be straightforward to implement using analytical methods or fast numerical algorithms.
- 7) The model should be relatively parsimonious.
- 8) It should be possible to use the model to generate sample paths and calculate prediction intervals.
- 9) The structure of the model should make it possible to incorporate parameter uncertainty in simulations.
- 10) At least for some countries, the model should incorporate a stochastic cohort effect.
- 11) The model should have a non-trivial correlation structure.

An important additional criterion is that the model is applicable for a full age range. Some models are designed for higher ages only (say 60 years or older). However, the portfolios of insurers and pension funds usually exist of policyholders from age 20 and older. One would want to model the mortality rates and their dependencies for the whole portfolio consistently, therefore the model should be applicable for the whole age ranges.

The existing models meet most of the above criteria. However, as far as we know, none of the existing models meet all of the above criteria (although some are close), see section 8 and Cairns et al (2007).

2.2 Stochastic mortality models

Stochastic mortality models either model the central mortality rate or the initial mortality rate (see Coughlan et al (2007)). The central mortality rate $m_{x,t}$ is defined as:

$$(2.1) \quad m_{x,t} = \frac{D_{x,t}}{E_{x,t}} = \frac{\text{\# deaths during calendar year } t \text{ aged } x \text{ last birthday}}{\text{average population during calendar year } t \text{ aged } x \text{ last birthday}}$$

The initial mortality rate q_x is the probability that a person aged x dies within the next year. The different mortality measures are linked by the following approximation:

$$(2.2) \quad q_x \approx 1 - e^{-m_x}$$

One of the most well known stochastic mortality models is the model of Lee and Carter (1992):

$$(2.3) \quad \ln(m_{x,t}) = a_x + b_x \kappa_t$$

where a_x and b_x are age effects and κ_t is a random period effect. Cairns et al (2007, 2008a and 2008b) noted several disadvantages of the Lee-Carter model:

- It is a 1-factor model, resulting in mortality improvements at all ages being perfectly correlated (trivial correlation structure).
- For countries where a cohort effect is observed in the past, the model gives a poor fit to historical data.
- The uncertainty in future death rates is proportional to the average improvement rate b_x . For high ages this can lead to this uncertainty being too low, since historical improvement rates have often been lower at high ages.
- The basic version of the model can result in a lack of smoothness in the estimated age effect b_x .

There is whole strand of literature on additions or modifications of the Lee-Carter model, for example Brouhns et al (2002), Lee and Miller (2001), Booth et al (2002), Girosi and King (2005), De Jong and Tickle (2006), Delwarde et al (2007) and Renshaw and Haberman (2003). Most of these models tackle one of the problems of the Lee-Carter model, but the other disadvantages still remain.

The first model that incorporated the cohort effect was proposed in Renshaw and Haberman (2006):

$$(2.4) \quad \ln(m_{x,t}) = a_x + b_x^1 \kappa_t + b_x^2 \gamma_{t-x}$$

where γ_{t-x} is a random cohort effect that is a function of the year of birth ($t-x$).

For countries where a cohort effect is observed in the past, this model provides a significant better fit to the historical data. However, CMI (2007) and Cairns et al (2007, 2008b) find that the Renshaw-Haberman suffers from a lack of robustness. Furthermore, although the model has an additional stochastic factor for the cohort effect, for most of the simulated mortality rates the correlation structure is still trivial. Especially when using a wide age range, the simulated cohort parameters are only relevant for the higher ages in the far end of the projection.

Currie (2006) introduced a simplification of the Renshaw-Haberman model that removes the robustness problem:

$$(2.5) \quad \ln(m_{x,t}) = a_x + \kappa_t + \gamma_{t-x}$$

However, the fit quality is less compared with the Renshaw-Haberman model, and the problem with the trivial correlation structure still remains.

When fitting models (2.4) and (2.5) to an age range of say 20-85, the modeled cohort effect can result in odd looking humps in the projected mortality rates over time. This problem will be further highlighted in the next paragraph.

Furthermore, Cairns et al (2008b) observe that for England & Wales and United States data, the fitted cohort effect appears to have a trend in the year of birth. This suggests that the cohort effect compensates the lack of a second age-period effect, as well as trying to capture the cohort effect in the data.

Cairns et al (2006a) introduced the following model:

$$(2.6) \quad \text{logit}(q_{x,t}) = \ln \left(\frac{q_{x,t}}{1-q_{x,t}} \right) = \kappa_t^1 + \kappa_t^2 (x - \bar{x})$$

where \bar{x} is the mean age in the sample range and (κ_t^1, κ_t^2) is assumed to be a bivariate random walk with drift. Cairns et al (2007) also introduced some additions on model (2.6), amongst others capturing the cohort effect. The models have multiple factors that result in a (desired) non-trivial correlation structure, while the structure of the model is relatively simple. However, those models are all designed for higher ages only. When using these models for full age ranges, the fit quality will be relatively poor and the projections are likely to be biologically unreasonable.

2.3 Problems with modeling cohort effect

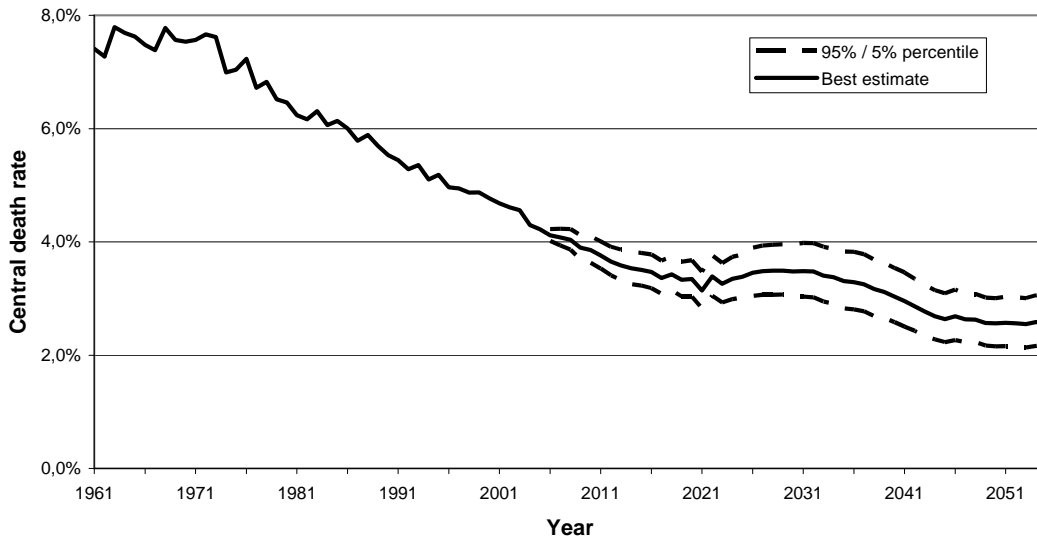
Various explanations have been put forward for cohort effects that have been identified in the past. For example, for the United Kingdom Willets (2004) mentions historical patterns of smoking behavior and the impact of early life experience on health in later life. He states that there are a number of reasons to believe that this cohort effect will have an enduring impact on rates of mortality improvement in future decades.

The investigations on cohort effects often concentrate on birth years until about 1945. This is natural, since in most cases the cohort effect is an effect on health in later life, so one needs observations of mortality rates for middle age and higher ages to verify the existence of the cohort effect. When applying models (2.4) and (2.5) to a full age range, say 20-85, cohort parameters are also fitted for birth years 1945-1980. This means that for these birth years, (cohort) movements for young ages (which can be volatile) are projected into the future. This affects the mortality rates for higher ages in a similar degree, since the cohort effect is usually modeled in a multiplicative way. However, given the possible nature of the movements for these specific birth years (for example AIDS, drug and alcohol abuse and violence) it is unclear whether these effects do have a persistent effect on the future mortality rates for these cohorts. And if so, it is questionable whether a high relative cohort effect for young ages will have a similar relative effect on mortality of higher ages, given the nature of the cohort effect for young ages².

Figure 1 shows a best estimate and percentiles of mortality rates for 75 years old males, using the Currie (2006) model applied to United States mortality for an age range of 20-84.

² Note that the Renshaw-Haberman tries to capture this in the parameter b_x^2 . However, this is based on the cohort effects for earlier birth years, which could have a significantly different nature.

Figure 1: projected mortality rates, 75 years old male – Currie (2006) model



The figure shows an odd-looking hump around 2020-2040 and flattening of projected mortality thereafter, corresponding with patterns in the fitted cohort parameters for birth years 1945-1980.

Given the considerations above and the odd-looking results when taking into account cohort effects of recent birth years, it might be wise to only include the cohort effect for early birth years (say until year 1945) in the fitting of the model. The cohort effect for later birth years (so > 1945) can be simulated from the fitted distributions. An additional advantage of this is that the simulation of the cohort effect becomes relevant for higher ages already in the beginning of the projection, leading to a non-trivial correlation structure.

3. A new stochastic mortality model

The models mentioned in the previous section all have some nice features:

- the a_x term of the Lee-Carter model makes it suitable for full age ranges
- the Renshaw-Haberman model addresses the cohort effect and fits well to historical data
- the Currie model has a simpler structure than the Renshaw-Haberman model, making it more robust
- the models of Cairns et al (2006a, 2007) have multiple factors, resulting in a non-trivial correlation structure, while the structure of the model is relatively simple

In this section a mortality model is proposed that combines those nice features, while eliminating the disadvantages mentioned in the previous section.

3.1 The proposed model

As for most other stochastic mortality models, the quantity of interest is the central mortality rate $m_{x,t}$. The proposed model for $m_{x,t}$ is:

$$(3.1) \quad \ln(m_{x,t}) = a_x + \kappa_t^1 + \kappa_t^2 (\bar{x} - x) + \kappa_t^3 (\bar{x} - x)^+ + \gamma_{t-x}$$

where $(\bar{x} - x)^+ = \max(\bar{x} - x, 0)$. The model has 4 stochastic factors, but has a similar relatively simple structure as the Currie (2006) and the Cairns et al (2006a, 2007) models.

The a_x is similar as in the Lee-Carter model and makes sure that the basic shape of the mortality curve over ages is in line with historical observations. Next to the a_x , the model has 4 stochastic factors $(\kappa_t^1, \kappa_t^2, \kappa_t^3, \gamma_{t-x})$. The parameters of the model can be fitted using the methodology described in section 4, after which suitable ARIMA-processes are selected for the various factors.

The factor κ^1 represents changes in the level of mortality for all ages. Following the reasoning in Cairns et al (2006b), the (long-term) stochastic process for this factor should not be mean reverting. Reason for this is that it is not expected that higher mortality improvements in some years will surely be compensated by lower mortality improvements in later years.

The factor κ^2 allows changes in mortality to vary between ages, to reflect the historical observation that improvement rates can differ for different age classes.

Furthermore, historical data seems to indicate that the dynamics of mortality rates at lower ages (up to age 40 / 50) can be (significantly) different at some times. For example, think of developments like AIDS, drugs and alcohol abuse, and violence. The factor κ^3 is added to capture these dynamics.

The factor γ_{t-x} is capturing the cohort effect, in the same way as the models of Currie (2006) and Cairns et al (2007). As mentioned in paragraph 2.2, the process for this factor should not have a trend. Therefore, a trendless mean reverting process will be assumed for γ_{t-x} .

Next to γ_{t-x} , the factors κ^2 and κ^3 allow the model to have a non-trivial correlation structure between ages. Fitting non-stationary ARIMA-process for factors κ^2 and κ^3 could result (in some scenarios) in projected scenarios where the shape of the mortality curve over ages is not biologically reasonable. Therefore, a stationary (mean reverting) process will be assumed for these factors.

In most cases mortality projections for a wide age range are needed. However, if one is only interested in higher ages (say age 60 and older), the factor κ^3 is not needed and can be left out. This reduces the model to:

$$(3.2) \quad \ln(m_{x,t}) = a_x + \kappa_t^1 + \kappa_t^2 (\bar{x} - x) + \gamma_{t-x}$$

This reduced model still has all the favorable characteristics of model (3.1), but is more suitable for high ages only.

3.2 Identifiability constraints

Just like most stochastic mortality models, the proposed mortality model has an identifiability problem, meaning that different parameterizations could lead to identical values for $\ln(m_{x,t})$. Note that the following parameterization leads to similar values for $\ln(m_{x,t})$:

$$(3.3) \quad \begin{aligned} \tilde{\gamma}_{t-x} &= \gamma_{t-x} + \psi_1 + \psi_2(t-x) \\ \tilde{\kappa}_t^1 &= \kappa_t^1 + \psi_1 - d\bar{x}\psi_2 \\ \tilde{\kappa}_t^2 &= \kappa_t^2 + d\psi_2 \\ \tilde{a}_x &= a_x + (1-d)\psi_2 \end{aligned}$$

where ψ_1 , ψ_2 and d are constants.

This can be resolved by setting identifiability constraints. We use the approach of Cairns et al (2007, model M6) for this, leading to the following constraints:

$$(3.4) \quad \begin{aligned} \sum_{c=c_0}^{c_1} \gamma_c &= 0 \\ \sum_{c=c_0}^{c_1} c \gamma_c &= 0 \\ \sum_t \kappa_t^3 &= 0 \end{aligned}$$

where c_0 and c_1 are the earliest and latest year of birth to which a cohort effect is fitted, and $c = t-x$. The rationale behind the choice of first two constraints is that if the function $\psi_1 + \psi_2(t-x)$ is fitted to γ_{t-x} the constraints ensure that $\hat{\psi}_1 = \hat{\psi}_2 = 0$. This results in a fitted process for γ_{t-x} that will fluctuate around 0 and there will be no constant trend up or down. This means that the constraints in (3.4) force the process of γ_{t-x} only to be used to capture the cohort effect and not to compensate lack of age-period effects. The third constraint is only used to normalize the estimates for κ^3 .

Other approaches for setting the identifiability constraints are also possible, see for example Cairns et al (2007) and Renshaw and Haberman (2006).

4. Fitting the model

An important aspect of stochastic mortality models is the quality of the fit of the model to historical mortality data. In this section the methodology for fitting the model is described, and a

comparison of fit quality with other models is made for mortality rates of the United States (US), England & Wales (E&W) and The Netherlands (NL).

4.1 Fitting methodology

Brouhns et al (2002) described a fitting methodology for the Lee-Carter model based on a Poisson model. The main advantage of this is that it accounts for heteroskedasticity of the mortality data for different ages. This method has been used more commonly after that, also for other models, see for example Renshaw and Haberman (2003, 2006) and Cairns et al (2007).

This fitting methodology will be applied to the model proposed in section 3. Therefore, the number of deaths is modeled using the Poisson model, implying:

$$(4.1) \quad D_{x,t} \sim \text{Poisson}(E_{x,t} m_{x,t})$$

where $D_{x,t}$ is the number of deaths, $E_{x,t}$ is the exposure (see (2.1)) and $m_{x,t}$ is modeled as in (3.1). The parameter set ϕ is fitted with maximum likelihood estimation, where the log-likelihood function of model (4.1) is given by:

$$(4.2) \quad L(\phi; D, E) = \sum_{x,t} \left\{ D_{x,t} \ln [E_{x,t} m_{x,t}(\phi)] - E_{x,t} m_{x,t}(\phi) - \ln(D_{x,t}!) \right\}$$

Because of the specific nature of the problem, there are (as far as we know) no commercial statistical packages that implement this Poisson regression with constraints. Therefore we have used the R-code of the (free) software package “Lifemetrics” as a basis for fitting (4.2)³. Another reason for using this is to make an honest comparison of the fit quality of the model proposed in this paper and existing models (also modeled in Lifemetrics), which is the topic of the next paragraph.

Besides estimates for a_x , the fitting procedure described above leads to time series of estimations of κ^1 , κ^2 , κ^3 and γ_{t-x} . The next step in fitting the model is selecting and fitting a suitable ARIMA-process to these time series (see paragraph 4.3).

4.2 Comparison of fit quality with existing models

To evaluate whether the proposed model fits historical data well, we have fitted the model to three different data sets and compared the fitting results with those of models from the Lifemetrics toolkit. The three used data sets are:

- United States, Males, 1961-2005, ages 20-84
- England & Wales, Males, 1961-2005, ages 20-89
- The Netherlands, Males, 1951-2005, ages 20-90

The data consists of numbers of deaths $D_{x,t}$ and the corresponding exposures $E_{x,t}$ and is extracted from www.mortality.org⁴.

³ See www.lifemetrics.com and <http://www.r-project.org/>. Lifemetrics is an (open source) toolkit for measuring and managing longevity and mortality risk, designed by J.P. Morgan.

⁴ Note that a longer history is available. We used these historic periods (for the U.S. and E&W) to be able to roughly compare results with Cairns et al (2007) and Cairns et al (2008).

As in Cairns et al (2007), the models are compared using the Bayes Information Criterion (*BIC*). The measure *BIC* provides a trade-off between fit quality and parsimony of the model. The *BIC* is defined as:

$$(4.3) \quad BIC = L(\hat{\phi}) - \frac{1}{2} K \ln N$$

where $\hat{\phi}$ is the maximum likelihood estimate of the parameter vector, N is the number of observations and K is the number of parameters being estimated.

Table 1 gives a comparison of the fitting results (in terms of *BIC*) of the model proposed in section 3 and existing models (fitted with the Lifemetrics toolkit).

Table 1: comparison BIC for proposed model and existing models

<i>BIC mortality models</i> *	U.S.	E&W	NL
Plat	-24.506	-18.151	-18.425
Renshaw-Haberman (2006)	-25.971	-18.062	-18.632
Currie (2006)	-37.489	-19.805	-18.597
Lee-Carter (1992)	-47.542	-22.949	-20.353
Cairns et al (2007, model M7)	-56.571	-27.730	-21.055
Cairns et al (2006)	-294.928	-66.744	-31.511

* higher *BIC* is more favorable

The table shows that for these specific data sets the proposed model gives the best fitting results, closely followed by the Renshaw-Haberman model. The *BIC* for the other models is (sometimes significantly) less. The models of Cairns et al (2006a, 2007) do not perform very well for this age range, since they are designed for higher ages only.

In the fitting process of the models above the cohort effect was taken into account for all birth years of the dataset. However, given the reasons mentioned in paragraph 2.3, for the remaining of this paper we will exclude the cohort effect for birth years later than 1945 in the fitting of the model. In general this will reduce the quality of the fit somewhat, as is shown in table 2.

Table 2: results BIC when excluding cohorts > 1945

<i>BIC mortality models</i> *	U.S.	E&W	NL
Plat	-24.506	-18.151	-18.425
Plat (excluding cohorts > 1945)	-32.392	-18.927	-18.378

* higher *BIC* is more favorable

The fitting results of the model are still good when excluding the cohort effect for birth years > 1945, certainly considering the fact that the *BIC* of the other models which include a cohort

effect (Renshaw-Haberman (2006), Currie (2006) and Cairns et al (2007)) would also be less favorable when excluding these birth years⁵.

Note that the proposed model nests the model of Currie (2006). For nested models, the use of a likelihood ratio test is more appropriate than the use of the *BIC* measure. The likelihood ratio (*LR*) test can be used to test the null hypothesis that the nested model (in this case, the Currie (2006) model) is the correct model against the alternative that the more general model (the model proposed in this paper) is correct. The likelihood ratio test statistic is:

$$(4.4) \quad \xi_{LR} = 2 \left[L(\hat{\phi}) - L(\tilde{\phi}) \right]$$

where $L(\hat{\phi})$ is the log-likelihood of the general model and $L(\tilde{\phi})$ of the nested model.

Under the null hypothesis, ξ_{LR} has a Chi-squared distribution with J degrees of freedom, J being the additional parameters being estimated in the general model compared to the nested model. Therefore, the null hypothesis can be rejected if:

$$(4.5) \quad \xi_{LR} > \chi_{J,\alpha}^2$$

where α is the significance level. Alternatively, the p -value can be determined for this test:

$$(4.6) \quad p = 1 - \chi_J^{-2} \left(2 \left[L(\hat{\phi}) - L(\tilde{\phi}) \right] \right)$$

The p -value is the probability of obtaining the observed value, assuming that the null hypothesis is true. If the p -value is lower than α , the null hypothesis is rejected. Table 3 shows the results of the likelihood ratio test for the three data sets.

Table 3: LR test, null hypothesis Currie (2006) model against proposed model

	Likelihood Ratio test statistic	Degrees of freedom	p-value
U.S.	26.677	89	< 0,0001
E&W	4.023	89	< 0,0001
Holland	1.245	109	< 0,0001

The table shows that for each data set the null hypothesis is rejected overwhelmingly. Therefore, the conclusion above (based on *BIC* results) that the model proposed in this paper is preferable to the nested Currie (2006) model is supported by the results from the likelihood ratio test.

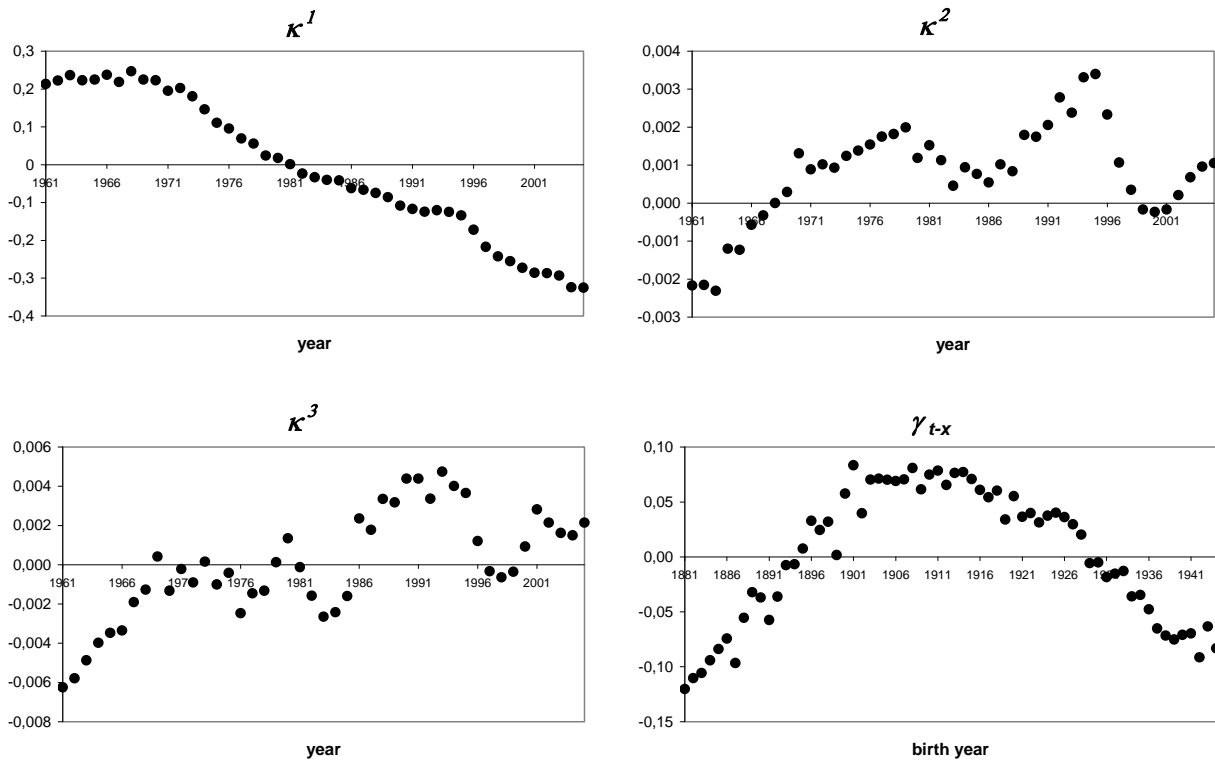
⁵ An alternative way of presentation could be to exclude the birth years > 1945 for all models that include the cohort effect and compare the *BIC* on that basis. However, the other models are all fitted with the Lifemetrics tool that includes all birth years.

4.3 Fitting the ARIMA processes – U.S. Males

In the remainder of this paper, we will focus on the population of U.S. males⁶. The next step in the process is selecting and fitting a suitable ARIMA-process to the time series of κ^1 , κ^2 , κ^3 and γ_{t-x} . The fitted parameters κ^1 , κ^2 , κ^3 and γ_{t-x} for U.S. males are given in figure 2. The figure shows that the pattern of the important parameter κ^1 is well-behaved. The patterns of the other parameters all reveal some autoregressive behavior.

Since the factor κ^1 drives a significant part of the uncertainty in mortality rates, its relatively regular behavior (for this particular dataset) will also show in the projected uncertainty (in other words, the confidence intervals will be relatively narrow).

Figure 2: estimated values of κ^1 , κ^2 , κ^3 and γ_{t-x}



Now for each of these time series all relevant ARIMA(p,d,q) processes for the range $p, d, q = 0, 1, 2, 3$ are fitted and the most favorable process in terms of BIC is selected. The selected ARIMA processes are:

- κ^1 : ARIMA(0,1,0)
- κ^2 , κ^3 and γ_{t-x} : ARIMA(1,0,0), no constant

It is commonly assumed (see Renshaw and Haberman (2006), CMI (2007) and Cairns et al (2008b), that the process for γ_{t-x} is independent of the other processes, so the parameters of this

⁶ To be able to compare simulation results with Cairns et al (2007), we can either use US males or E&W males. The choice for U.S. males is more or less arbitrary.

process can be fitted independently using Ordinary Least Squares (OLS). The other processes can be fitted simultaneously using Seemingly Unrelated Regression (SUR, see Zellner (1963)).

Table 4 gives the fitted parameters, standard errors, t-ratios and *BIC*'s and table 5 shows the fitted standard deviations (on the diagonal) and correlations.

Table 4: fitted parameters for the process $y_{t+1} = \theta_1 + \theta_2 y_t$ *

<i>Fit results</i>	κ^1	κ^2	κ^3	γ_{t-x}
θ_1	-0,0131 0,0022	-5,925		
θ_2		0,9539 0,0495	19,280	0,9366 0,0361
BIC	120,83	267,65	229,21	163,55

* In each cell for θ_1 and θ_2 , top: fitted parameter, bottom left: standard error, bottom right: t-ratio

Table 5: fitted standard deviations (on the diagonal) and correlations

	κ^1	κ^2	κ^3	γ_{t-x}
κ^1	0,0150	0,2539	0,0274	0
κ^2	0,2539	0,0005	0,0144	0
κ^3	0,0274	0,0144	0,0012	0
γ_{t-x}	0	0	0	0,0175

5. Mortality projections – U.S. Males

Cairns et al (2008b) performed an extensive assessment of the out-of-sample performance of several stochastic mortality models, focusing on England & Wales and U.S. Males between 60 and 90 years old. The main criteria used in this assessment were biological reasonableness and robustness of the (stochastic) forecasts. Based on these criteria and specifically for these datasets, they concluded that the models of Lee and Carter (1992), Renshaw and Haberman (2006) and Cairns et al (2007, model M8) did not perform in a satisfactory way. Furthermore, they concluded that the models of Currie (2006) and Cairns et al (2006, 2007 model M7) did produce plausible results and seem robust.

This section shows the simulation results and results of robustness tests for the proposed mortality model.

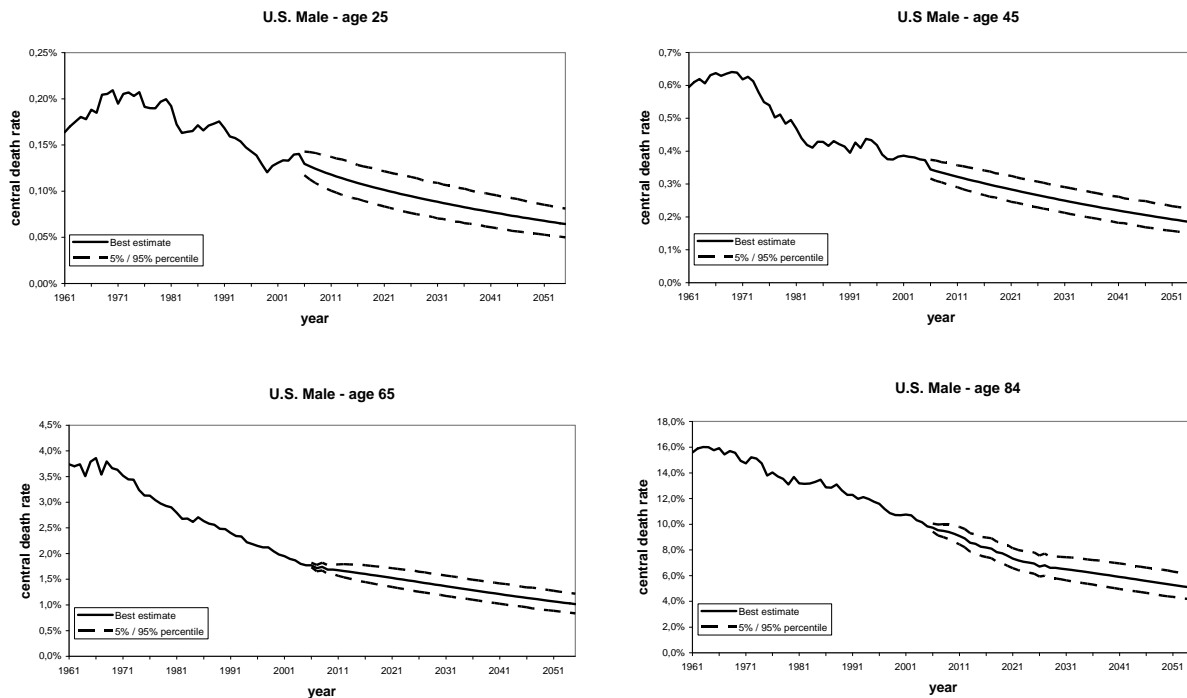
5.1 Simulation results – U.S. Males

Using the fitted ARIMA processes and the fitted values for a_x and γ_{t-x} (see Appendix A), future mortality rate scenarios for U.S. males can be constructed using Monte Carlo simulation. Figure 3 shows simulation results for ages 25, 45, 65 and 84 for U.S. males⁷. The best estimate projection is given and the 5% and 95% percentiles.

⁷ Simulation results for England & Wales and the Netherlands are given in Appendix B.

The results are biologically plausible. For higher ages, the widths of the confidence intervals are broadly similar as the models of which Cairns et al (2008b) concluded that they produced biologically plausible results. The results for younger ages (25 and 45) also seem plausible, where the observed historical variability is reflected in the confidence intervals.

Figure 3: simulation results for U.S. Males



5.2 Robustness of simulation results

Some models suffer from a lack of robustness. For example, Cairns et al (2007, 2008b) find that the Renshaw-Haberman model is not robust for changes in range of years. They link this to the shape of the likelihood function. Robust models probably have a unique maximum that remains broadly unchanged when the range of years or ages is changed. Models that lack robustness possibly have more than one maximum, so when changing the range of years or ages the optimizer can jump from one local maximum to another, yielding different parameter estimates.

The model proposed in this paper is tested for robustness using the same test as in Cairns et al (2008b). This means that the simulation results above are compared with those of two sensitivities. These sensitivities are:

- 1) The model is fitted only to historical data from 1981-2005 (instead of 1961-2005)
- 2) The model (3.1) is fitted to historical data from 1961-2005, but the stochastic models for κ^1 , κ^2 , κ^3 and γ_{t-x} are only fitted to a restricted set of parameter estimates (being only the final 24 $\kappa^{(i)}$'s and the final 45 γ_{t-x} 's)

Of course, if there is a change in trend or variability in the period 1981-2005 compared to 1961-2005, it is inevitable, for all models, that the simulation results will be somewhat different.

The results are given in appendix C and are not significantly different as the results shown in paragraph 5.1. The confidence intervals for age 25 are wider, due to the higher variability for younger ages in the past 25 years. Conclusion is that the proposed model is robust for the sensitivities given above.

Furthermore, a backtest is carried out, meaning that the model is fitted to historical data from 1961-1986 and the forecast results are compared with the actual observations for the period 1987-2005. Also for this backtest, the proposed model performs adequately (see the results in appendix C).

5.3 Comparison with other models

Paragraph 5.1 and 5.2 showed that the proposed model produces plausible results and seems robust. Cairns et al (2008b) came to the same conclusion for the models of Currie (2006) and Cairns et al (2006, 2007 model M7).

The models of Cairns et al (2006, 2007 model M7) are designed for higher ages, so will not produce plausible results for lower ages. Compared to those models the proposed model has the advantage that it does produce plausible results for a full age range.

Compared to the model of Currie (2006) the proposed model has the advantage that it has a non-trivial correlation structure. This is important because often insurers and pension funds have different type of exposures for younger or middle ages (term insurance, pre-retirement spouse option) than for higher ages (pensions, annuities). Aggregating these different types of exposures can only be done sufficiently if the model has a non-trivial correlation structure. Assuming an almost perfect correlation between ages, as in the Currie (2006) model, will possibly lead to an overstatement of the diversification benefits that arise when aggregating these exposures.

6. Risk neutral specification of the model

The model proposed in section 3 is set up in the so-called real-world measure, suitable for assessing risks for example in the context of Solvency 2. For pricing instruments of which the payoff depends on future mortality rates, a risk adjusted pricing measure has to be defined. A common approach is to specify a risk neutral measure Q that is a suitable basis for pricing, see for example Milevsky and Promislow (2001), Dahl (2004), Schrage (2006), Cairns et al (2006a, 2006b) and Biffis et al (2006). The risk neutral specification proposed below is in line with the approach of Cairns et al (2006a).

Note that the market for longevity or mortality instruments is currently (very) far from complete. Consequence of this is that the risk neutral measure Q is not unique. Given the absence of any market price data, it seems wise to keep the specification of the risk neutral process relatively

simple. For the same reason, it is difficult to judge whether one risk neutral mortality model is better than another.

6.1 Risk neutral dynamics

The stochastic process for the factors κ^1 , κ^2 , κ^3 and γ_{t-x} in the real world measure P can generally be specified as:

$$(6.1) \quad K_t = \Theta(K_{t-1}, \varepsilon_{t-1}) + \Sigma Z_t^P$$

Where K_t is the vector with factors κ^1 , κ^2 , κ^3 and γ_{t-x} , $\Theta(K_{t-1}, \varepsilon_{t-1})$ is the drift of the process, $\Sigma\Sigma'$ is the covariance matrix and Z^P is a 4 x 1 vector with standard normal random variables under measure P .

Now the proposed dynamics under the risk neutral measure Q are:

$$(6.2) \quad \begin{aligned} K_t &= \Theta(K_{t-1}, \varepsilon_{t-1}) + \Sigma [Z_t^Q - \lambda] \\ &= \Theta(K_{t-1}, \varepsilon_{t-1}) - \Sigma \lambda + \Sigma Z_t^Q \end{aligned}$$

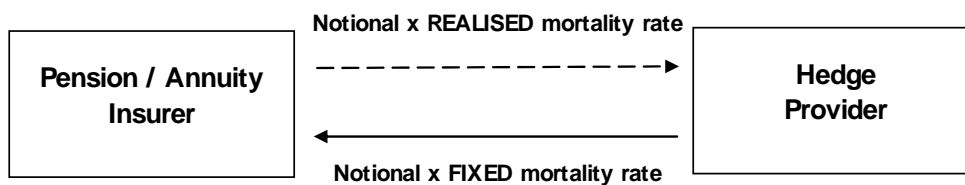
where the vector λ represents the “market price of risk” associated with the process K_t . Like Cairns et al (2006a), we assume that the market price of risk is constant over time. When market prices for longevity or mortality derivatives are available, the vector λ should be calibrated in such a way that the theoretical prices under the measure Q are approximately equal to market prices.

6.2 Calibration of the market price of risk

Currently, there is no developed market for longevity derivatives. However, Loeys et al (2007) have the opinion that q-forwards are most likely to become the basis of a longevity market. Therefore, in this paragraph the risk neutral model (6.2) is calibrated to q-forward prices. Of course, calibration to other instruments such as longevity bonds or survivor swaps, would also be possible.

A q-forward is a simple capital market instrument with similar characteristics as an interest rate swap. The instrument exchanges a realized mortality rate in a future period for a pre-agreed fixed mortality rate. This is shown in figure 4. The pre-agreed fixed mortality rate is based on a projection of mortality rates, coming from the Lifemetrics toolkit.

Figure 4: working of a q-forward



For example, when the realized mortality rate is lower than expected, the pension / annuity insurer will receive a payment which (partly) compensates the increase of the expected value of the insurance liabilities (caused by the decreasing mortality rates).

The basis for the instrument is the (projected) mortality of a country population, not the mortality of a specific company or portfolio. This makes the product and the pricing very transparent compared to traditional reinsurance.

Although there have been some transactions involving q-forwards, currently no market quotes for q-forwards are publicly available. However, Loeys et al (2007) give an indication and examples on how such an instrument would be priced in practice. In absence of real market data, we calibrate the model to q-forward prices resulting from these examples.

Loeys et al (2007) give the following formula for setting the fixed q-forward rate:

$$(6.3) \quad q_{forward} = (1 - horizon * Sharpe\ ratio * q_{vol}) * q_{expected}$$

where they have used 10 years for the horizon of the derivative, 25% for the Sharpe ratio and the volatility q_{vol} based on historical data. Table 6 shows the results for q-forwards with a horizon of 10 and starting ages of 35, 45, 55 and 65, where $q_{expected}$ is based on model (3.1). Since in this paper the central mortality rate $m_{x,t}$ is modeled, the results are also translated these terms, which makes the calibration easier.

Table 6: indication q-forward rate for horizon 10 and translation to m-forward

Age start	Age end	q_{vol}	$q_{expected}$	$q_{forward}$	$m_{expected}$	$m_{forward}$	h
35	45	2,31%	0,306%	0,288%	0,307%	0,289%	0,060
45	55	1,53%	0,709%	0,682%	0,712%	0,685%	0,039
55	65	1,01%	1,618%	1,578%	1,632%	1,590%	0,026
65	75	1,47%	3,542%	3,412%	3,606%	3,471%	0,038

Now interpreting $m_{forward}$ as the expectation under the risk neutral measure and $m_{expected}$ as the expectation under real world measure leads to:

$$(6.4) \quad E^Q(m_{x_{end}, t_{end}}) = g E^P(m_{x_{end}, t_{end}})$$

where x_{end} and t_{end} are age and year at the end of the contract and g can be extracted from the market (or in this case, from table 6). Taking logarithms leads to:

$$(6.5) \quad \ln[E^Q(m_{x_{end}, t_{end}})] = \ln[g] + \ln[E^P(m_{x_{end}, t_{end}})]$$

Because the only difference between the processes under Q and P is in the drift term, we can assume that:

$$(6.6) \quad \ln \left[E^Q \left(m_{x_{end}, t_{end}} \right) \right] - \ln \left[E^P \left(m_{x_{end}, t_{end}} \right) \right] = E^Q \left(\ln \left[m_{x_{end}, t_{end}} \right] \right) - E^P \left(\ln \left[m_{x_{end}, t_{end}} \right] \right) = \ln [g]$$

Now since this difference in the drift term is the matrix $-\Sigma\lambda$, for a horizon k the following holds:

$$(6.7) \quad E^Q \left(\ln \left[m_{x_{end}, t_{end}} \right] \right) - E^P \left(\ln \left[m_{x_{end}, t_{end}} \right] \right) = - \sum_{t=1}^k W_t \Sigma \lambda$$

where W_t is the matrix of weights that are used to translate the values for κ^1 , κ^2 , κ^3 and γ_{t-x} into values for $\ln(m_{x,t})$. This leads to:

$$(6.8) \quad h = \sum_{t=1}^k W_t \Sigma \lambda$$

where $h = -\ln[g]$, of which the values are given in the table above. From (6.8) the market prices of risk can be solved:

$$(6.9) \quad \hat{\lambda} = \left(\sum_{t=1}^k W_t \Sigma \right)^{-1} h$$

Now we use this formula to calibrate the market prices of risk to the q-forwards specified above. The weights matrices W_t vary slightly for each year t depending on the development of the age:

$$(6.10) \quad W_t = \begin{pmatrix} 1 & (\bar{x} - 45) & (\bar{x} - 45)^+ & 1 \\ 1 & (\bar{x} - 55) & (\bar{x} - 55)^+ & 1 \\ 1 & (\bar{x} - 65) & (\bar{x} - 65)^+ & c_t \\ 1 & (\bar{x} - 75) & (\bar{x} - 75)^+ & 0 \end{pmatrix}$$

where $c_t = 0$ for $t < 5$ and $c_t = 1$ otherwise. Reason for this time dependence is that the simulated cohort effect gradually comes into the projections for age 65. The right bottom item of W_t is 0 because for age 75 the cohort effect does not play a role within the horizon of 10 year for this age.

Applying formula (6.9) using (6.10), the results in table 5 and the vector h from table 6 leads to the calibrated market prices of risk given in table 7.

Table 7: market prices of risk

	λ
κ^1	1,2430
κ^2	0,9793
κ^3	-0,5756
γ_{t-x}	-0,7338

When the market develops and a number of q-forward prices are available, the market prices of risk can be calibrated by minimizing the squared errors between the theoretical prices and the market prices:

$$(6.11) \quad \hat{\lambda} = \underset{\lambda}{\text{Min}} \sum_{i=1}^p \left(h_i - \sum_{t=1}^k W_t^i \Sigma \lambda \right)^2$$

where p is the number of q-forwards the model is calibrated to.

7. Parameter Uncertainty

As mentioned in the criteria for stochastic mortality models in paragraph 2.1, the structure of the model should make it possible to incorporate parameter uncertainty in simulations. There are three possible approaches for including this parameter uncertainty:

- 1) Using a formal Bayesian framework, see Cairns (2000) and Cairns et al (2006a)
- 2) Simulate the parameter values using the estimates and the standard errors obtained in the estimation process
- 3) Applying a bootstrap procedure such as described in Brouhns et al (2005) and Renshaw and Haberman (2008)

In a Bayesian framework a prior, possibly non-informative, distribution is assumed for the parameters. Combining this prior distribution with the sample data and the assumed density function of a particular stochastic process leads to a posterior distribution. This posterior distribution can be used to assess the parameter uncertainty.

Approach 2) uses the standard errors of the fitted parameters to incorporate the parameter uncertainty. When least squares or maximum likelihood estimation is used the estimators are either normally or asymptotically normally distributed.

Approach 3) uses bootstrapping techniques, either applied to $D_{x,t}$ (semi-parametric bootstrap) or to the residuals $D_{x,t} - \tilde{D}_{x,t}$ (residual bootstrap),

While a formal Bayesian approach is more elegant than approach 2) and 3), it generally leads to significantly more complexity. Market Chain Monte Carlo (MCMC) methods or importance sampling might be necessary, because the posterior distribution often does not belong to a known class of probability density functions (see for example Kleibergen and Hoek (1996)). Since the approaches should not lead to significantly different parameter uncertainty, it is questionable whether it is worth increasing the complexity of the model significantly for slightly more elegance. Therefore, approach 1) does not have our preference.

By using approach 2), parameter uncertainty can be incorporated in the model proposed in this paper. For the stochastic processes of κ^1 , κ^2 , κ^3 and γ_{t-x} the estimates and standard errors given in

table 4 can be used as the moments of the normal distributions of the parameters. For the parameter estimates of the γ_{t-x} 's (until birth year 1945) and the a_x 's the standard errors have to be calculated separately. Starting point for this (see for example Verbeek (2008)) is the vector of first derivatives of the log-likelihood function, the so-called score vector $s(\phi)$:

$$(7.1) \quad s(\phi) = \frac{\partial L(\phi)}{\partial \phi} = \sum_{i=1}^N \frac{\partial L_i(\phi)}{\partial \phi} = \sum_{i=1}^N s_i(\phi)$$

where ϕ is the parameter set, $L_i(\phi)$ is de log-likelihood function for observation i and N is the number of observations. Now the covariance matrix V_{par} to be used can be estimated with:

$$(7.2) \quad V_{par} = \left(\sum_{i=1}^N s_i(\hat{\phi}) s_i(\hat{\phi})' \right)^{-1}$$

The standard errors for the γ_{t-x} 's (until birth year 1945) and the a_x 's are the square roots of the relevant diagonal elements.

Approach 2) is the most practical method. However, Renshaw and Haberman (2008) noted that the confidence intervals for the Lee-Carter and Renshaw-Haberman models vary for different versions of identifiability constraints when using this method. This phenomenon was not seen when using approach 3). Although the question remains whether their conclusion still holds for other models (such as the one proposed in this paper) and different sort of constraints (such as the ones used in this paper and in Cairns et al (2007, model M6), approach 3) seems the most appropriate method for addressing parameter uncertainty in the model proposed in this paper.

8. Conclusions

All well known stochastic mortality models have nice features but also disadvantages. In this paper a stochastic mortality model is proposed that aims at combining the nice features from existing models, while eliminating the disadvantages.

The paper shows that the fit of the model to historical data is better than the well-known mortality models. Also, the model has 4 stochastic factors, leading to a (desired) non-trivial (but not too complex) correlation structure between ages. Due to a (Lee-Carter type) variable that describes the shape of the mortality curve over ages and the inclusion of a separate stochastic factor for young ages, the model is applicable to a full age range. Furthermore, the model captures the cohort effect and has no robustness problems. The paper also describes how to incorporate parameter uncertainty into the model.

In paragraph 2.1, a list of criteria for stochastic mortality models is given. Table 8 shows whether the existing models and the proposed model meet those criteria. A large part of the table is based on Cairns et al (2007) and the conclusions in Cairns et al (2008b).

Table 8: comparison of stochastic mortality models – satisfaction of criteria *

Satisfaction of criteria mortality models	Renshaw & Haberman	Currie	Lee & Carter	Cairns et al (2006)	Cairns et al (2007, M7)	Plat
1) Positive mortality rates	+	+	+	+	+	+
2) Consistency historical data	+	+	+ / -	+	+	+
3) Long-term biological reasonableness	+	+	+	+	+	+
4) Robustness	-	+	+	+	+	+
5) Forecasts biological reasonable	+	+	+/-	+	+	+
6) Ease of implementation	+	+	+	+	+	+
7) Parsimony	+/-	+/-	+	+	+/-	+/-
8) Possibility generating sample paths	+	+	+	+	+	+
9) Allowance for parameter uncertainty	+	+	+	+	+	+
10) Incorporation cohort effects	+	+	-	-	+	+
11) Non-trivial correlation structure	+/-	+/-	-	+	+	+
12) Applicable for full age range	+/-	+/-	+/-	-	-	+

* +: meets criterion, +/-: partly meets criterion, -: does not meet criterion

The table shows that, apart from partly meeting the parsimony criteria, the proposed model meets all of the criteria. None of the existing models meet all of the criteria. Of the existing models, the model of Currie (2006) is most close to meeting all criteria. However, the advantages of the proposed model compared to the model of Currie (2006) are:

- Better fit to historical data
- Non-trivial correlation structure, which is important in solvency calculations
- Better applicable to a full age range, amongst others due to the inclusion of a separate factor for younger ages

So by combining the nice features of the existing models, the proposed model has eliminated the disadvantages of those models, and as a result the model meets all of the criteria set for stochastic mortality models.

For pricing purposes, a risk neutral version of the model is given, that can be used for pricing. This model is calibrated to some indicative prices for longevity derivatives.

Appendix A: U.S. Male - estimates for a_x and γ_{t-x}

age	a_x
20	-6,3983
21	-6,3328
22	-6,3382
23	-6,3483
24	-6,3826
25	-6,3812
26	-6,3777
27	-6,3645
28	-6,3301
29	-6,3426
30	-6,3080
31	-6,2548
32	-6,2078
33	-6,1571
34	-6,1303
35	-6,0643
36	-6,0026
37	-5,9383
38	-5,8493
39	-5,8183
40	-5,7386
41	-5,6532
42	-5,5658
43	-5,4921
44	-5,4221
45	-5,3288
46	-5,2423
47	-5,1539
48	-5,0575
49	-4,9965
50	-4,8950
51	-4,8109
52	-4,7155

age	a_x
53	-4,6403
54	-4,5587
55	-4,4660
56	-4,3843
57	-4,3008
58	-4,1935
59	-4,1249
60	-4,0197
61	-3,9471
62	-3,8376
63	-3,7746
64	-3,7040
65	-3,6100
66	-3,5483
67	-3,4648
68	-3,3801
69	-3,3068
70	-3,2088
71	-3,1444
72	-3,0406
73	-2,9656
74	-2,8864
75	-2,7994
76	-2,7193
77	-2,6375
78	-2,5568
79	-2,4671
80	-2,3649
81	-2,2824
82	-2,1863
83	-2,0906
84	-1,9992

birth year	γ_{t-x}
1881	-0,1202
1882	-0,1103
1883	-0,1057
1884	-0,0940
1885	-0,0839
1886	-0,0743
1887	-0,0967
1888	-0,0553
1889	-0,0321
1890	-0,0370
1891	-0,0574
1892	-0,0361
1893	-0,0074
1894	-0,0067
1895	0,0075
1896	0,0329
1897	0,0244
1898	0,0321
1899	0,0016
1900	0,0576
1901	0,0834
1902	0,0396
1903	0,0703
1904	0,0711
1905	0,0702
1906	0,0690
1907	0,0705
1908	0,0809
1909	0,0615
1910	0,0749
1911	0,0785
1912	0,0655
1913	0,0765

birth year	γ_{t-x}
1914	0,0773
1915	0,0708
1916	0,0609
1917	0,0544
1918	0,0603
1919	0,0340
1920	0,0554
1921	0,0366
1922	0,0400
1923	0,0313
1924	0,0376
1925	0,0402
1926	0,0361
1927	0,0297
1928	0,0203
1929	-0,0057
1930	-0,0050
1931	-0,0184
1932	-0,0152
1933	-0,0128
1934	-0,0360
1935	-0,0347
1936	-0,0477
1937	-0,0651
1938	-0,0717
1939	-0,0751
1940	-0,0709
1941	-0,0697
1942	-0,0914
1943	-0,0632
1944	-0,0831
1945	-0,0702

Appendix B: simulation results England & Wales and the Netherlands

In this appendix the simulation results for England & Wales (E&W) and the Netherlands are given. The best estimate projection is given and the 5% and 95% percentiles. Information about the fitted parameters and underlying ARIMA processes is available upon request.

Figure B1: simulation results for England & Wales males

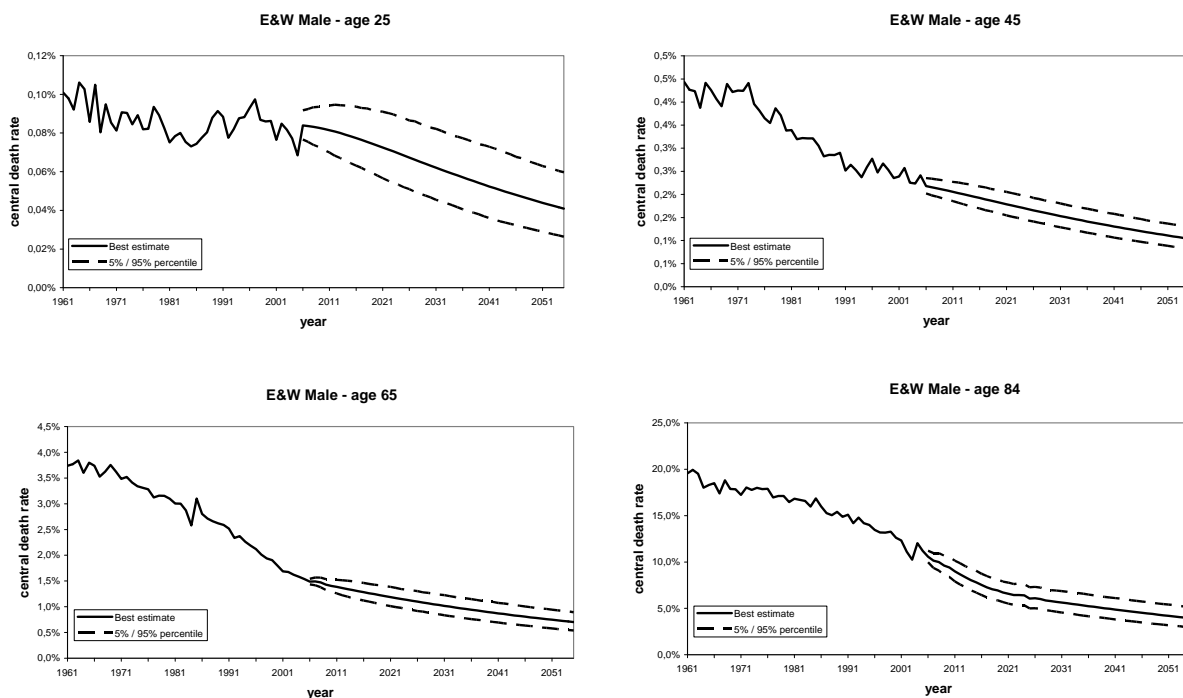
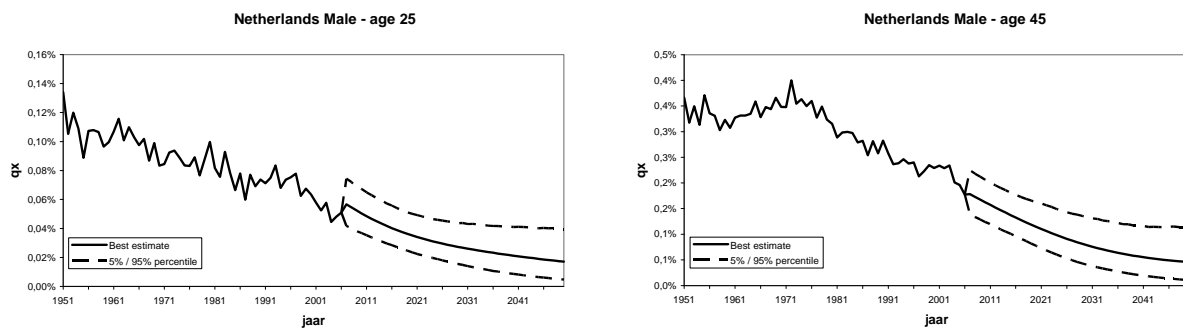
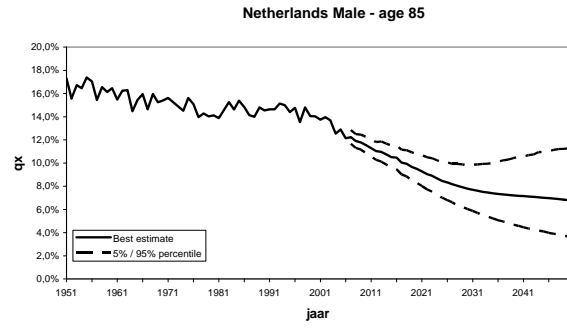
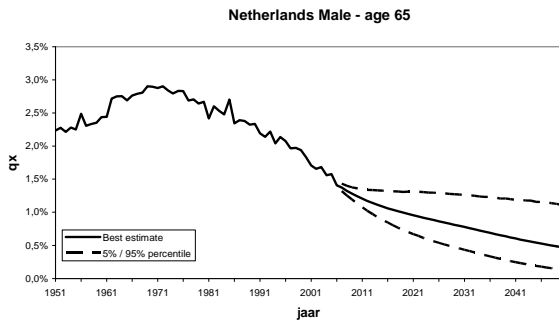


Figure B2: simulation results the Netherlands





Appendix C: simulation results robustness tests

In this appendix the simulation results are given for the sensitivities that have been specified to test the robustness of the model:

- 1) The model is fitted only to historical data from 1981-2005 (instead of 1961-2005)
- 2) The model (3.1) is fitted to historical data from 1961-2005, but the stochastic models for κ^1 , κ^2 , κ^3 and γ_{t-x} are only fitted to a restricted set of parameter estimates (being only the final 24 $\kappa^{(i)}$'s and the final 45 γ_{t-x} 's)

Figure C.1: simulation results sensitivity 1

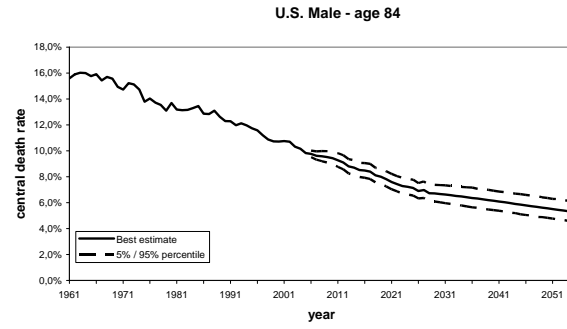
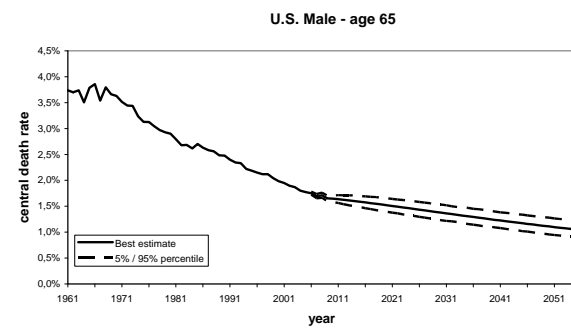
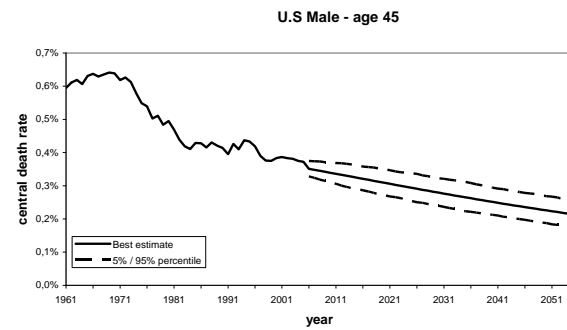
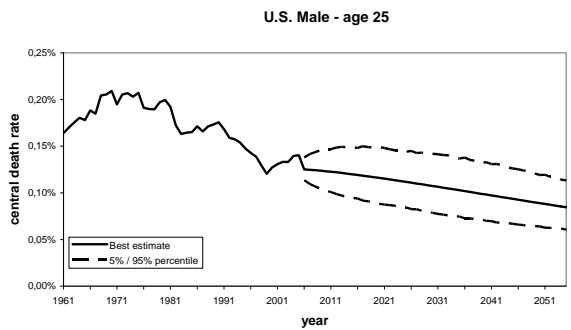


Figure C.2: simulation results sensitivity 2)

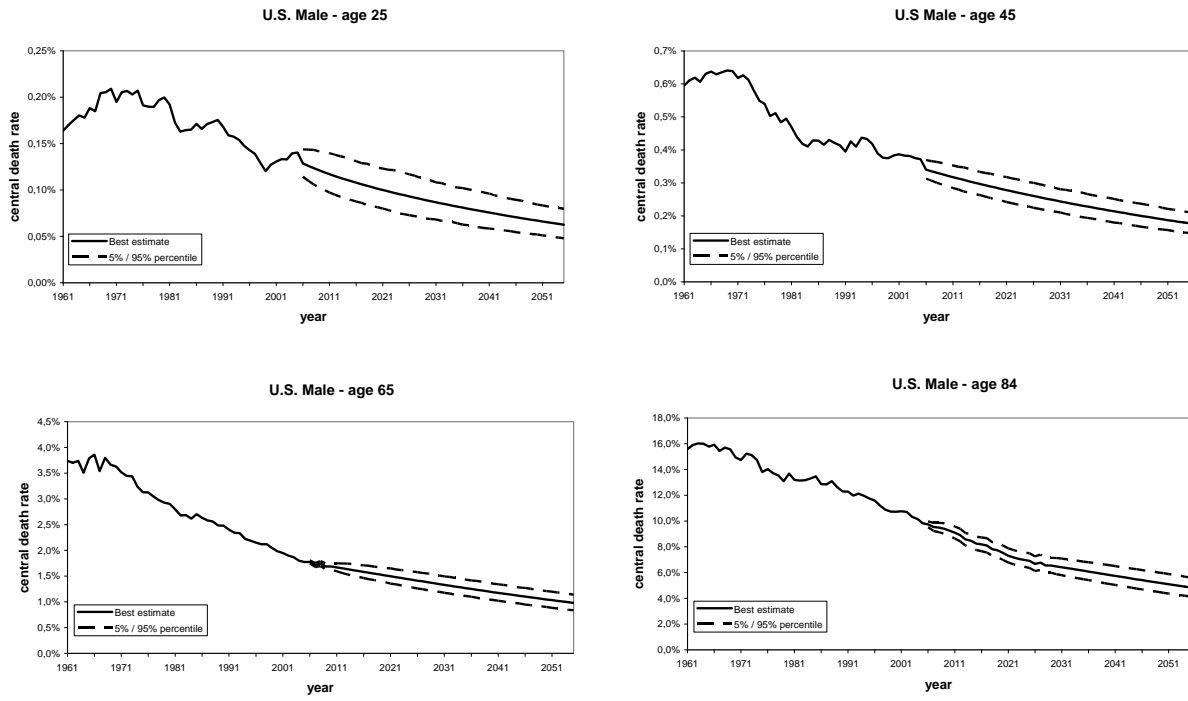
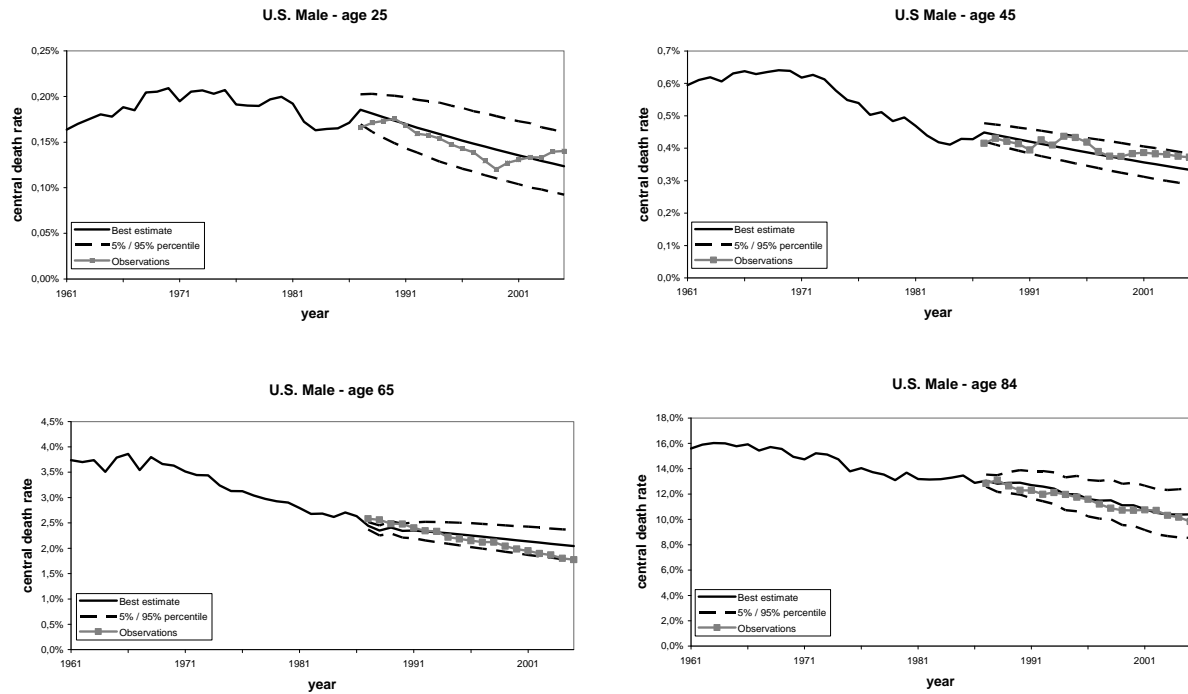


Figure C3: simulation results backtest



References

- BIFFIS, E., M. DENUIT, AND P. DEVOLDER (2006): Stochastic mortality under measure changes, *Pensions Institute Discussion Paper PI-0512*
- BOOTH, H., J. MAINDONALD AND L. SMITH (2002): Applying Lee-Carter under conditions of variable mortality decline, *Population Studies* 56, 325-336
- BROUHNS, N., M. DENUIT, AND J.K. VERMUNT (2002): A Poisson log-bilinear regression approach to the construction of projected life tables, *Insurance: Mathematics and Economics* 31, 373-393
- BROUHNS, N., M. DENUIT, AND I. VAN KEILEGOM (2005): Bootstrapping the Poisson log-bilinear model for mortality forecasting, *Scandinavian Actuarial Journal* 3, 212-224
- CAIRNS, A.J.G. (2000): A discussion of parameter and model uncertainty in insurance, *Insurance: Mathematics and Economics* 27, 313-330
- CAIRNS, A.J.G., D. BLAKE, AND K. DOWD (2006a): A two-factor model for stochastic mortality with parameter uncertainty: Theory and Calibration, *Journal of Risk and Insurance* 73, 687-718
- CAIRNS, A.J.G., D. BLAKE, AND K. DOWD (2006b): Pricing death: Frameworks for the valuation and securitization of mortality risk, *ASTIN Bulletin* 36, 79-120
- CAIRNS, A.J.G., D. BLAKE, K. DOWD, G. D. COUGHLAN, D. EPSTEIN, A. ONG AND I. BALEVICH (2007): A quantitative comparison of stochastic mortality models using data from England & Wales and the United States, *Working paper, Heriot-Watt University, and Pensions Institute Discussion Paper PI-0701*
- CAIRNS, A.J.G., D. BLAKE, AND K. DOWD (2008a): Modelling and Management of Mortality Risk: A Review, *Working paper, Heriot-Watt University, and Pensions Institute Discussion Paper PI-0814*
- CAIRNS, A.J.G., D. BLAKE, K. DOWD, G. D. COUGHLAN, D. EPSTEIN AND M. KHALLAF-ALLAH (2008b): The plausibility of mortality density forecasts: an analysis of six stochastic mortality models, *Working paper, Heriot-Watt University, and Pensions Institute Discussion Paper PI-0801*
- CONTINUOUS MORTALITY INVESTIGATION (2007): Stochastic projection methodologies: Lee-Carter model features, example results and implications, *Working paper* 25
- COUGHLAN, G., D. EPSTEIN, A. ONG, A. SINHA, I. BALEVICH, J. HEVIA-PORTOCARRERO, E. GINGRICH, M. KHALAF-ALLAH AND P. JOSEPH (2007): Lifemetrics Technical Document, available at: <http://www.jpmorgan.com/pages/jpmorgan/investbk/solutions/lifemetrics>
- CURRIE, I.D. (2006): Smoothing and forecasting mortality rates with P-splines, *Talk given at the Institute of Actuaries, June 2006*, available at: <http://www.ma.hw.ac.uk/~iain/research/talks.html>
- DAHL, M. (2004): Stochastic mortality in life insurance: Market reserves and mortality-linked insurance contracts, *Insurance: Mathematics and Economics* 35, 113-136
- DELWARDE, A., M. DENUIT AND P. EILERS (2007): Smoothing the Lee-Carter and Poisson log-bilinear models for mortality forecasting: A penalized log-likelihood approach, *Statistical Modelling* 7, 29-48
- DE JONG, P., AND L. TICKLE (2006): Extending the Lee-Carter model of mortality projection, *Mathematical Population Studies* 13, 1-18
- KLEIBERGEN, F., AND H. HOEK (1996): Bayesian Analysis of ARMA Models using Noninformative Priors, *Tinbergen Institute discussion paper*
- LEE, R.D., AND L.R. CARTER (1992): Modelling and forecasting U.S. mortality, *Journal of the*

American Statistical Association 87, 659-675

LEE, R., AND T. MILLER (2001): Evaluating the performance of the Lee-Carter model for forecasting mortality, *Demography* 38, 537-549

LOEYS, J., N. PANIGIRTZOGLU AND R.M. RIBEIRO (2007): Longevity: a market in the making, available at: <http://www.jpmorgan.com/pages/jpmorgan/investbk/solutions/lifemetrics>

MILEVSKY, M.A., AND S.D. PROMISLOW (2001): Mortality derivatives and the option to annuities, *Insurance: Mathematics and Economics* 29, 299-318

RENSHAW, A.E., AND S. HABERMAN (2003): Lee-Carter mortality forecasting with age-specific enhancement, *Insurance: Mathematics and Economics* 33, 255-272

RENSHAW, A.E., AND S. HABERMAN (2006): A cohort-based extension to the Lee-Carter model for mortality reduction factors, *Insurance: Mathematics and Economics* 38, 556-570

RENSHAW, A.E., AND S. HABERMAN (2008): On simulation-based approaches to risk measurement in mortality with specific reference to Poisson Lee-Carter modelling, *Insurance: Mathematics and Economics* 42, 797-816

SCHRAGER, D.F. (2006): Affine stochastic mortality, *Insurance: Mathematics and Economics* 38, 81-97

VERBEEK, M. (2008): Modern Econometrics, 3th edition, *John Wiley & Sons, Ltd*

WILLETS, R.C. (2004): The cohort effect: Insights and explanations, *British Actuarial Journal* 10, 833-877

ZELLNER, A. (1963): Estimators of Seemingly Unrelated Regressions: Some Exact Finite Sample Results, *Journal of the American Statistical Association* 58, 977-992