Modelling Longevity Bonds: Analysing the Swiss Re Kortis Bond

Andrew Hunt*

Cass Business School, City University London

Corresponding author: andrew.hunt.1@cass.city.ac.uk

David Blake

Pensions Institute, Cass Business School, City University London

May 2015

Abstract

A key contribution to the development of the traded market for longevity risk was the issuance of the Kortis bond, the world's first longevity trend bond, by Swiss Re in 2010. We analyse the design of the Kortis bond, develop suitable mortality models to analyse its payoff and discuss the key risk factors for the bond. We also investigate how the design of the Kortis bond can be adapted and extended to further develop the market for longevity risk.

JEL Classification: C15, C32, G13

Keywords: Mortality modelling, age/period/cohort models, general procedure, cointegration, cohort effects, Kortis bond

*Material in this paper was presented under the title "Projecting mortality: Coherence and co-integration in multi-population projections" at the Longevity 9 conference in Beijing, China, in 2013. We are grateful to participants there, at a seminar at Cass Business School, to Bent Nielsen and Michele Bergamelli for discussions regarding identifiability and cointegration, and to Andrés Villegas for many useful discussions on this and related topics. We would also like to thank Daniel Harrison FIA at Swiss Re and the anonymous referee for their very valuable comments regarding this paper.

1 Introduction

The traded market for longevity risk continues to grow and develop. As the risks posed by increasing longevity for the providers of pensions and annuities have gained greater prominence, a variety of different vehicles have been proposed and implemented to transfer longevity risk to the capital markets. These have included bonds, swaps and forwards, each linked to different measures of mortality rates and survivorship.

A key contribution to the development of the market was the issuance of the Kortis bond by Swiss Re in 2010. Unlike previous mortality and longevity securitisations, the Kortis bond is linked to the divergence in mortality improvement rates between two countries, rather than to mortality rates directly or to survivorship amongst a cohort. As such, it was promoted as the first "longevity trend bond". The bond might herald a distinctly new way of transferring the risk of faster than expected reductions in mortality rates, from insurers and reinsurers to investors willing to hold these risks as part of a diversified portfolio.

The development of new longevity-linked securities has been aided by and, in turn, encouraged the development of increasingly sophisticated mortality models. These are necessary in order to estimate accurately the risk present in such securities. In particular, they need to project mortality rates with complex correlation structures, robustly estimated cohort effects and dependencies between different populations. Such projections also need to be "well-identified", in the sense that they do not depend on the arbitrary choices we make in order to fit the model to historical data.

In this paper, we analyse the Kortis bond in detail and critique several of its features. In Section 2, we discuss the design of the Kortis bond and some of the reasons why this particular structure was chosen. We then develop a model suitable for projecting the index to which the payoff of the bond is linked, first by discussing the models used to fit the historical data in Section 3.1. We then outline the methods used to project the period and cohort terms in these models into the future in Sections 3.2 and 3.3, which were developed more fully in Hunt and Blake (2015b) and Hunt and Blake (2015a), respectively. In Section 4, we compare the results of our model to those from the proprietary model used by Risk Management Solutions (RMS) on behalf of Standard and Poors' in the initial pricing of the bond. We go on to analyse the key factors determining this payoff in Section 5. Section 6 discusses how the design of the Kortis bond can be adapted and extended in order to develop a deeper traded market for longevity-linked securities. Finally, Section 7 concludes.

2 The Swiss Re Kortis bond

In December 2010, Swiss Re issued an index-linked note with a principal of \$50m through the Kortis Capital special purpose vehicle. This note, the "Kortis bond", pays quarterly coupons at a floating rate of 5.0% above LIBOR with a return of principal at maturity in January 2017 (extendable until July 2019). It was promoted as the first "longevity trend bond", since the principal is linked to an index measuring the divergence in mortality rates between male populations in England & Wales[1] and the United States. At maturity, if this Longevity Divergence Index Value (LDIV) is greater than 3.4%, then the principal of the bond is reduced linearly until full exhaustion of the principal if the LDIV is greater than 3.9%. In this respect, it is similar to the Vita bond issued by Swiss Re in 2004 and discussed in Blake et al. (2006) and Bauer and Kramer (2007) and valued in Beelders and Colarossi (2004).

The Kortis bond was designed to hedge Swiss Re's exposure to longevity risk in the two countries, which typically arises from reinsuring life assurance policies in the US and annuity policies in the UK. These risk sources would be broadly considered as a partial "natural hedges" of each other (as discussed in Cox and Lin (2007) and Wang et al. (2009)), in the sense that faster reductions in mortality rates would be expected to increase the present value of UK annuity business, but decrease the present value of US life assurance business. Nevertheless, there is substantial basis risk between these populations, since

- trends in mortality can be different in the US and UK, owing to the different lifestyles and medical systems operating on either side of the

[1] Two of the four constituent nations of the UK, together comprising approximately 90% of the population. Data for England & Wales is collected separately and more rapidly than data for Scotland and Northern Ireland, and so is used in the calculation of the index for the Kortis bond.

Atlantic, and

- holders of annuity policies are typically older than life assurance policyholders and rates of longevity improvement differ across ages.

The implication of basis risk is that the natural hedge between the two sets of policies will be imperfect, since they will not experience perfectly correlated price movements. The Kortis bond was designed to partially hedge this basis risk, so that sufficient divergence in mortality rates in the two populations would lead to a lower payoff on the bond, which offsets the adverse experience and higher reserve requirements faced by Swiss Re. To achieve this, the LDIV is constructed in several steps.[2]

First, the observed improvement (i.e., reduction) in mortality rates in each population for men at different ages is defined as

$$\text{Improvement}_n^{(p)}(x,t) = 1 - \left[\frac{m_{x,t}^{(p)}}{m_{x,t-n}^{(p)}}\right]^{\frac{1}{n}} \tag{1}$$

where $m_{x,t}^{(p)}$ is the crude mortality rate observed at age x in year t for population p and n is the averaging period of the bond. An averaging period of eight years is used to calculate the observed improvement, so that the Kortis bond is indexed to changes in mortality rates over the period 1 January 2009 to 31 December 2016.

Second, an improvement index is calculated for each year and country using

$$\text{Index}(t,p) = \frac{1}{1+x_2-x_1}\sum_{x=x_1}^{x_2} \text{Improvement}_8^{(p)}(x,t) \tag{2}$$

This index represents the average improvement rate observed across ages x_1 to x_2 in the eight years to t in population p. In order to hedge the basis risk arising from the different average ages of the relevant populations in each country, Index(t,EW) is calculated using ages between 75 and 85 (i.e.,

[2] Details of the construction and valuation of the Kortis bond are taken from Standard and Poor's (2010) and are also presented in Blake et al. (2013).

$x_1 = 75$ and $x_2 = 85$), whilst Index(t,US) is calculated using ages between 55 and 65. Finally, the LDIV is calculated for year t as

$$\text{LDIV}(t) = \text{Index}(t, \text{EW}) - \text{Index}(t, \text{US}) \tag{3}$$

The principal of the Kortis bond is then reduced by a "principal reduction factor" (PRF) given by

$$\text{PRF} = max\left(min\left(\frac{\text{LDIV}(2016) - 3.4\%}{3.9\% - 3.4\%}, 1\right), 0\right) \tag{4}$$

where 3.4% is referred to as the "point of attachment" and 3.9% is referred to as the "point of exhaustion". We see, therefore, that the bond is structured so that if the rate of improvement seen in England & Wales between 2008 and 2016 is significantly higher than that observed in the US, then Swiss Re's liability with respect to the bond will be reduced, offsetting the adverse experience and higher reserves required on the annuity book. The points of attachment and exhaustion were set so that it is unlikely that the principal of the Kortis bond will be reduced at maturity, implying that the purpose of the Kortis bond is to only hedge tail basis risk. This may be because modern regulatory regimes, such as Solvency II in the EU[3] and the Swiss Solvency Test in Switzerland, place high capital charges on tail risks held on life insurers' balance sheets, encouraging the mitigation of such risks.

In exchange for assuming the risk that the principal is reduced, holders of the Kortis bond are paid quarterly floating coupons at a margin of 5.0% above the three-month LIBOR value.[4] Lane and Beckwith (2011, 2012, 2013, 2014) published average indicative quotations for the spread on the Kortis bond (and other insurance-linked securities) at the end of each quarter, which are shown in Figure 1.[5] These indicate that the "market" spread over LIBOR on the Kortis bond has decreased over the period since the bond was issued, in line with most other fixed interest securities. However, as this information is based on indicative quotations rather than actual trades it is unclear what prices (if any) the Kortis bond has traded at since December 2010.

[3]See EIOPA (2014).

[4]For comparison, it is interesting to note that the Vita bond was issued with a coupon rate of 1.35% above LIBOR.

[5]We are indebted to the anonymous referee for bringing this source of information to our attention.

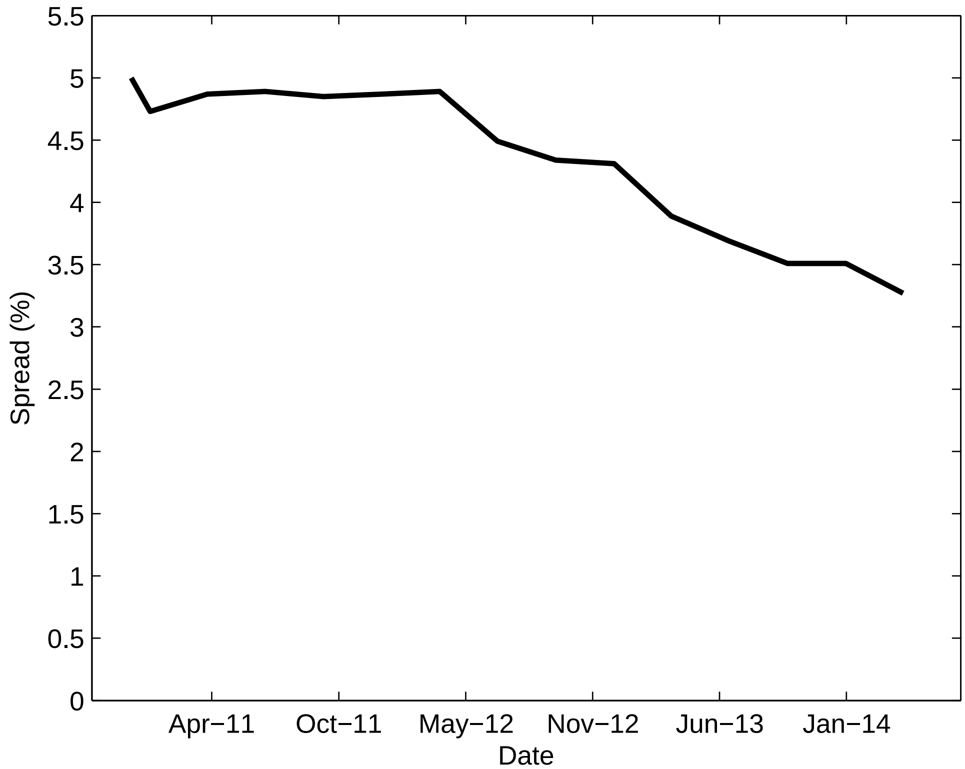


Figure 1: Spread of Kortis bond over LIBOR

It is worth noting that Swiss Re is both the issuer of the bond and the hedger of risk in the transaction, whilst the bond's investors are the hedge providers and assume the risk that the principal is reduced in return for receiving the spread above LIBOR. It is also structured as an out-of-the-money option on the LDIV, which not only means that only the tail basis risk is transferred from Swiss Re's perspective, but also reassures investors that they will see a full return of principal in the majority of scenarios. These features potentially make the bond more appealing for all parties.

The calculation of the LDIV is also, in many ways, very natural. Actuaries have, traditionally, talked of "mortality improvement rates" when discussing mortality risk, i.e., the year-on-year percentage change in mortality rates.[6] However, observed year-on-year improvement rates are subject to considerable uncertainty and so could not be used directly as the index for the Kortis bond. Instead, they are averaged, first over the eight year averaging period in Equation 1 and then over the ten different ages in Equation 2. This process reduces much of the "noise" in individual mortality rates, so that a clearer picture of the underlying improvement rates can be observed. The final step in Equation 3 is to compare the improvement rates in England & Wales with those in the US to check for divergence. However, this process compares individuals from different years of birth and so introduces a sensitivity to cohort effects into our measurement of the LDIV, as discussed in Section 5.

Figure 2 shows the values for the LDIV that would have been calculated between 1958 and 2010 based on data for England & Wales and the US from the Human Mortality Database (Human Mortality Database (2014)). It can be seen that the historical LDIV has not reached the point of attachment within the past 50 years, which supports the conjecture that the Kortis bond was designed to hedge tail basis risk. However, it is important to consider the evolution of the underlying mortality rates used to calculate the LDIV rather than just look at the aggregate movements of the index.

When the Kortis bond was issued, Standard and Poor's rated the bond as BB+. This was based in part on the modelling work of RMS which used its proprietary mortality model based on "vitagion categories" to project mortality rates for England & Wales and the US.[7] Table 1 gives the distribution of the PRF as calculated by RMS and given in Standard and Poor's (2010).[8] These give a conditional expected loss (i.e., the mean loss on the bond given a loss occurs) of approximately 62%, as calculated by Lane and Beckwith (2011). This is broadly comparable with the expected losses on other insurance-linked securities shown in Lane and Beckwith (2011).

[6]For example, see Continuous Mortality Investigation (2002) and also Haberman and Renshaw (2012) and Mitchell et al. (2013) for recent academic studies which take the same approach.

[7]See Blake et al. (2013) for a brief description of this approach.

[8]The figures in Table 1 are cumulative, calculated from the annualised exceedance probabilities in Table 3 of Standard and Poor's (2010).

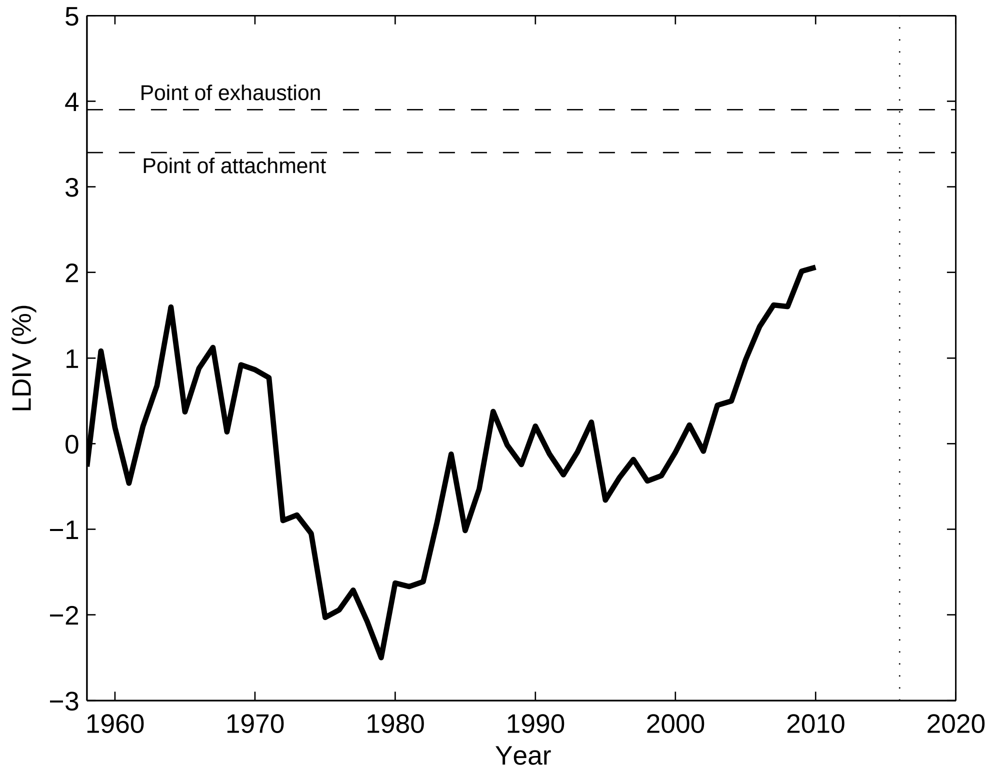


Figure 2: Historical LDIV

3 Mortality models for England & Wales and the US

3.1 Fitting mortality models for each population

The first stage in assessing the Kortis bond is to construct suitable mortality models for England & Wales and the US. To do this, we apply the "general procedure" described in Hunt and Blake (2014) to data from the Human Mortality Database for both countries from 1950 to 2008 and for ages 50 to 100. We use data to 2008, to ensure comparability with the results of RMS, which would only have data available to 2008 when valuing the Kortis bond at issue in 2010.

LDIV $\geq$	PRF $\geq$	Probability
3.4%	0%	5.31%
3.5%	20%	4.32%
3.6%	40%	3.48%
3.7%	60%	2.82%
3.8%	80%	2.28%
3.9%	100%	1.81%

Table 1: Distribution of the LDIV as calculated by RMS (Source: Standard and Poor's (2010))

In doing so, we note that there appear to be some anomalies in the data for men in the US aged 70 in 1977 and 1978 and men aged 78 in 1978 and 1979. These anomalies were also found in Cairns et al. (2014). These materially affect the estimation of the higher order age/period terms and cohort parameters in the model. We therefore remove these data cells from the analysis by setting their weights to zero in the estimation process. We also note that the data for England & Wales includes revisions as a result of the 2011 census, as discussed in Cairns et al. (2014), which have reduced the exposures to risk at higher ages in the last years of the data compared with the population data that was available to RMS in 2010. However, these changes are concentrated on individuals aged 85 and higher who do not impact the value of the Kortis bond. We therefore believe that these data issues have not materially affected the results presented in this study.

The general procedure constructs a bespoke mortality model within the class of age/period/cohort models discussed in Hunt and Blake (2015e) of the form

$$\ln(\mu_{x,t}^{(p)}) = \alpha_x^{(p)} + \sum_{i=1}^{N_p} f^{(i,p)}(x;\theta^{(i,p)})\kappa_t^{(i,p)} + \gamma_{t-x}^{(p)} \qquad (5)$$

where

- age, x, is in the range $[50, 100]$, period, t, is in the range $[1950, 2008]$, and, consequently, that year of birth, $y = t - x$, is in the range $[1850, 1958]$;

- $\alpha_x^{(p)}$ is a static function of age for population $p = \{EW, US\}$ for England & Wales and the US, respectively;

- $\kappa_t^{(i,p)}$ are N_p period functions governing the evolution of mortality with time;

- $f^{(i,p)}(x; \theta^{(i,p)})$ are parametric age functions (in the sense of having a specific functional form selected a priori) modulating the impact of the mortality dynamics over the age range, potentially with free parameters $\theta^{(i,p)}$;[9] and

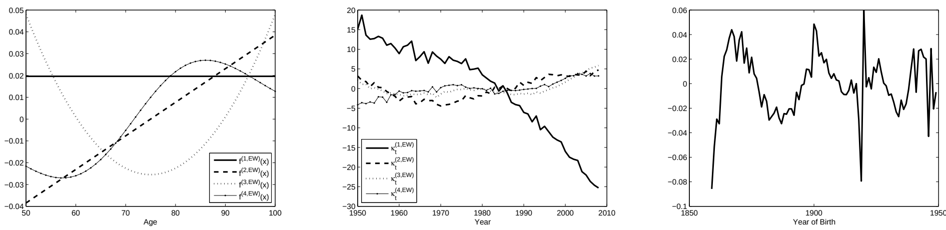
- $\gamma_y^{(p)}$ is a cohort function describing mortality effects which depend upon a cohort's year of birth and follow that cohort through life as it ages.

The general procedure sequentially adds age/period terms in order to construct a mortality model tailored to the specific features of the data, before adding a cohort term if this is statistically justified. Once the model is constructed, it is refitted so that all terms are freely estimated to provide the best fit to the data. In so doing, we obtain the age/period terms described in Table 2 and shown in Figures 3a, 3b, 4a and 4b.[10] The general procedure produces a model with four age/period terms for men in England & Wales, and six terms for men in the US. The two models also include static age functions and cohort parameters shown in Figures 3c and 4c. Further details of the age functions used in this model and tests of the goodness of fit to data are given in Appendix A.

The general procedure reproduces the age/period terms in model M7 in Cairns et al. (2009) for both England & Wales and the US, but extends these with a number of higher order age/period terms. M7 was found to give a good fit to both of these datasets in Cairns et al. (2009). However, additional age/period terms are necessary to ensure that the cohort parameters in the model are robustly estimated and are not attempting to erroneously capture unexplained or unaccounted for age/period structure. Since the cohort parameters are critical for modelling the Kortis bond, as discussed in Section

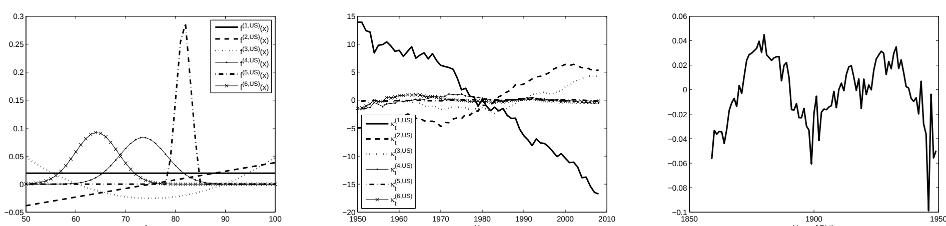
[9] For simplicity, the dependence of the age functions on $\theta^{(i,p)}$ is suppressed in notation used in this paper, although it has been allowed for when fitting the model to data.

[10] Demographic significance, as used in Table 2, is defined in Hunt and Blake (2015e) as the interpretation of the components of a model in terms of the underlying biological, medical or socio-economic causes of changes in mortality rates which generate them.



(a) Age functions (b) Period functions (c) Cohort function

Figure 3: Age, period and cohort functions for England & Wales



(a) Age functions (b) Period functions (c) Cohort function

Figure 4: Age, period and cohort functions for the US

Term	England & Wales		US	
	Description	Demographic Significance	Description	Demographic Significance
$f^{(1)}(x)\kappa_t^{(1)}$	Constant age function	General level of mortality	Constant age function	General level of mortality
$f^{(2)}(x)\kappa_t^{(2)}$	Linear age function	Slope of mortality curve	Linear age function	Slope of mortality curve
$f^{(3)}(x)\kappa_t^{(3)}$	Quadratic age function	Curvature of mortality curve	Quadratic age function	Curvature of mortality curve
$f^{(4)}(x)\kappa_t^{(4)}$	Rayleigh age function	Postponement of old age mortality	Gaussian age function	Mortality effect centred around age 73
$f^{(5)}(x)\kappa_t^{(5)}$	N/A	N/A	Linear spline age function	Mortality effect centred around age 82
$f^{(6)}(x)\kappa_t^{(6)}$	N/A	N/A	Gaussian age function	Mortality effect centred around age 64

Table 2: Terms in the models constructed using the general procedure

5, this is of vital importance.

Many mortality models are not fully identified. This means that we can find transformations of the parameters in the model which leave the fitted mortality rates unchanged.[11] To uniquely specify the parameters when fitting the model to data, we therefore impose identifiability constraints. These constraints are arbitrary, in the sense that they do not affect the fit to data, but allow us to impose our desired demographic significance on the parameters. These issues are discussed in detail in Hunt and Blake (2015c) and Hunt and Blake (2015d).

[11] For this reason, they were called "invariant transformations" in Hunt and Blake (2015c,d).

In the context of the models generated by the general procedure for England & Wales and the US, we impose the following standard identifiability constraints[12]

$$\sum_t \kappa_t^{(i)} = 0 \;\forall i \tag{6}$$

$$\sum_x |f^{(i)}(x)| = 1 \;\forall i \tag{7}$$

These identifiability constraints, respectively, allow us to:

- set a consistent level for each of the period functions, so that they represent deviations from an "average" level of mortality in the period, and
- select age functions a priori so that they have a consistent normalisation scheme. This enables us to compare the magnitudes of the period functions both with each other and between populations and also to gauge their relative importance.

However, using the results of Hunt and Blake (2015d), we observe that the following transformations of the parameters leave the fitted mortality results unchanged[13]

$$\{\hat{\alpha}_x, \hat{\kappa}_t^{(1)}, \hat{\kappa}_t^{(2)}, \hat{\kappa}_t^{(3)}, \hat{\gamma}_y\} = \{\alpha_x - a_0, \kappa_t^{(1)}, \kappa_t^{(2)}, \kappa_t^{(3)}, \gamma_y + a_0\} \tag{8}$$

$$\begin{aligned}\{\hat{\alpha}_x, \hat{\kappa}_t^{(1)}, \hat{\kappa}_t^{(2)}, \hat{\kappa}_t^{(3)}, \hat{\gamma}_y\} = \{&\alpha_x + a_1(x - \bar{x}), \kappa_t^{(1)} - a_1(t - \bar{t}),\\ &\kappa_t^{(2)}, \kappa_t^{(3)}, \gamma_y + a_1(y - \bar{y})\}\end{aligned} \tag{9}$$

$$\begin{aligned}\{\hat{\alpha}_x, \hat{\kappa}_t^{(1)}, \hat{\kappa}_t^{(2)}, \hat{\kappa}_t^{(3)}, \hat{\gamma}_y\} = \{&\alpha_x - a_2\left((x - \bar{x})^2 - \sigma_y + \sigma_t\right), \kappa^{(1)} - a_2\left((t - \bar{t})^2 - \sigma_t\right),\\ &\kappa_t^{(2)} + 2a_2(t - \bar{t}), \kappa_t^{(3)}, \gamma_y + a_2\left((y - \bar{y})^2 - \sigma_y\right)\}\end{aligned} \tag{10}$$

$$\begin{aligned}\{\hat{\alpha}_x, \hat{\kappa}_t^{(1)}, \hat{\kappa}_t^{(2)}, \hat{\kappa}_t^{(3)}, \hat{\gamma}_y\} = \{&\alpha_x + a_3\left((x - \bar{x})^3 + 3\sigma_x(x - \bar{x}) - \rho_y(x - \bar{x})\right),\\ &\kappa_t^{(1)} - a_3\left((t - \bar{t})^3 + 3\sigma_x(t - \bar{t}) - \rho_y(t - \bar{t})\right),\\ &\kappa_t^{(2)} + a_3\left((t - \bar{t})^2 - \sigma_t\right), \kappa_t^{(3)} - 3a_3(t - \bar{t}),\\ &\gamma_y + a_3\left((y - \bar{y})^3 - \rho_y(y - \bar{y})\right)\}\end{aligned} \tag{11}$$

---

[12] Dependence on population, p, has been suppressed for clarity.

[13] Constants $\bar{x}$, σ_x and ρ_x, and similarly for $\bar{t}$, $\bar{y}$, etc, are defined in Appendix A. Also note that for ease of exposition, Equations 8, 9, 10 and 11 do not incorporate the normalisation factors required on the age functions in order to ensure that $\sum_x |f^{(i)}(x)| = 1 \;\forall i$. These also need to be included in the full computation of the model.

All of these transformations involve adding a deterministic function of year of birth to the cohort parameters and adding corresponding deterministic functions of age and period to the static age function and period parameters. This leaves the fitted mortality rates unchanged due to the collinearity between age, period and cohort (i.e., $y = t - x$). The transformations, therefore, can be used to change the observed trends in the parameters, which will depend upon the identifiability constraints imposed, but not their stochastic behaviour around these trends.[14] The degrees of freedom represented by the free parameters a_0, a_1, a_2 and a_3 in these transformations need to be used to impose four identifiability constraints on the cohort parameters when fitting the model. We choose these to be

$$\sum_y n_y \gamma_y = 0 \tag{12}$$

$$\sum_y n_y \gamma_y (y - \bar{y}) = 0 \tag{13}$$

$$\sum_y n_y \gamma_y \left((y - \bar{y})^2 - \sigma_y \right) = 0 \tag{14}$$

$$\sum_y n_y \gamma_y \left((y - \bar{y})^3 - \rho_y (y - \bar{y}) \right) = 0 \tag{15}$$

where n_y is the number of observations for each cohort in the data. The justification for these constraints is that they remove polynomial trends up to cubic order in the cohort parameters at the fitting stage, so that they conform better with the demographic significance described in Hunt and Blake (2015e), i.e., that the cohort parameters should be centred around zero and not have any long term trends. However, it is important to note that the choice of these constraints is still arbitrary and, while convenient for the purposes of fitting the model to data, should not affect our projected mortality rates. To ensure this, we need to be careful when selecting the time series processes we use to project the parameters in Sections 3.2.2 and 3.3 below.

[14]This, in turn, implies that the trends observed in the fitted parameters have no meaning independent of the specific identifiability constraints imposed. Nevertheless, deviations from these trends do have independent meaning, as discussed in Section 3.2.2.

3.2 Multi-population projections of the period functions

In order to analyse the Kortis bond, it is clear that we require a mortality model that can capture the co-movement in mortality rates consistently between the different populations. Such a model needs to be able to capture the correlations between mortality rates observed in the different populations, but also give projections which do not depend on the arbitrary choices used to identify the model. A number of different multi-population mortality models have been developed to achieve this, starting with that of Carter and Lee (1992), which built upon the model of Lee and Carter (1992). This framework was later extended in Li and Lee (2005) and Hyndman et al. (2013) and tested in Li and Hardy (2011). Alternative modelling frameworks have been developed in Jarner and Kryger (2011), Dowd et al. (2011) and Cairns et al. (2011b). However, many of these models make strong assumptions about the evolution of mortality rates in the future or lead to projected mortality rates which are not independent of the identifiability constraints that have been used, as discussed in Hunt and Blake (2015b) and Section 3.2.1. To avoid these problems, we have developed an alternative approach based on cointegration, discussed in 3.2.2.

3.2.1 Coherence in projections

One issue with most of the models above is that they impose "coherence" in the sense defined in Hyndman et al. (2013) (hereafter "HBY") on the populations, namely that relative mortality rates in the different populations tend to a ratio that is constant in time, i.e.

$$\mathbb{E}\left[\frac{\mu_{x,t}^{(p_1)}}{\mu_{x,t}^{(p_2)}}\right] \rightarrow R_x \tag{16}$$

which is a function of age only. This definition of coherence means that the relative difference in mortality rates is mean reverting. This, in turn, implies that we could predict the relative divergence in mortality rates between the populations in both the short and intermediate terms as they revert to the mean from the last observed value. In the context of modelling the payoff of the Kortis bond, which depends on the divergence in mortality rates between two populations, this is a strong assumption to make.

To see this, let us take

$$R_{x,t} = \frac{\mu_{x,t}^{(EW)}}{\mu_{x,t}^{(US)}}$$

as the relative mortality rates between England & Wales and the US at age x and in year t. Figure 5 shows these relative mortality rates across time at selected ages. There appears to be a negative trend for most ages, meaning that mortality rates have diverged between England & Wales and the US over the historical data period used to estimate the model. The payoff of the Kortis bond will depend upon whether this divergence continues in the near future or whether the relative mortality rates mean-revert to some long-run level.

In the very long run, HBY convergence to a long-run mean may be a reasonable assumption and plausibly justified on biological grounds. However, there is little evidence in the historical data to suggest that there is mean reversion to a constant difference in relative mortality rates in the shorter run. An absence of mean-reversion has also been found in other related datasets, for example, Villegas and Haberman (2014) who document divergences in relative rates between sub-populations of England & Wales over a 30-year period. As discussed in Hunt and Blake (2015b), it therefore does not seem reasonable to assume HBY convergence for the short-run projections we require to analyse the Kortis bond.

In addition, the results of Hunt and Blake (2015b) show that imposing coherence often conflicts with a more important desire that the projected mortality rates do not depend upon the arbitrary identifiability constraints imposed when fitting the model to data. It is especially important that the projected mortality rates are well-identified when modelling longevity-linked securities such as the Kortis bond in order to ensure that these arbitrary choices do not have financial consequences. We therefore wish to use a mortality projection method which does not impose coherence, namely the cointegration technique developed in Hunt and Blake (2015b).

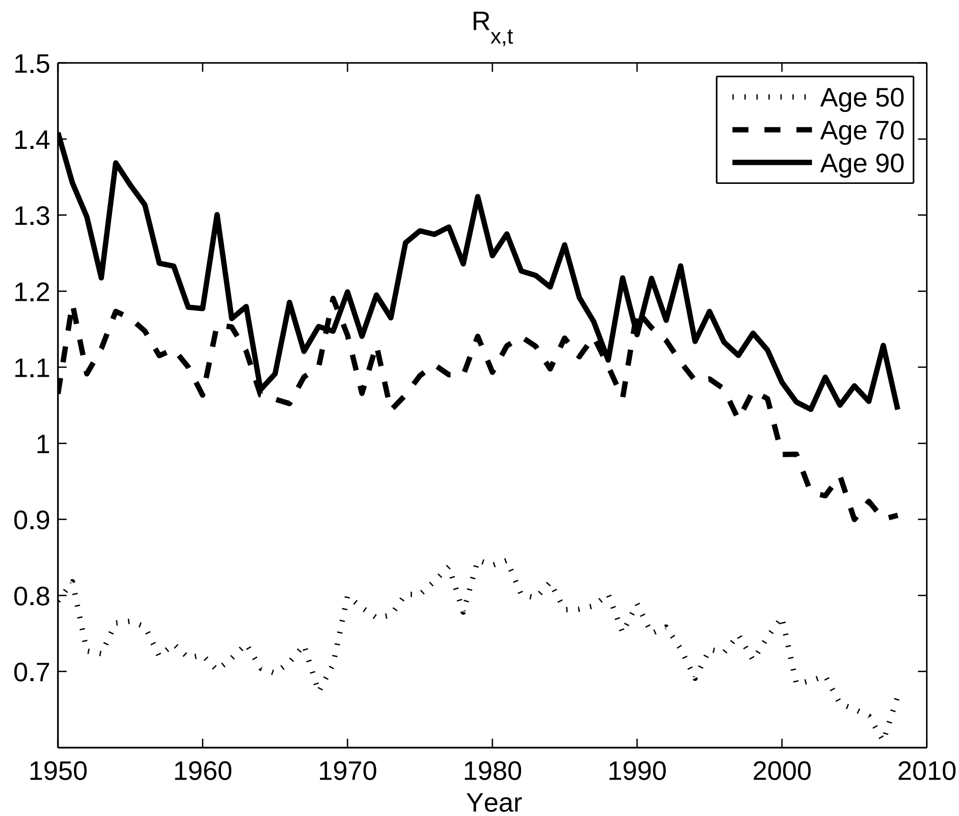


Figure 5: Ratio of mortality rates in England & Wales and the US at different ages

3.2.2 A cointegration process for the period functions

Modelling the period functions using a cointegration framework enables us to project the two populations in a manner which allows for covariation between mortality rates in the two populations, but does not impose HBY coherence. This is similar to the approach proposed in Carter and Lee (1992) (implemented in Li and Hardy (2011)) and Yang and Wang (2013)), which used the Lee and Carter (1992) model. It also has some similarity with the "gravity" model proposed in Dowd et al. (2011) as discussed in Zhou et al. (2014). However, these simpler models only have one age/period term for each population and therefore our approach needs to make a number of modifications.

First, we note that we chose our age/period terms using the general procedure to ensure they have distinct demographic significance. It is therefore appropriate to project them in a manner which retains this distinctiveness in order to achieve projections with greater biological reasonableness.[15] This means that we do not look for cointegrating relationships between all ten age/period terms in the models for England & Wales and the US if this is not appropriate.

Second, the general procedure resulted in the first three age/period terms in each population having the same form and demographic significance, as shown in Table 2. Therefore, it is reasonable to allow for the possibility of cointegration between these terms and so we pair them together. However, there is less justification for pairing the other population-specific terms in the model. Accordingly, we choose to look for cointegration only between the first three pairs of terms.

Forming these three pairs of period functions into vectors

$$\boldsymbol{\kappa}_t^{(i)} = \begin{pmatrix} \kappa_t^{(i,EW)} \\ \kappa_t^{(i,US)} \end{pmatrix}$$

we build on the results of Hunt and Blake (2015b) and use cointegrating time series to model the period functions. To illustrate this for $\boldsymbol{\kappa}_t^{(1)}$, we use a time series of the form

$$\Delta\boldsymbol{\kappa}_t^{(1)} = \nu^{(1)} \begin{pmatrix} 1 \\ t \\ t^2 \end{pmatrix} + \alpha^{(1)} \left(\beta^{(1)\top} \boldsymbol{\kappa}_{t-1}^{(1)} + \beta_0^{(1)} t^3 \right) + \boldsymbol{\epsilon}_t^{(1)} \tag{17}$$

where $\nu^{(1)}$ is a matrix of regression coefficients associated with the unconstrained polynomial trends up to quadratic order ($X_t = \begin{pmatrix} 1 & t & t^2 \end{pmatrix}^\top$), $\alpha^{(1)}$ and $\beta^{(1)}$ are $2 \times r^{(1)}$ matrices with the interpretation that the rows of $\beta^{(1)\top}\boldsymbol{\kappa}_{t-1}^{(1)}$ represent the $r^{(1)}$ stationary cointegrating relationships, $\beta_0^{(1)}$ is the coefficient of the constrained cubic trend and $\boldsymbol{\epsilon}_t^{(1)}$ is a vector of time series innovations.

[15] Introduced in Cairns et al. (2006b) and defined as "*a method of reasoning used to establish a causal association (or relationship) between two factors that is consistent with existing medical knowledge*"

This means that we need to allow for unconstrained trends up to quadratic order and constrained cubic trends in $\boldsymbol{\kappa}_t^{(1)}$ in order for the projected mortality rates to be well-identified (i.e., that they do not depend upon the arbitrary identifiability constraints used). A similar analysis leads to the conclusion that we need to allow for unconstrained trends up to linear order and constrained quadratic trends in $\boldsymbol{\kappa}_t^{(2)}$, and an unconstrained constant and constrained linear trends in $\boldsymbol{\kappa}_t^{(3)}$. See Hunt and Blake (2015b) for a more detailed discussion of this issue.

In order to find the number of cointegrating relationships (the sizes of the matrices $\alpha^{(i)}$ and $\beta^{(i)}$ for each pair of period functions), we calculate test statistics for the fitted time series, as described in Johansen (1991). Many authors, such as in Darkiewicz and Hoedemakers (2004), Lazar and Denuit (2009), Salhi and Loisel (2009), Gaille and Sherris (2011) and Arnold-Gaille and Sherris (2013) used statistical tests based on the asymptotic distributions of these test statistics to find the number of cointegrating relationships. Standard critical statistics for these tests have been calculated and are freely available. However, the values of these test statistics are highly dependent on the nature of the deterministic trends allowed for in the model. Trends beyond linear order are not generally present in the situations in economics for which cointegration was developed[16] and so standard critical statistics are not available for the circumstances we face. In addition, we need to be aware that tests based on asymptotic analyses may have low power when used with limited data and a large number of free parameters. As we are fitting our model to approximately 60 years of data, it is likely that the standard test statistics will not be valid for use with our model.

Therefore, we test for the number of cointegrating relationships using bootstrapping to generate our own critical statistics. For each pair of period functions, we compare the calculated test statistic (found via the Johansen procedure as described in Juselius (2006)) against the 5^{th} percentile of the empirical distribution found from 5,000 Monte Carlo simulations under the null hypothesis that there are no stationary cointegrating relationships, i.e.,

[16]Juselius (2006, p. 100) assumed the higher order trends are the result of model misspecification and suggests that *"although quadratic trends may sometimes improve the fit ... it seems preferable to find out what caused this approximate quadratic growth and, if possible include more appropriate information in the model"*.

$r^{(i)} = 0$. If the null hypothesis is rejected, we then repeat the bootstrapping procedure under the null hypothesis that $r^{(i)} = 1$, etc. This procedure sets a high threshold for finding cointegrating relationships, which means we may find fewer of them than a procedure based on asymptotic critical values.

In addition, to allow for parameter uncertainty, we use the residual bootstrapping procedure of Koissi et al. (2006) to generate resampled death counts for each population, and use these to refit the mortality models in Section 3.1. To each re-estimated set of period functions, we then refit the cointegrating model and re-test for the number of stationary cointegrating relationships. This means that we allow both the period functions and the number of cointegrating relationships to be subject to parameter uncertainty in our analysis.

It is worth noting that, if we find no cointegrating relationships for $\boldsymbol{\kappa}_t^{(i)}$, then $\alpha^{(i)}$ and $\beta^{(i)}$ are null matrices and Equation 17 simplifies to

$$\Delta \boldsymbol{\kappa}_t^{(i)} = \nu^{(i)} X_t^{(i)} + \boldsymbol{\epsilon}_t^{(i)} \tag{18}$$

where $X_t^{(i)}$ are the unconstrained deterministic trends. This is equivalent to using a well-identified random walk with drift for the period functions, which is consistent with the approach used in Cairns et al. (2011a) and Haberman and Renshaw (2011).

Using this procedure on the fitted parameters, we find one cointegrating relationship between $\kappa_t^{(1,EW)}$ and $\kappa_t^{(1,US)}$ 100% of the time when allowing for parameter uncertainty. This means that we always find a cointegrating relationship between the dominant period functions for England & Wales and the US, which control the level of mortality rates across ages. We find one cointegrating relationship between $\kappa_t^{(2,EW)}$ and $\kappa_t^{(2,US)}$, which determine the slopes of the mortality curves in both countries, about 7% of the time. Finally, 75% of the time we find one cointegrating relationship between $\kappa_t^{(3,EW)}$ and $\kappa_t^{(3,US)}$, which controls the curvature of the mortality curves.

The presence of cointegrating relationships between the period functions acts to reduce the modelled basis risk between the two populations which arises from the particular age/period term in the models. For instance, consider $\boldsymbol{\kappa}_t^{(1)}$, for which we always find a cointegrating relationship. In order to be well-identified, the cointegration techniques we use allow for different trends in the level of mortality in the two populations. However, the presence of a cointegrating relationship in $\boldsymbol{\kappa}_t^{(1)}$ ensures that deviations from the trend in either population are correlated, i.e., if we project faster than anticipated decreases in $\kappa_t^{(1,EW)}$, then we are more likely to project faster than anticipated decreases in $\kappa_t^{(1,US)}$. This reduces the potential for unexpected deviations from the difference in the levels of the mortality curves of the two populations, and hence the basis risk arising from the "level" factor in the mortality model. Similarly, we usually find a cointegrating relationship for $\boldsymbol{\kappa}_t^{(3)}$, indicating that there is relatively little basis risk arising from unexpected changes in the curvature of the mortality curve.

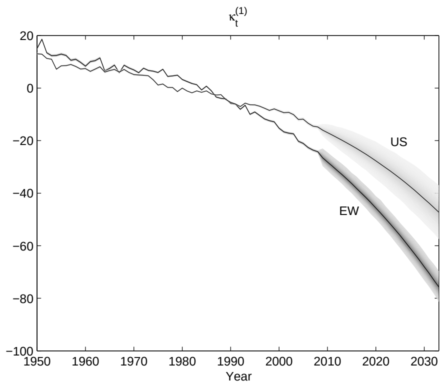
In contrast, we rarely find a cointegrating relationship for $\boldsymbol{\kappa}_t^{(2)}$. This means that our projections allow for considerable basis risk arising from changes in the slope of the mortality curve of either country. This basis risk is important when comparing populations at different ages, such as life assurance policyholders and annuitants. However, we note that the LDIV has been structured to provide a good hedge for basis risk arising from just this source. Therefore, we find that the Kortis bond is highly effective at hedging the factor we believe to be the primary cause of basis risk between the two populations.

For the other period functions in the models, we project them using a standard random walk with drift

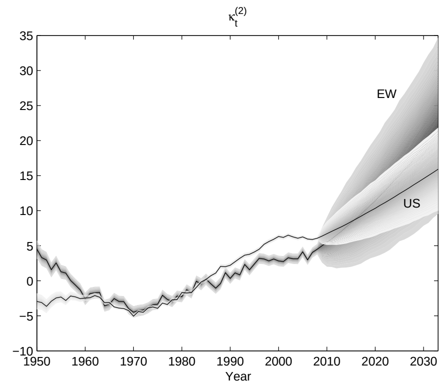
$$\Delta\kappa_t^{(i,p)} = \nu^{(i,p)} + \epsilon_t^{(i,p)},\ p = EW \text{ and } i = 4,\ p = US \text{ and } i = 4, 5, 6 \qquad (19)$$

To allow for covariation in the different populations, we let the innovations, $\epsilon_t^{(i,p)}$, including those from the cointegrating relationships, be contemporaneously correlated.

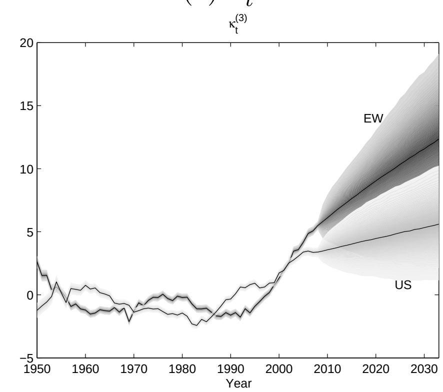
Since the fitted residuals from the processes used to project the period functions appear to be slightly leptokurtic, we do not assume that future innovations are normally distributed. Instead, we generate future innovations by bootstrapping the residuals from fitting the cointegrating relationships, taking care to preserve the observed contemporaneous correlation structure across different period functions. We do this in order to preserve the consistency between our observations of the past and our projections of the future, which is fundamental to the extrapolative approach to modelling mortality



(a) $\kappa_t^{(1)}$



(b) $\kappa_t^{(2)}$



(c) $\kappa_t^{(3)}$

Figure 6: Projections of the cointegrating period functions

(as discussed in Hunt and Blake (2015b)). This has an important impact on our analysis of the payoff of the Kortis bond as this depends upon the tail of the distribution of projected mortality rates and so is sensitive to any normality in the projected period functions. This is discussed in Section 5.

We then project the cointegrated time series into the future using the above time series processes and bootstrapped residuals. Projections of the period functions using the cointegrating model are shown in Figure 6. As discussed in Hunt and Blake (2015b), it is important to realise that the cointegrated time series can still appear to diverge from each other. This is because the cointegrating relationships allow for different deterministic trends in the series in order to be well-identified. These deterministic trends are set via the identifiability constraints imposed, and therefore our projected period functions should not depend upon them. Hence different trends can lead to apparent divergence in the parameters which cannot be constrained by the well-identified cointegrating relationships. However, cointegration means that the variation around these trends is correlated between the two populations, e.g., faster than trend improvements in mortality in the US are more likely to occur at times of faster than trend improvements in mortality in England & Wales and vice versa. Hence, the presence of the cointegrating relationships act to reduce the projected basis risk between the two populations and this is highly significant for the modelling of the Kortis bond.

3.3 Projecting the cohort parameters using a Bayesian framework

As a result of how Index(2016,EW) and Index(2016,US) are defined in Equation 2, the value of the LDIV in 2016 (and therefore the payoff of the Kortis bond) depends on some cohort parameters that are estimated from the historical data and others that are projected. The LDIV is therefore very sensitive to any discontinuity between historical and projected cohort parameters, and this is likely to occur using traditional methods of projecting the cohort parameters.

In Hunt and Blake (2015a), we developed a new technique for projecting cohort parameters within a Bayesian framework. This used an assumed data generating process for the ultimate cohort parameters (which can only be known once the cohort is extinct). This, in turn, allowed for the fact that the estimates of the more recent cohort parameters were based on fewer observations and had fewer deaths per observation. This means that the estimates for these more recent cohort parameters are highly uncertain. We then derived expressions for the means and variances of the ultimate cohort parameters given our incomplete observations to date and assuming an underlying mean-reverting process generating the ultimate cohort parameters. In this way, discontinuities between the historical and projected parameter estimates were avoided, with the former having a suitable allowance for their uncertainty. It is natural to use this framework again when modelling the Kortis bond.

We first assume that the cohort parameters in England & Wales and the US are independent. As the cohort parameters represent lifelong mortality effects which are specific to different years of birth, it is not unreasonable to assume that these are independent for men living on opposite sides of the Atlantic Ocean. If we test for correlation between the two sets of estimated cohort parameters, we find empirically low correlations, especially when the outliers for years of birth 1946/47 are removed.[17] In practice, the assumption of independence simplifies the use of the Bayesian framework of Hunt and Blake (2015a) considerably. However, this assumption may not be reasonable if we were to try to project two more closely related populations, such as men and women in England & Wales.

In Hunt and Blake (2015a), we assume

$$\gamma_y^{(p)} | \mathcal{F}_{t,y}, \beta^{(p)}, \rho_{(p)}, \sigma^2_{(p)} \sim N\left(M^{(p)}(y,t), V^{(p)}(y,t)\right) \tag{20}$$

with

$$M^{(p)}(y,t) = \sum_{s=0}^{\infty}\left[\prod_{r=0}^{s-1}(1-D^{(p)}_{t-y+r})\right]\rho^s_{(p)}\left[\underline{\gamma}^{(p)}_{y-s}(t) + (1-D^{(p)}_{t-y+s})\beta^{(p)}(X^{(\gamma)}_{y-s} - \rho_{(p)}X^{(\gamma)}_{y-s+1})\right] \tag{21}$$

$$V^{(p)}(y,t) = \sum_{s=0}^{\infty}\left[\prod_{r=0}^{s-1}(1-D^{(p)}_{t-y+r})^2\right](1-D^{(p)}_{t-y+s})\rho^{2s}_{(p)}\sigma^2_{(p)} \tag{22}$$

and where

- $D_x^{(p)}$ is the proportion of a cohort in population p assumed to still be alive at age x (assumed to be independent of year of birth);
- $\rho_{(p)}$ and $\sigma^2_{(p)}$ are the autocorrelation and variance, respectively, of the AR(1) process assumed to be driving the evolution of the cohort parameters in population p;

[17] We believe these outliers are caused by data artefacts related to the uneven pattern of births immediately after the Second World War rather than genuine mortality effects, see Richards (2008) and Cairns et al. (2014). We therefore use indicator variables to ensure these outliers do not impact our projected cohort parameters.

- $X_y^{(\gamma)}$ is a set of deterministic regressors introduced in order to ensure that our projections are well-identified, i.e., based on the analysis of Hunt and Blake (2015d) and Section 3.1, we need to use an AR(1) process around a cubic drift, so that $X_y^{(\gamma)}$ contains polynomial trends up to cubic order;

- $\beta^{(p)}$ is a set of regression coefficients for the cohort parameters in population p with respect to the deterministic trends $X_y^{(\gamma)}$;[18]

- $\underline{\gamma}_y^{(p)}(t)$ are found from the fitted cohort parameters at time t, $\overline{\gamma}_y^{(p)}(t)$, using $\underline{\gamma}_y(t) = D_{t-y}\overline{\gamma}_y(t)$; and

- $M^{(p)}(y,t)$ and $V^{(p)}(y,t)$ are the mean and variance, respectively, of the ultimate cohort parameter for year y based on information received at time t (i.e., $\mathcal{F}_{t,y}$).

In practice, this framework can be used to generate stochastic ultimate cohort parameters by using

$$\begin{aligned} \gamma_y^{(p)} &= \underline{\gamma}_y^{(p)}(t) + (1 - D_{t-y}^{(p)}) \left(\beta^{(p)} X_y + \rho_{(p)}(\gamma_{y-1}^{(p)} - \beta^{(p)} X_{y-1}) \right) + \epsilon_y^{(p)} \qquad (23) \\ \epsilon_y^{(p)} &\sim N\left(0, (1 - D_{t-y}^{(p)})\sigma_{(p)}^2\right) \end{aligned}$$

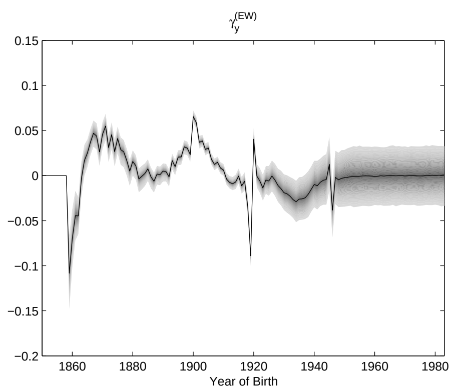
with $D_{t-y}^{(p)} = 0$ for future years of birth.

The cohort parameters estimated using this framework (and allowing for parameter uncertainty) are shown in Figure 7. It is important to note the uncertainty that is present in the cohort parameters used in the calculation of LDIV(2016), i.e., years of birth 1923 to 1941 for England & Wales and 1943 to 1961 for the US. This would be significantly underestimated using traditional methods of projecting the cohort parameters.

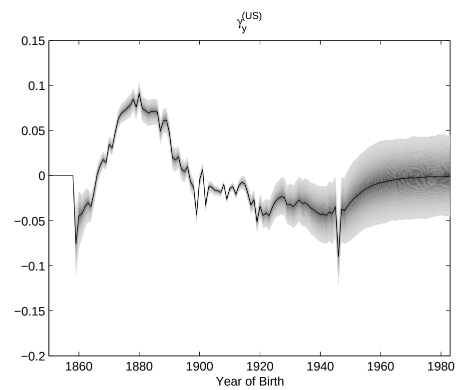
4 Projecting the LDIV

With the mortality and projection models discussed in Section 3.1, we are now able to project mortality rates for England & Wales and the US allowing

[18]These depend upon our identifiability constraints. In practice, we use Equations 8, 9, 10 and 11 to impose a new set of identifiability constraints, such that $\beta^{(p)} = 0$, to simplify matters considerably. This new constraint overrides the old constraints in Equations 12, 13, 14 and 15 that were imposed when fitting the model.



(a) $\gamma_y^{(EW)}$



(b) $\gamma_y^{(US)}$

Figure 7: Projections of the cohort parameters

for any dependence that we find between them. We perform 5,000 Monte Carlo simulations to project mortality rates 25 years into the future (i.e., to 2033) for both populations and calculate the LDIV in each case. The fan chart of the 98% confidence interval for the projected LDIV is shown in Figure 8.

Figure 8 also shows the observed values of the LDIV for 2009 and 2010 as crosses. Our projections show these observations lie close to the centre of the 98% confidence interval for those years, implying that the model gives reliable projections of the LDIV in the very short term. Figure 8 also shows that it is likely that the LDIV reached a maximum of approximately 2.2% in 2011 and will decrease in value from that point. It is therefore unlikely that the principal will be reduced when the LDIV is calculated for the purposes of determining the PRF in 2016. However, our projections show that the projected LDIV is highly uncertain. The distribution of projected values of the LDIV in 2016 is shown in Figure 9. We find that the median LDIV value is 1.4% with a distribution that is roughly symmetric (skewness of just −0.02) around this value with a standard deviation of 1.2%.

Table 3 shows the probabilities that we find for the PRF exceeding different thresholds, along with those calculated by RMS. For comparison, the median annual default probability on B-rated bonds is approximately 3.6%

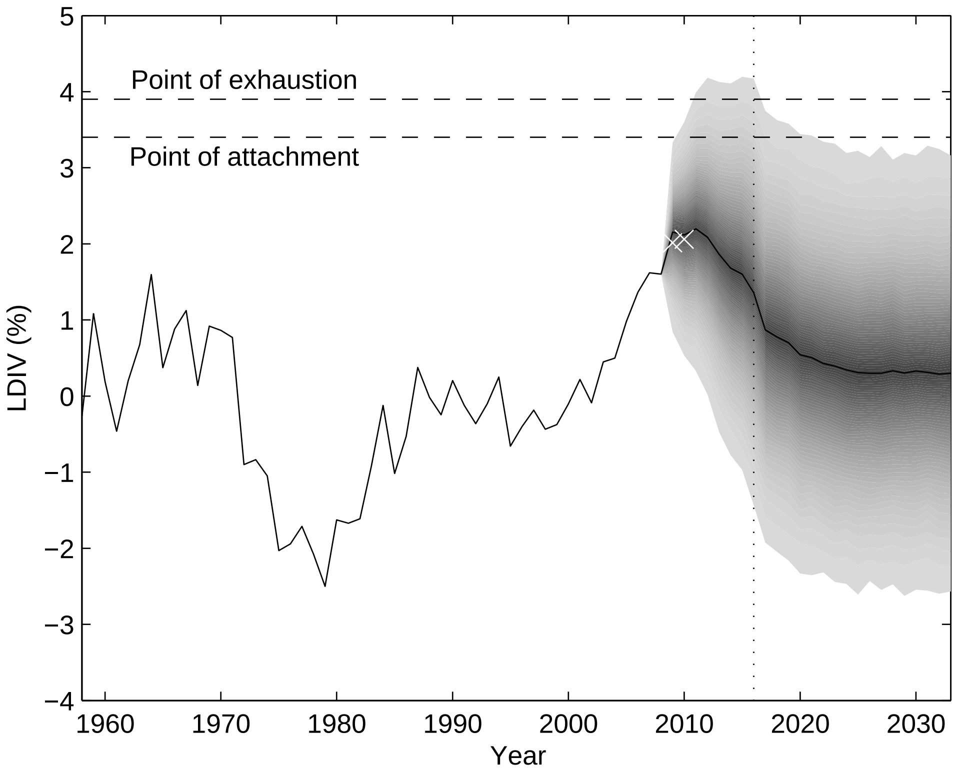


Figure 8: Fan chart of the projected LDIV showing the 98% confidence interval

p.a..[19] Our estimated probabilities are broadly consistent with, but slightly lower than those calculated by RMS. Further, the conditional expected loss we calculate is broadly consistent with that from the RMS analysis, despite a slightly lower chance that the principal is reduced. This is due to the upper bound that the point of exhaustion places on potential losses from the bond. However, without further information on the projection model used by RMS, it is impossible to draw any firm conclusions concerning the difference between our results and theirs.

[19]Source: S&P 2011 Annual Global Corporate Default Study and Rating Transitions, Table 4, `http://www.standardandpoors.com/ratings/articles/en/us/?articleType=HTML&assetID=12453308`

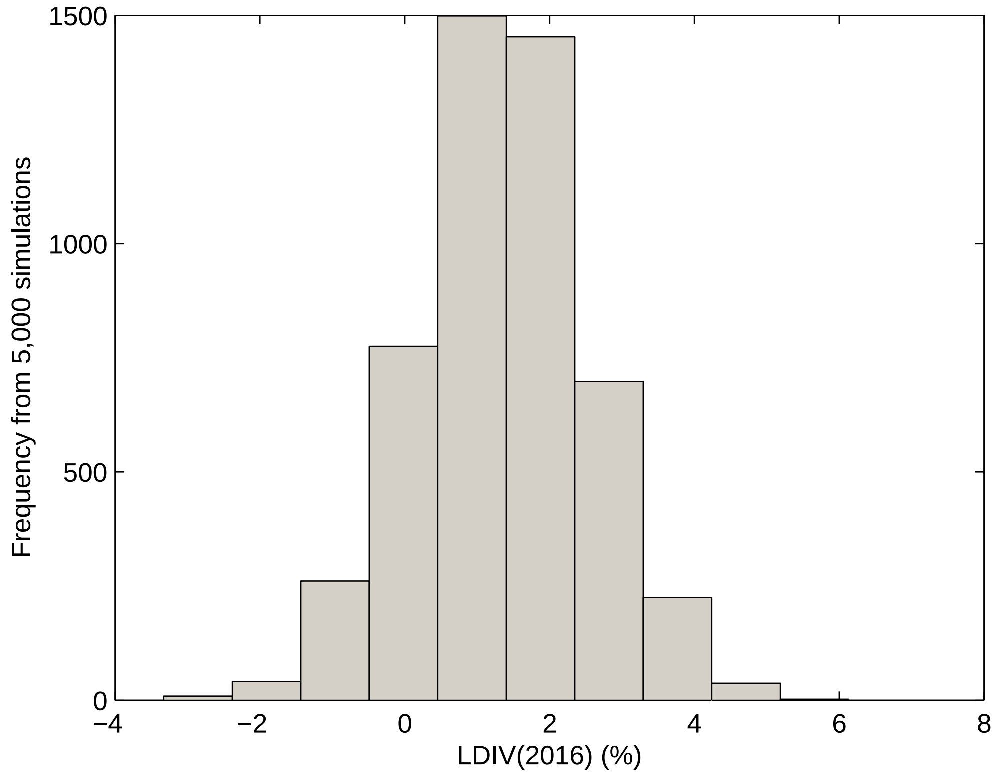


Figure 9: Distribution of the LDIV in 2016

Note that we do not attempt to put a value on the Kortis bond or estimate a suitable margin above LIBOR for the coupon payments. To do so would require having a "market price of risk"[20] to transform the "real-world" probability distribution estimated here into a "risk neutral" distribution from which discounted expected values could be obtained. Studies which use this approach include Lin and Cox (2005), Cairns et al. (2006a), Denuit et al. (2007), Plat (2009) and Chen and Cox (2009). However, as there is currently no actively traded market in longevity-linked securities or derivatives, such market prices of risk are not available.

[20]Or equivalently, a Sharpe ratio.

PRF $\geq$	Estimated Probability	
	Our results	RMS results
0%	4.34%	5.31%
20%	3.58%	4.32%
40%	3.04%	3.48%
60%	2.48%	2.82%
80%	2.10%	2.28%
100%	1.76%	1.81%
Expected loss	2.82%	3.27%
Conditional expected loss	65%	62%

Table 3: Selected quantiles of the distribution of the PRF

5 Analysis of the projected LDIV

The payoff on the Kortis bond is a complicated function of mortality rates for different populations at different times and at different ages. The appropriate analysis of the payoff, therefore, requires a sophisticated mortality model, such as the ones used by RMS or derived in this study, which is motivated by the need to understand the dynamics of mortality rates at the specific ages of interest for the bond. We analyse the main factors determining the payoff on the Kortis bond in Section 5.1 and investigate the main sources of risk to that payoff in Section 5.2.

5.1 Age/period/cohort analysis of the Kortis bond

We begin by considering the age/period/cohort models used to project mortality rates in England & Wales and the US. Since our population mortality models are of the form

$$\ln(\mu_{x,t}^{(EW)}) = \alpha_x^{(EW)} + \sum_{i=1}^{4} f^{(i,EW)}(x)\kappa_t^{(i,EW)} + \gamma_{t-x}^{(EW)} \tag{24}$$

$$\ln(\mu_{x,t}^{(US)}) = \alpha_x^{(US)} + \sum_{i=1}^{6} f^{(i,US)}(x)\kappa_t^{(i,US)} + \gamma_{t-x}^{(US)} \tag{25}$$

we have from Equation 1 that

$$\text{Improvement}_8^{(p)}(x,t) = 1 - \exp\left[\frac{1}{8}\left(\sum_{i=1}^{N_p} f^{(i,p)}(x)(\kappa_t^{(i,p)} - \kappa_{t-8}^{(i,p)}) + \gamma_{t-x}^{(p)} - \gamma_{t-8-x}^{(p)}\right)\right] \tag{26}$$

which gives the improvement rate observed in each population over the eight-year period at different ages. Equation 2 calculates the index value in each population as the average of the improvement rates across the different ages using the arithmetic mean. For algebraic simplicity, we approximate this by the geometric mean to give[21]

$$\begin{aligned}\text{Index}(t,y) &\approx 1 - \exp\left[\frac{1}{8}\sum_{i=1}^{N_p} \bar{f}^{(i,p)}(\kappa_t^{(i,p)} - \kappa_{t-8}^{(i,p)}) + \frac{1}{8}\left(\bar{\gamma}_t^{(p)} - \bar{\gamma}_{t-8}^{(p)}\right)\right] \\ &\approx -\left[\frac{1}{8}\sum_{i=1}^{N_p} \bar{f}^{(i,p)}(\kappa_t^{(i,p)} - \kappa_{t-8}^{(i,p)}) + \frac{1}{8}\left(\bar{\gamma}_t^{(p)} - \bar{\gamma}_{t-8}^{(p)}\right)\right]\end{aligned} \tag{27}$$

where a bar denotes the arithmetic mean over the relevant ages.[22] Finally, substituting this into Equation 3 and expanding to first order we find that[23]

$$\begin{aligned}\text{LDIV}(t) \approx \frac{1}{8}\sum_{i=1}^{6}&\left(\bar{f}^{(i,US)}(\kappa_t^{(i,US)} - \kappa_{t-8}^{(i,US)}) - \bar{f}^{(i,EW)}(\kappa_t^{(i,EW)} - \kappa_{t-8}^{(i,EW)})\right) \\ &+ \frac{1}{8}\left(\bar{\gamma}_t^{(US)} - \bar{\gamma}_{t-8}^{(US)} - \bar{\gamma}_t^{(EW)} + \bar{\gamma}_{t-8}^{(EW)}\right)\end{aligned} \tag{28}$$

In the long run, the cohort parameters in each population will mean-revert around zero due to the identifiability constraints imposed, and therefore will not contribute to the long-run evolution of the LDIV. However, they will be important in the short run as the historical cohort parameters pass through the relevant ages for calculating the LDIV. Indeed, the distinctive peak of the LDIV in Figure 8 between 2009 and 2012 is mostly caused by the decrease in the value of the cohort parameters observed in England & Wales in the 1930s

[21] We have tested the accuracy of this approximation and in practice it makes no significant difference to the results obtained.

[22] That is, $\bar{f}^{(i,p)} = \frac{1}{1+x_{2,p}-x_{1,p}}\sum_{x=x_{1,p}}^{x_{2,p}} f^{(i,p)}(x)$ and $\bar{\gamma}_t^{(p)} = \frac{1}{1+x_{2,p}-x_{1,p}}\sum_{x=x_{1,p}}^{x_{2,p}} \gamma_{t-x}^{(p)}$.

[23] For clarity, we have implicitly set $f^{(5,EW)}(x) = f^{(6,EW)}(x) = 0\ \forall x$ in order to have the same number of age/period terms for each population.

identified in Figure 3c, which corresponds to the "golden cohort" identified in Willets (1999, 2004) and discussed in Murphy (2009). This illustrates why the models used for England & Wales have to have well-specified cohort parameters, as provided by the use of the general procedure, and why the projection of these parameters must allow fully for their uncertainty, such as in the Bayesian approach of Hunt and Blake (2015a).

For the period functions generated by the random walk process in Equation 19, i.e., the population-specific period functions, this can be solved to give

$$\frac{1}{8}(\kappa_t^{(i,p)} - \kappa_{t-8}^{(i,p)}) = \nu^{(i,p)} + \frac{1}{8}\sum_{s=t-7}^{t} \epsilon_s^{(i)} \tag{29}$$

Therefore, $\frac{1}{8}(\kappa_t^{(i,p)} - \kappa_{t-8}^{(i,p)})$ has constant mean and standard deviation.

From Equation 18, we also see that when we find no cointegrating relationships, Equation 17 simplifies to a standard random walk with drift. This will be the case in the majority of simulations for $\boldsymbol{\kappa}_t^{(2)}$. In this case, $\frac{1}{8}(\kappa_t^{(i,p)} - \kappa_{t-8}^{(i,p)})$ will also have constant standard deviation. However, since $\boldsymbol{\kappa}_t^{(2)}$ requires a linear drift term, we find $\frac{1}{8}\mathbb{E}(\kappa_t^{(2,p)} - \kappa_{t-8}^{(2,p)}) = \frac{1}{8}\nu^{(2,p)}\sum_{s=t-7}^{t} X_s^{(2)}$, which also changes linearly with time.

The case where we find one cointegrating relationship between the pair of period parameters is slightly more complicated. From Equation 17, we have (where j is the polynomial order of the constrained trends for $\boldsymbol{\kappa}_t^{(i)}$)

$$\begin{aligned}
\boldsymbol{\kappa}_t^{(i)} &= \nu^{(i)} X_t^{(i)} \boldsymbol{\kappa}_{t-1}^{(i)} + \alpha^{(i)} \left(\beta^{(i)\top}\boldsymbol{\kappa}_{t-1}^{(i)} + \beta_0^{(1)} t^j\right) + \boldsymbol{\epsilon}_t^{(i)} \\
&= \tilde{\nu}^{(i)} \tilde{X}_t^{(i)} + \left(\mathtt{I} + \alpha^{(i)}\beta^{(i)\top}\right)\boldsymbol{\kappa}_{t-1}^{(i)} + \boldsymbol{\epsilon}_t^{(i)}
\end{aligned}$$

where we have combined the polynomial terms inside and outside of the cointegrating relationship into a single term[24] and $\mathtt{I}$ is the identity matrix. For example, for $\boldsymbol{\kappa}_t^{(1)}$, we require $X_t^{(1)} = \begin{pmatrix}1, & t, & t^2\end{pmatrix}^\top$ in Equation 17, and a term in t^3 (i.e., $j = 3$) within the cointegrating relationship. These have been

[24]To do this, we set $\tilde{\nu}^{(i)} = \begin{pmatrix}\nu^{(i)}, & \alpha^{(i)}\beta_0^{(i)}\end{pmatrix}$.

combined into a single term in $\tilde{X}^{(1)}_t = \left(1, \quad t, \quad t^2, \quad t^3\right)^\top$.

Using this rearrangement, we find

$$\begin{aligned}
\boldsymbol{\kappa}^{(i)}_t &= \tilde{\nu}^{(i)} \tilde{X}^{(i)}_t + \left(\mathtt{I} + \alpha^{(i)} \beta^{(i)\top}\right) \boldsymbol{\kappa}^{(i)}_{t-1} + \boldsymbol{\epsilon}^{(i)}_t \\
&= \sum_{s=t-7}^{t} \left(\mathtt{I} + \alpha^{(i)} \beta^{(i)\top}\right)^{t-s} \tilde{\nu}^{(i)} \tilde{X}^{(i)}_s + \sum_{s=t-7}^{t} \left(\mathtt{I} + \alpha^{(i)} \beta^{(i)\top}\right)^{t-s} \epsilon^{(i)}_s \\
&\quad + \left(\mathtt{I} + \alpha^{(i)} \beta^{(i)\top}\right)^{8} \boldsymbol{\kappa}^{(i)}_{t-8} \qquad (30)
\end{aligned}$$

$$\begin{aligned}
\frac{1}{8}\mathbb{E}\left(\boldsymbol{\kappa}^{(i)}_t - \boldsymbol{\kappa}^{(i)}_{t-8}\right) &= \frac{1}{8} \sum_{s=t-7}^{t} \left(\mathtt{I} + \alpha^{(i)} \beta^{(i)\top}\right)^{t-s} \tilde{\nu}^{(i)} \tilde{X}^{(i)}_s \\
&\quad + \frac{1}{8} \left((\mathtt{I} + \alpha^{(i)} \beta^{(i)\top})^{8} - \mathtt{I}\right) \mathbb{E}\boldsymbol{\kappa}^{(i)}_{t-8} \qquad (31)
\end{aligned}$$

In order for the cointegrating relationships to be stationary, we require the determinant $det(\alpha^{(i)}\beta^{(i)\top}) < 1$ and so the second term in Equation 31 will be relatively small. Therefore, we see that the term $\frac{1}{8}(\kappa^{(i,p)}_t - \kappa^{(i,p)}_{t-8})$ is driven by the higher-order deterministic terms in $\tilde{X}^{(i)}_t$ and, in the long run, will come to dominate the evolution of the LDIV. Accordingly, we anticipate that the evolution of the first and second period functions will come to dominate the level of the LDIV in the long run, albeit with a significant contribution from the cohort parameters in the short run before they mean-revert.

Turning to the variability of the LDIV, we have assumed that the cohort parameters are stationary and, accordingly, $\frac{1}{8^2}\mathbb{V}ar\left(\bar{\gamma}^{(US)}_t - \bar{\gamma}^{(US)}_{t-8} - \bar{\gamma}^{(EW)}_t + \bar{\gamma}^{(EW)}_{t-8}\right)$ will be constant. We also see from Equation 29 that the variability of those period parameters projected using a random walk or which lack a cointegrating relationship will be $\frac{1}{8^2}\mathbb{V}ar\left(\kappa^{(i,p)}_t - \kappa^{(i,p)}_{t-8}\right) = \frac{1}{8}\mathbb{V}ar(\epsilon^{(i,p)}_t)$, which is also constant. The case when there is one cointegrating relationship is, again, slightly more complicated. From Equation 30, we find

$$\begin{aligned}
\frac{1}{8^2}\mathbb{V}ar\left(\boldsymbol{\kappa}^{(i)}_t - \boldsymbol{\kappa}^{(i)}_{t-8}\right) &= \frac{1}{8}\left(\mathtt{I} + \alpha^{(i)} \beta^{(i)\top}\right)^\top \mathbb{V}ar(\boldsymbol{\epsilon}^{(i)}_t) \left(\mathtt{I} + \alpha^{(i)} \beta^{(i)\top}\right) \\
&\quad + \frac{1}{8^2}\left((\mathtt{I} + \alpha^{(i)} \beta^{(i)\top})^8 - \mathtt{I}\right)^\top \mathbb{V}ar\left(\boldsymbol{\kappa}^{(i)}_{t-8}\right) \left((\mathtt{I} + \alpha^{(i)} \beta^{(i)\top})^8 - \mathtt{I}\right)
\end{aligned}$$

The first of these terms is constant, whilst the second increases linearly with time but is very small in practice as $det(\alpha^{(i)}\beta^{(i)\top}) < 1$. Combining these sources of risk, we would expect the variability of the LDIV to remain roughly constant in time - which is consistent with Figure 8. In addition, we would also expect the period parameters to explain the majority of the uncertainty in our projected mortality rates, as they explained the majority of the structure present in the historical data (consistent with the demographic significance in Hunt and Blake (2015e)).

5.2 Decomposition of sources of risk

To decompose the risks that we model in the LDIV, we use the tripartite classification of risks proposed in Cairns (2000):

- Process risk - the stochastic uncertainty due to the projection of the time series processes driving the dynamics of the model (i.e., due to the randomness in the projections of the period and cohort parameters).
- Parameter uncertainty - the uncertainty due to the fact that the parameters in the model and the time series processes are not known with certainty, but are estimates based on finite data. We allow for this using the residual bootstrapping procedure of Koissi et al. (2006), which generates multiple realisations of the different fitted parameters that can then be projected.
- Model risk - the uncertainty caused by our model being an approximation to the true underlying processes driving changing mortality rates. We do not allow for this directly, but try to illustrate the potential impact of model risk, as described below.

We also allow for idiosyncratic risk in the projected mortality rates, i.e., the risk associated with the LDIV being calculated from the future death counts in finite (albeit large) populations.

To perform this risk decomposition, we allow for each risk in sequence, whilst assuming all other factors are deterministic. For example, to allow for process risk in the projected period parameters, we project them stochastically, but assume that all historical parameters are known with certainty, the cohort parameters are projected deterministically and the observed death counts are assumed to be equal to their expectations under the Poisson assumption.[25] Finally, we allow for all of the different sources of risk in the model simultaneously to give the results presented in Section 4. As can be seen in Figure 10, although the risks associated with parameter uncertainty and idiosyncratic risk are relatively small, the fact that the number of cointegrating relationships we found for both $\boldsymbol{\kappa}_t^{(2)}$ and $\boldsymbol{\kappa}_t^{(3)}$ in Section 3.2.2 was sensitive to parameter uncertainty means that it is still important that our modelling includes these risk sources.

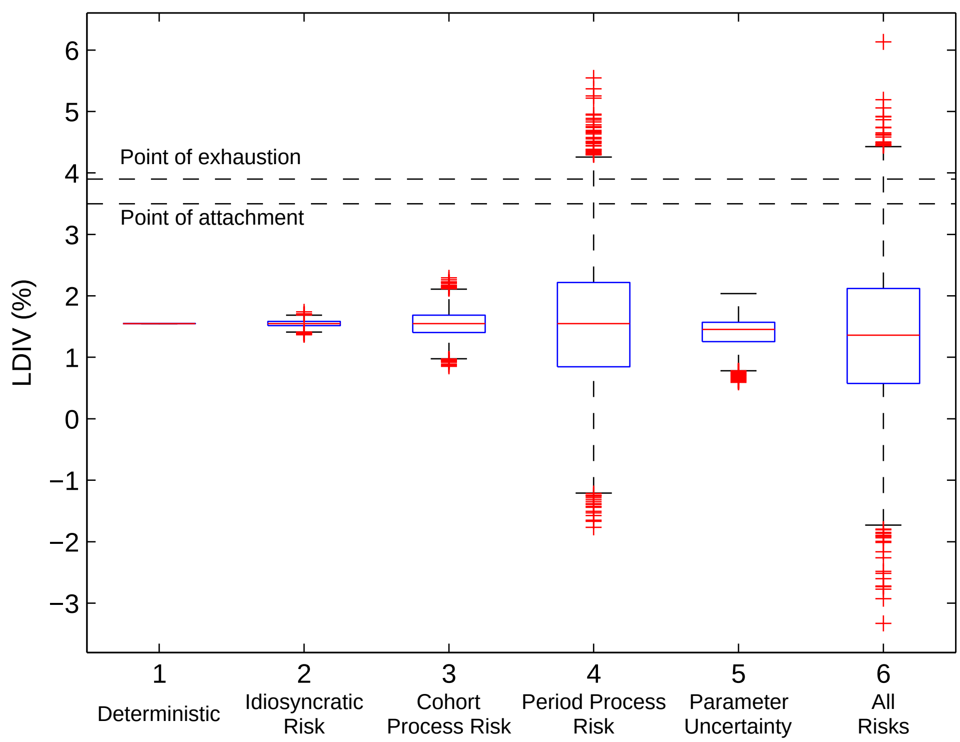


Figure 10: Boxplot of projected LDIV in 2016 allowing for different sources of risk

[25]This means that for parameter uncertainty, we generate multiple realisations of the fitted parameters and fit time series processes to each of these, but these time series processes are projected deterministically.

Note that we cannot directly allow for potential model risk without using a range of alternative models. Model risk is likely to be a significant risk, as we are dealing with the behaviour of a complicated index in the tail of its projected distribution. However, comparison with the results found by RMS can provide some insight into the potential magnitude of this risk, as the model used by RMS is based on fundamentally different principles to those adopted in this study. RMS find projected probabilities for the reduction of the principal which are slightly higher than those from our analysis, as shown in Table 3, which indicates that they would consider the Kortis bond a slightly more risky investment than we do. However, without further information regarding the full distribution of the LDIV projected by RMS and more details regarding the proprietary model they use to derive this, we are unable to properly reconcile the differences.

To illustrate some of the potential impact of model risk, we investigate the effect on the LDIV if some alternative methods were used to project the period parameters. First, we consider how our estimates of the LDIV would be affected if we do not allow for possible cointegration between the period functions. In this case, the period functions are projected using multivariate random walks with time-dependent drift functions of sufficient polynomial order to ensure that the projections are well-identified under the transformation in Equation 11. This means that the drift of $\kappa_t^{(1,p)}$ is of quadratic order and the drift of $\kappa_t^{(2,p)}$ is linear.

Figure 11 shows the distribution of the LDIV in 2016, in one case, allowing for cointegration between the period functions and, in the other, using a multivariate random walk to model them. We see that allowing for cointegration significantly reduces the projected risk in the LDIV. We find that, in this case, the estimated probability that the principal is reduced is approximately 4.6% (compared with approximately 4.3% when cointegration is allowed for) and the estimated probability that the principal is exhausted is approximately 2.4% (c.f., 1.8% in Table 3). We expect the use of cointegration to lower the projected probability of the principal being reduced, as doing so allows for the possibility that deviations from trend in one population affect deviations from trend in the other. This reduces the modelled basis risk between the populations and, therefore, increases the projected probability that the Kortis bond's principal will be paid in full.

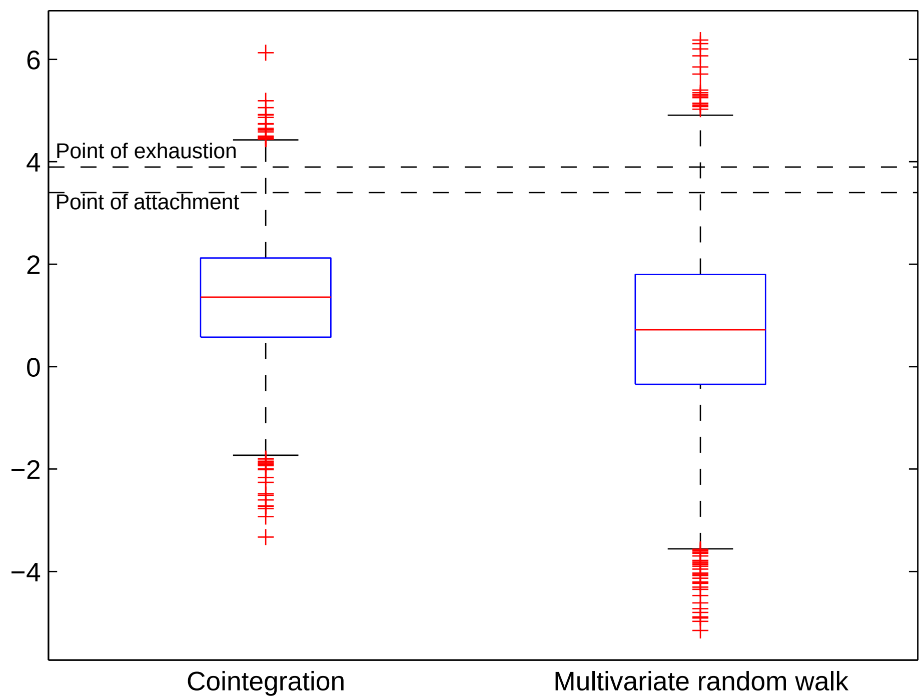


Figure 11: Boxplot of the projected LDIV in 2016 using cointegration and multivariate random walk processes

Second, we consider how our estimates of the LDIV would be affected if we assumed that the time series innovations were normally distributed (but still contemporaneously correlated), as opposed to bootstrapping the residuals from the fitting process as described in Section 3.2.2. We allow for the same cointegrating relationships. Figure 12 shows the distribution of the LDIV in 2016 using either bootstrapped or normally distributed innovations. As can be seen, the projected probabilities of the LDIV exceeding both the points of attachment and exhaustion are significantly higher in the case when we use bootstrapped innovations. We find that using normally distributed innovations reduces the probability of the principal being reduced to 2.5%, whilst the probability of complete exhaustion of the principal was found to be 0.8%. This indicates that using non-normal distributions for time series innovations, as in Wang et al. (2011), may be desirable, especially when trying to model extreme mortality scenarios.

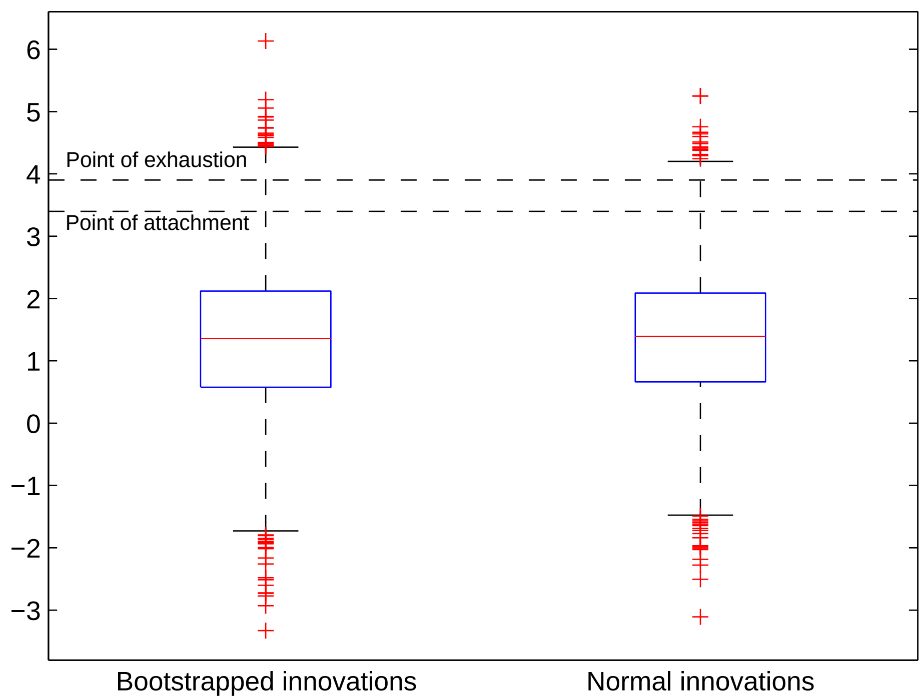


Figure 12: Boxplot of the projected LDIV in 2016 using bootstrapped and normally distributed innovations

6 Developing the market for longevity bonds

The Kortis bond is important and interesting because it is the prototype of a new breed of longevity bond indexed to the rates of improvement observed in mortality in different populations rather than in mortality rates directly (in the case of, say, the Vita bond) or in the survivorship of a cohort (in the case of the original survivor bond proposed by Blake and Burrows (2001) and marketed unsuccessfully by the European Investment Bank and BNP Paribas in 2004). It is therefore useful to try to generalise the structure of the Kortis bond and see which aspects of its design can be adapted for future bond issues.

The structure of the Kortis bond has a number of advantages. First, the issuer is the party seeking to hedge risk, which means that the bond can be structured for efficient risk transfer with an adjustable spread above LIBOR being set to meet the demands of risk-seeking investors. Second, its structure is similar both to existing catastrophe bonds and, in particular, to the mortality catastrophe bonds pioneered by the Swiss Re Vita bond. This means the structure will be familiar to investors and this increases the appeal and hence the marketability of the bond. Third, the LDIV refers to mortality improvement rates, which are a natural way of discussing longevity risk and commonly used by actuaries in practice. Fourth, it is a relatively short-term note, which makes it more marketable to insurance-linked securities investors, hedge funds and other investors with relatively short-term investment horizons. Finally, the floating rate coupons offer a significant margin above LIBOR, which may make it very attractive in the current low-yield investment environment.

These advantages need to be built upon by developing the structure of Kortis bond, if additional longevity bonds are going to be issued successfully. However, there are several limitations of this structure, which would need to be addressed in order to develop the market for longevity bonds further.

6.1 Developing the Kortis structure

Although the Kortis bond has many attractive features, we believe that its structure and the specification of the LDIV and population-specific indices can be developed in a number of ways to allow for a wider range of longevity-linked bonds and more effective transfer of longevity risk to the capital markets.

First, the Kortis bond is a relatively short-term financial instrument when compared against the term of the longevity risk in life assurance and annuity liabilities. While this may make it more marketable, it might not be ideal for hedging the long-term basis risk between the two populations. However, the design of the bond could easily be extended to allow for this. Figure 13 shows the mean projected LDIV along with estimated attachment and exhaustion points based on the same percentiles of the projected distribution of the LDIV for different maturities of the bond.[26]

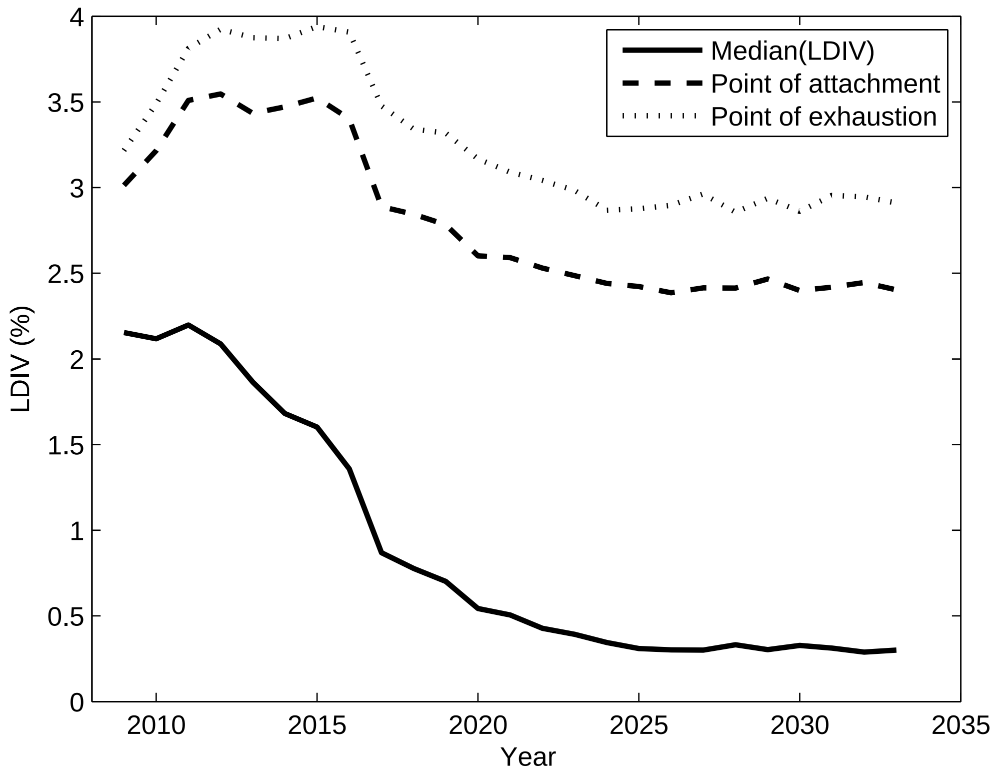


Figure 13: Projected attachment and exhaustion points for different maturity dates

We find the attachment point reduces slowly to around 2.4% as the term of the bond is lengthened, with a broadly similar pattern observed for the

[26]Assuming the same eight-year averaging period for the improvement rates in both populations.

exhaustion point, which converges to around 2.9%. It is worth noting, however, that estimates for the exhaustion point are subject to considerable uncertainty, as they are within the tail of the distribution of the forecast LDIV. A useful measure of the riskiness of the LDIV is the range over which the principal is reduced, which is roughly stable at 0.5% into the future. This is consistent with the results of Section 5.1, which suggested that the variability of the LDIV should not increase with time.[27] This means that it might be possible to extend the maturity of the Kortis bond significantly without excessively increasing its longevity-risk exposure. This implies that families of bonds could be issued, each with different maturities, in order to hedge longevity basis risk over a rolling period. Thus the Kortis structure could be used to hedge long-term longevity risk, whilst still being marketable to investors.

Second, we noted in Section 2 that Equation 2 calculates an average improvement rate for England & Wales and the US. It is therefore feasible to issue bonds constructed with reference to these indices for each country separately in order to hedge the risk of higher than expected improvements within the specific country, rather than on an index which is designed to measure the divergence in mortality rates between two different populations. This may be desirable for some bond issuers. For instance, an insurer writing annuity business in England & Wales could issue a bond linked to $\text{Index}(t, \text{EW})$ in order to hedge the longevity risk in its annuity book, with the principal of the bond being reduced if $\text{Index}(t, \text{EW})$ is greater than a given threshold. Compound indices could also be established in the same manner as the LDIV for hedging annuity books written across different countries, e.g., $\text{Index}_{60/40}(t) = 0.6\text{Index}(t, \text{EW}) + 0.4\text{Index}(t, \text{US})$ may be of interest to an insurer that writes 60% of its annuity business in the UK and 40% in the US. We therefore envisage a family of indices, with a market for different longevity bonds referenced to them.

Third, it is natural to ask whether a market participant could trade the LDIV directly through forward or swap contracts rather than through a bond. However, we suspect that this is not practical, given the infrequent nature of the updates to the LDIV and the fact that, as an average over an eight-year

[27]This is due to the eight-year averaging of the indices and the assumed time series processes, based on random walks, for the period functions, as discussed in Section 5.1.

period, the value of the LDIV can be estimated with reasonable accuracy several years in advance. Nevertheless, derivatives on the LDIV may be feasible if the averaging period were reduced to, say, one year, in the same way that derivatives on annual inflation rates exist.

Finally, we saw in Section 3.2.2 that the specification of the indices on different age ranges in England & Wales and the US meant that it was more sensitive to divergences in $\kappa_t^{(2,EW)}$ and $\kappa_t^{(2,US)}$. Since we found very little evidence to support cointegration between these period functions, they were responsible for much of the basis risk between the two populations. In contrast, if the indices for England & Wales and the US were defined on the same age range, the LDIV would be more sensitive to divergences in $\kappa_t^{(1,EW)}$ and $\kappa_t^{(1,US)}$. However, these are likely to be small, due to the presence of the cointegrating relationship. Therefore, we see that, by making relatively small changes to the definition of the indices used to construct the Kortis bond, we could calibrate a bond to be effective at transferring the risk arising from the specific features of the mortality curve we believe are generating the basis risk. This flexibility and ability to tailor the nature of the risk transferred may increase the appeal of the Kortis structure for hedging longevity risk.

6.2 Limitations of the Kortis structure

While we believe that the structure of the Kortis bond can be extended to develop the market, there are, however, a number of potential disadvantages to this structure which are more difficult to overcome. These limitations often make the modelling of the bond quite difficult, and so may restrict the appeal of this structure to only highly sophisticated investors.

One limitation of the Kortis structure is that the LDIV is strongly dependent on the cohort effects present. This runs counter to the interpretation of the Kortis bond as a "longevity trend" bond, which leads naturally to it being interpreted in terms of changes in mortality rates across different periods. Whether cohort effects genuinely exist within populations or whether they are simply artefacts of a (possibly mis-specified) model is still a controversial issue in some quarters (for instance, see Murphy (2009, 2010)). Therefore, the dependence of the LDIV on cohort effects might be considered a problem. In addition, to the extent that they exist, cohort effects can be both difficult to estimate robustly from the historical data due to their relatively small magnitude, and challenging to project plausibly into the future However, the combined use of the general procedure and the Bayesian framework can help with both of these issues. More fundamentally, however, it may not be desirable for the payoff from the bond to be strongly dependent on a quantity which can, in theory, be estimated significantly in advance of the bond's maturity. This would lead to informational asymmetries between the issuer and different buyers in the market, which may reduce demand and marketability to the extent of potentially making the market unviable.

Unfortunately, the dependence on cohort effects is inherent in the structure of the Kortis bond due to the collinearity of age, period and year of birth. Because *period* = *year of birth* + *age*, it is impossible to look at changes caused by one explanatory variable (in this case, period) whilst holding the other two constant. The definition of the LDIV compares mortality rates in different periods but at a fixed age, and therefore also allows the cohort to vary. A structure which held year of birth constant (and so did not depend on cohort effects) would compare mortality rates at different ages. As mortality rates tend to vary more by age than by year of birth, doing so would introduce new (and greater) problems than the dependence on cohort effects. This would also be counter to the relatively natural interpretation of the LDIV as measuring the divergence in mortality improvement rates. The dependence of the LDIV on cohort effects is, therefore, largely unavoidable.

Strong dependence on cohort effects is also a feature of many of the other mortality-linked securities that have been proposed, although this has not widely been commented upon. For instance, derivatives based on the survivorship of a given cohort will be strongly dependent on any differences between individuals born in that particular year of birth compared to neighbouring cohorts. This will manifest itself as systematically higher or lower survival rates than would otherwise be expected. Similarly, derivatives such as q-forwards, as proposed in Coughlan et al. (2007), will depend on the relevant cohort effects for the reference age at maturity. It is therefore essential that robust methods for detecting, analysing and projecting cohort effects, such as the general procedure and the use of the Bayesian framework, be developed in order to support the development of the market for longevity risk.

Another potential disadvantage of the structure of the Kortis bond is that it is designed as an option to hedge tail basis risk, i.e., reduction of the principal is an unlikely event and, in most scenarios, the bond will be redeemed at par. In order to model an option on an index, one requires at least two moments of the projected distribution of the index - mean and variance - to value the payoff. In contrast, modelling a forward on the same index requires only the first moment of the projected distribution to be modelled. Therefore, more complex models (which require both a greater number of assumptions to be made and a greater number of parameters to be estimated from historical data) are required for the modelling of embedded options - such as that in the Kortis bond - than if a similar structure with an embedded forward were used. This, and the relatively narrow window between the principal being paid in full and completely exhausted, leaves any analysis of the bond highly sensitive to model risk. Relatively small changes in the median of the projected LDIV can easily double the forecast probability that the bond is not redeemed at par, or just as easily reduce the forecast probability to zero. In addition, almost all models struggle to accurately forecast extreme scenarios due to the small number of extreme events in the historical data. In presenting our analysis, we have endeavoured to make our assumptions clear and justify their appropriateness. However, further studies using different methods to analyse the Kortis bond are necessary to reliably assess the probability that the principal is reduced.

Model risk is compounded by a lack of intuition regarding tail scenarios. With a conventional catastrophe bond (such as one linked to hurricane damage), investors can fairly easily imagine the sorts of catastrophes which would cause the bond's principal to be reduced and monitor whether such events have occurred. In contrast, it is relatively difficult to imagine how a scenario where the principal of the Kortis bond is exhausted looks different for the investor from one where the bond is paid back in full. This leaves valuations prior to the maturity of the bond almost entirely dependent on modelling assumptions, and hence model risk, with relatively few "sense checks" which can be performed on the reasonableness of the model output.

Related to this is the payoff structure of the Kortis bond. In both our analysis and that of RMS, we find that there is a very low probability of a partial reduction in the principal of the bond. We find that there is a 95.66% probability of the principal being paid back in full, a 1.76% chance of the principal being fully exhausted and, therefore, only a 2.58% probability of a partial repayment. This magnifies the impact of model risk, as small changes in the probability of the LDIV being greater than the point of attachment in 2016 can lead to dramatic changes in the valuation of the bond. In addition, if the observed LDIV approaches 3.4% before 2016, the value of the bond would fall precipitously. This may cause investors to try to sell the bond if this looks likely, leading to liquidity problems. Whilst this issue is not immediately significant, since there are very few longevity bonds currently trading, it is likely to become important if a market develops based on deep out-of-the-money longevity options, such as those proposed in Michaelson and Mulholland (2014). Thus, a market based on longevity options may suffer from systematic liquidity issues in the event of adverse experience, not unlike those affecting the markets for structured credit during the recent global financial crisis.

It is likely that Swiss Re focused on tail basis risk as this risk attracts high capital requirements under many modern solvency regimes and so retaining it is not capital efficient. However, this conflicts with a desire to generate a deep and liquid market for longevity-linked securities, which requires simpler securities that can be used to gauge market expectations of likely developments in longevity trends rather than extreme tail events. "Middle-of-the-range" scenarios are also less sensitive to model risk, with greater comparability between forecasts made using different modelling approaches, helping to assuage fears about the appropriateness of any given model and allowing a consensus to develop. We therefore believe that securities with embedded forwards on the LDIV, rather than those containing options based on the tail distribution of the index, would overcome this problem to some extent. This would also enable investors to share in some of the "upside" risk if the LDIV is observed to be lower than forecast.

In light of this analysis, it is interesting to note that Swiss Re has not followed up the Kortis bond with a family of similar longevity notes, in the same manner as with the Vita bond. This may be due to the features of the bond discussed above, which led to only modest investor interest in the bond. Nevertheless, we note that the Kortis bond was fully subscribed and that the coupon of 5.0% above LIBOR would be considered high-yield in the current environment of low interest rates. It may therefore be that other factors have been responsible for fact that the longevity trend bond structure pioneered by the Kortis bond has not been developed further. One potential explanation might be the postponement until 2016 of the implementation of Solvency II reducing (at least temporarily) the regulatory pressure to hedge tail basis risks. Alternatively, a more competitive reinsurance market may have increased the scope to transfer and hedge risks within the insurance sector rather than through the capital markets. Nevertheless, as longevity risk is a potentially huge global risk, such concerns should only postpone, rather than prevent, the development of the market in longevity risk transfer.

7 Conclusions

The longevity-risk market has evolved steadily since the first life book securitisations discussed in Cowley and Cummins (2005). In these securitisations, mortality was one of many risks being transferred. Pure longevity securities have been proposed in a number of different forms, from derivatives based on survivorship within a group such as in Blake and Burrows (2001) and Dowd et al. (2006), through derivatives based on mortality rates directly, such as the Swiss Re Vita bond and the q-forwards proposed in Coughlan et al. (2007). The Kortis bond represents a distinctly different approach to securitising longevity risk and is the first in a potentially large market of instruments based on improvements in mortality rates. Such a market would help the holders of longevity risk hedge their exposure and allow the emergence of a market-based consensus on the direction and level of risk posed by changes in longevity.

The Kortis bond attempts to hedge the basis risk caused by a divergence in the improvements in mortality between populations in England & Wales and the United States. It therefore requires mortality models which can estimate mortality rates robustly in both populations and project them into the future allowing for any dependence between them. In this paper, we demonstrate this by using the "general procedure" of Hunt and Blake (2014) to construct appropriate mortality models for each population and using well-identified cointegration techniques to project them.

In so doing, we find that there is approximately a 4% probability of the principal of the Kortis bond being reduced due to divergent mortality trends in the two populations, which is consistent with the analysis performed by RMS in pricing the bond at issue. We also find that one of the key factors determining the expected value of the LDIV is the historical cohort parameters estimated for each population. It is therefore very important that these are estimated robustly and projected with an appropriate allowance for their degree of risk. However, the uncertainty of the index is primarily governed by the process risk in the projections of the period effects. The estimated uncertainty caused by the basis risk between the evolution of mortality in England & Wales and the US is reduced by allowing for cointegration between the different populations since this captures the common co-movements between them.

We find that the design of the Kortis bond has many attractive features which could be used, developed and extended in the construction of other types of longevity bonds. These include the similarity in structure to existing insurance-linked securities (such as the Vita catastrophe bond) and the indexation to mortality improvement rates (which are familiar to actuaries and practitioners). We also identify a number of ways in which this structure could be adapted and extended in order to meet the needs of other holders of longevity risk. Therefore, we hope the Kortis bond inspires a new generation of longevity bonds and so contributes to the creation of a deeper market in longevity risk transfers.

A Models constructed by the "general procedure"

In Hunt and Blake (2014), a "general procedure" for constructing mortality models tailored to the specific features of individual datasets was proposed. In outline, the general procedure

- starts from a simple, static mortality model with a non-parametric[28] static age function;
- sequentially adds age/period terms to the model to detect and capture the age/period structure in the data;

[28]Defined in Hunt and Blake (2015e) as one fitted without any imposing any a priori structure across ages.

  - structure is detected by adding a non-parametric age/period term which will identify the feature explaining the largest proportion of the remaining structure in the data;
  - then this term is simplified into a parametric[29] form which identifies the same feature in the data more parsimoniously and with greater demographic significance;
  - then the statistical significance and robustness of the term is tested;

- finally adds a cohort term once all age/period structure has been captured by the model;

- tests the standardised deviance residuals of the model for any remaining structure, independence and normality.

This procedure was applied to data from the Human Mortality Database (2014) for England & Wales and the US for ages 50 to 100 and years 1950 to 2008 in order to construct mortality models capable of capturing all the relevant information in the data and therefore allowing it to be projected appropriately. In particular, experiments with simpler models such as those of Lee and Carter (1992) and Plat (2009) were felt to combine too many different (but correlated) features of the data into a single term (in the case of the model of Lee and Carter (1992)) or gave cohort parameters which both lacked demographic significance and captured possible residual age/period structure in the data (in the case of the Plat (2009) model).

A brief description of the terms in the models and their demographic significance is shown in Table 2. A fuller list of the parametric age functions in the "toolkit" developed as part of the general procedure is given in the Appendix of Hunt and Blake (2014). The algebraic form of these terms is reproduced in Table 4.

In the table, it is assumed that ages, x, run from 1 to X, whilst years, t, run from 1 to T. The normalisation factor is applied to ensure that $\sum_x |f(x)| = 1$ for all age functions and therefore that we can meaningfully compare the magnitudes of the period functions. Where there are free parameters in the

[29] Defined in Hunt and Blake (2015e) as one taking a specific functional form that is defined by an algebraic formula.

Name	Function $f(x) \propto$	Normalisation $N(\theta)$	Free Parameters
Constant	1	$\frac{1}{X}$	none
Linear	$x - \bar{x}$	$\frac{1}{\bar{x}(\bar{x}+1)}$	none
Quadratic	$(x - \bar{x})^2 - \sigma_x$	$\frac{1}{12}X(X+2)^2$	none
Spike	$(x - (x_c - a))\, I(x_c - a \leq x < x_c) +$ $((x_c + a) - x)\, I(x_c \leq x < x_c + a)$	$\frac{1}{a}$	x_c - peak a - width
Normal	$\exp\left(-\frac{(x-\hat{x})^2}{\sigma^2}\right)$	$\frac{1}{\sigma}$	$\hat{x}$ - location σ - width
Rayleigh	$(x - \hat{x})\exp(-\rho^2(x - \hat{x})^2)$	$0.5\rho^2$	$\hat{x}$ - location ρ - width^{-1}

Table 4: Age functions used in models constructed by the general procedure for England & Wales and the USA

age function,[30] the normalisation factor is a function of these parameters to ensure that the age function is "self-normalising". This means that the normalisation scheme $\sum_x |f^{(i,p)}(x, \theta^{(i,p)})| = 1$ always holds as the free parameters, $\theta^{(i,p)}$, vary when fitting the model. See Hunt and Blake (2015c) for a more detailed discussion of self-normalisation.

In Table 4 and the transformation of Equations 8, 9, 10 and 11, we also define constants

$$\bar{x} = \frac{1}{X}\sum_x x = 0.5(X+1)$$
$$\sigma_x = \frac{1}{X}\sum_x (x - \bar{x})^2$$
$$\rho_x = \frac{1}{\sigma_x}\sum_x (x - \bar{x})^4$$

[30]For instance, the normal function has parameters controlling the location and width of the curve

with similar definitions for $\bar{t}$, $\bar{y}$, etc. These allow the polynomial age functions to be easily orthogonalised, i.e.

$$\begin{aligned}
\sum_x (x-\bar{x}) &= 0 \\
\sum_x \left((x-\bar{x})^2 - \sigma_x\right) &= 0 \\
\sum_x \left((x-\bar{x})^3 - \rho_x(x-\bar{x})\right) &= 0 \\
\sum_x (x-\bar{x})\left((x-\bar{x})^2 - \sigma_x\right) &= 0 \\
\sum_x (x-\bar{x})\left((x-\bar{x})^3 - \rho_x(x-\bar{x})\right) &= 0 \\
\sum_x \left((x-\bar{x})^2 - \sigma_x\right)\left((x-\bar{x})^3 - \rho_x(x-\bar{x})\right) &= 0
\end{aligned}$$

In practice, this means that we use the Legendre polynomials in preference to the standard polynomials as age functions.

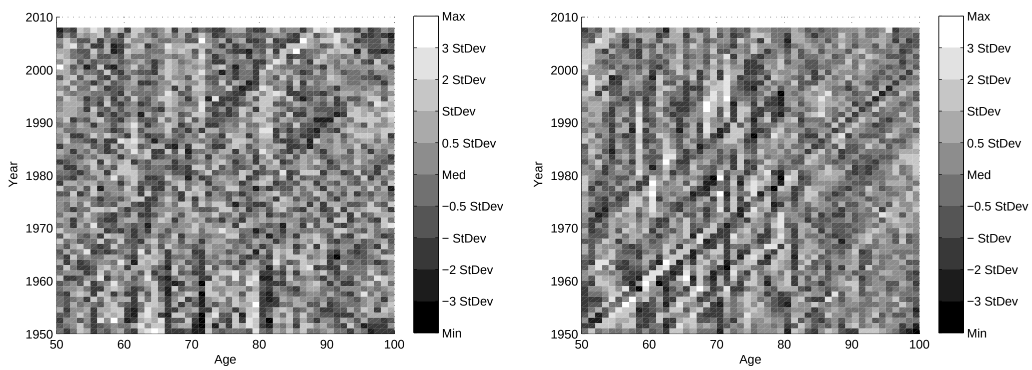
When fitting the final models, we obtain the parameters shown in Figures 3 and 4. These models have Bayes Information Criteria of -1.86×10^4 and -2.41×10^4 for England & Wales and the US, respectively, with 389 and 509 free parameters.[31] We also test the standardised deviance residuals from fitting the model as part of the general procedure. The moments of the residuals and a Jarque-Bera test of their normality are given in Table 5. We can see that the residuals are relatively close to normality (a p-value of 2.25%) for England & Wales, but are considerably more leptokurtic for the US.

Observations of the heat maps for the residuals in Figure 14 indicate that the residuals for England & Wales show no systematic patterns. However, there is a distinctive pattern of residuals from the US data diagonally along cohorts, especially for years of birth around 1900 which are not adequately captured by the cohort term. It may be that a cohort term which varies as individuals age (either using a non-parametric function such as the model of

[31] For comparison, the Lee and Carter (1992) model fitted to the same data obtains BICs of -2.42×10^4 and -4.24×10^4 with 158 free parameters for both populations.

	Residual mean	Standard deviation	Residual skewness	Residual kurtosis	Jarque-Bera statistic
EW	0.00	0.93	-0.01	3.25	7.68
US	0.01	0.91	0.06	3.71	65.34

Table 5: Moments of the residuals for England & Wales and the US

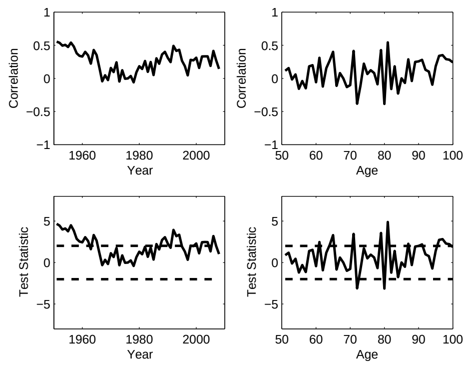


(a) England & Wales (b) US

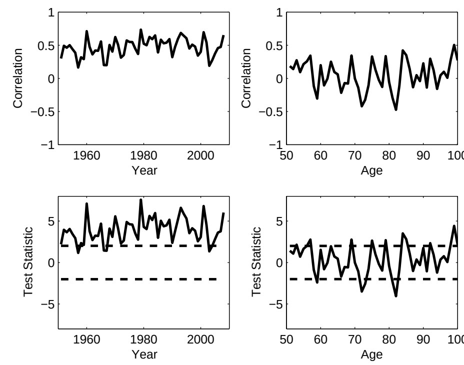
Figure 14: Heat maps of fitted residuals

Renshaw and Haberman (2006) or a linear parametric function as in Model M8 of Cairns et al. (2009)) might be more appropriate. However, we have not investigated this further. Tests of the independence of the standardised deviance residuals across ages and periods are shown in Figure 15 and indicate that, whilst the assumption of independence over age and period is rejected for both populations, there is considerably more year-on-year dependence for the residuals from the US model.

In summary, the use of the general procedure has allowed us to construct a mortality model for England & Wales which provides a superior fit to the data compared with alternative models. For the US, however, there is still considerable structure remaining in the residuals. As the general procedure seeks to find the most suitable model for the data from the class of age/period/cohort mortality models described in Hunt and Blake (2015e), we need to go beyond this class in order to construct an improved model for the US data.



(a) England & Wales



(b) US

Figure 15: Year-on-year and age-on-age correlations of fitted residuals

The problems with the US seem to arise chiefly as a result of issues with the underlying data, as indicated by our need to remove some datapoints as anomalies. It may be that these anomalies are caused by more heterogeneity present in the US population than that of England & Wales. If so, a possible extension to the modelling approach would be to adopt the negative binomial modelling framework of Li et al. (2009) which introduces additional parameters and a new fitting procedure to capture this heterogeneity. Alternatively, we might look to modify or extend the structure of the cohort term based on the evidence of the heat map in Figure 14b. Another suggestion would be to use the approach of Cairns et al. (2014) to adjust the underlying exposures for potential errors or allow stochastically for uncertainty in the underlying population exposure data if we believe the anomalies are due to errors in data collection. All of these approaches, however, go beyond the class of age/period/cohort models and the general procedure described in Hunt and Blake (2014) and so have been left for future work.

References

Arnold-Gaille, S., Sherris, M., 2013. Forecasting mortality trends allowing for cause-of-death mortality dependence. North American Actuarial Journal 17 (4), 273–282.

Bauer, D., Kramer, F., 2007. Risk and valuation of mortality contingent catastrophe bonds. Tech. rep., University of Ulm.

Beelders, O., Colarossi, D., 2004. Modelling mortality risk with extreme value theory: The case of Swiss Re's mortality-indexed bonds. Global Association of Risk Professionals, 26–30.

Blake, D., Burrows, W., 2001. Survivor bonds: Helping to hedge mortality risk. Journal of Risk and Insurance 68 (2), 339–348.

Blake, D., Cairns, A. J., Coughlan, G. D., Dowd, K., MacMinn, R., 2013. The new life market. Journal of Risk and Insurance 80 (3), 501–558.

Blake, D., Cairns, A. J., Dowd, K., 2006. Living with mortality: Longevity bonds and other mortality-linked securities. British Actuarial Journal 12 (1), 153–197.

Cairns, A. J., 2000. A discussion of parameter and model uncertainty in insurance. Insurance: Mathematics and Economics 27 (3), 313–330.

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

Cairns, A. J., Blake, D., Dowd, K., 2006b. Pricing death: Frameworks for the valuation and securitization of mortality risk. ASTIN Bulletin 36 (1), 79–120.

Cairns, A. J., Blake, D., Dowd, K., Coughlan, G. D., Epstein, D., Khalaf-Allah, M., 2011a. Mortality density forecasts: An analysis of six stochastic mortality models. Insurance: Mathematics and Economics 48 (3), 355–367.

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

Cairns, A. J., Blake, D., Dowd, K., Coughlan, G. D., Khalaf-Allah, M., 2011b. Bayesian stochastic mortality modelling for two populations. ASTIN Bulletin 41 (1), 29–59.

Cairns, A. J., Blake, D., Dowd, K., Kessler, A., 2014. Phantoms never die. Tech. rep., Herriot Watt University, Edinburgh.

Carter, R., Lee, D., 1992. Modeling and forecasting US sex differentials. International Journal of Forecasting 8, 393–411.

Chen, H., Cox, S. H., 2009. Modeling mortality with jumps: Applications to mortality securitization. Journal of Risk and Insurance 76 (3), 727–751.

Continuous Mortality Investigation, 2002. Working Paper 1 - An interim basis for adjusting the "92" series mortality projections for cohort effects. URL http://www.actuaries.org.uk/research-and-resources/pages/cmi-working-paper-

Coughlan, G. D., Epstein, D., Sinha, A., Honig, P., 2007. q-forwards: Derivatives for transferring longevity and mortality risks. JPMorgan Pension Advisory Group.

Cowley, A., Cummins, J., 2005. Securitization of life insurance assets and liabilities. Journal of Risk and Insurance 72 (2), 193–226.

Cox, S. H., Lin, Y., 2007. Natural hedging of life and annuity mortality risks. North American Actuarial Journal 11 (3), 1–15.

Darkiewicz, G., Hoedemakers, T., 2004. How the co-integration analysis can help in mortality forecasting. Tech. rep., Catholic University of Leuven.

Denuit, M. M., Devolder, P., Goderniaux, A.-M., 2007. Securitization of longevity risk: Pricing survivor bonds with Wang transform in the Lee-Carter framework. Journal of Risk and Insurance 74 (1), 87–113.

Dowd, K., Cairns, A. J., Blake, D., 2006. Mortality-dependent financial risk measures. Insurance: Mathematics and Economics 38 (3), 427–440.

Dowd, K., Cairns, A. J., Blake, D., Coughlan, G. D., 2011. A gravity model of mortality rates for two related populations. North American Actuarial Journal 15 (2), 334–356.

EIOPA, 2014. Technical specification for the preparatory phase (Part I). Tech. Rep., Frankfurt am Main, Germany.

Gaille, S., Sherris, M., 2011. Modelling mortality with common stochastic long-run trends. The Geneva Papers on Risk and Insurance Issues and Practice 36 (4), 595–621.

Haberman, S., Renshaw, A., 2011. A comparative study of parametric mortality projection models. Insurance: Mathematics and Economics 48 (1), 35–55.

Haberman, S., Renshaw, A., 2012. Parametric mortality improvement rate modelling and projecting. Insurance: Mathematics and Economics 50 (3), 309–333.

Human Mortality Database, 2014. Human Mortality Database. Tech. rep., University of California, Berkeley and Max Planck Institute for Demographic Research.
URL `www.mortality.org`

Hunt, A., Blake, D., 2014. A general procedure for constructing mortality models. North American Actuarial Journal 18 (1), 116–138.

Hunt, A., Blake, D., 2015a. A Bayesian approach to modelling and projecting cohort effects. Work in Progress.

Hunt, A., Blake, D., 2015b. Identifiability, cointegration and the gravity model. Work in Progress.

Hunt, A., Blake, D., 2015c. Identifiability in age/period mortality models. Work in Progress.

Hunt, A., Blake, D., 2015d. Identifiability in age/period/cohort mortality models. Work in Progress.

Hunt, A., Blake, D., 2015e. On the structure and classification of mortality models. Work in Progress.

Hyndman, R., Booth, H., Yasmeen, F., 2013. Coherent mortality forecasting: The product-ratio method with functional time series models. Demography 50 (1), 261–83.

Jarner, S. F., Kryger, E. M., 2011. Modelling mortality in small populations: The SAINT model. ASTIN Bulletin 41 (2), 377–418.

Johansen, S., 1991. Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. Econometrica 59 (6), 1551–1580.

Juselius, K., 2006. The cointegrated VAR model: Methodology and applications. Oxford University Press.

Koissi, M., Shapiro, A., Hognas, G., 2006. Evaluating and extending the Lee-Carter model for mortality forecasting: Bootstrap confidence interval. Insurance: Mathematics and Economics 38 (1), 1–20.

Lane, M. N., Beckwith, R., 2011. Prague Spring or Louisiana morning? Annual review for the four quarters, Q2 2010 to Q1 2011. Tech. rep., Lane Finacial LLC.

Lane, M. N., Beckwith, R., 2012. More return; More risk: Annual review for the four quarters, Q2 2011 to Q1 2012. Tech. rep., Lane Financial LLC.

Lane, M. N., Beckwith, R., 2013. Soft markets ahead!? Annual review for the four quarters, Q2 2012 to Q1 2013. Tech. rep., Lane Financial LLC.

Lane, M. N., Beckwith, R., 2014. Straw hats in winter: Annual review for the four quarters, Q2 2013 to Q1 2014. Tech. rep., Lane Financial LLC.

Lazar, D., Denuit, M. M., 2009. A multivariate time series approach to projected life tables. Applied Stochastic Models in Business and Industry 25 (6), 806–823.

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

Li, J. S.-H., Hardy, M. R., 2011. Measuring basis risk in longevity hedges. North American Actuarial Journal 15 (2), 177–200.

Li, J. S.-H., Hardy, M. R., Tan, K. S., 2009. Uncertainty in mortality forecasting: An extension to the classical Lee-Carter approach. ASTIN Bulletin 39 (1), 137–164.

Li, N., Lee, R. D., 2005. Coherent mortality forecasts for a group of populations: An extension of the Lee-Carter method. Demography 42 (3), 575–594.

Lin, Y., Cox, S. H., 2005. Securitization of mortality risks in life annuities. Journal of Risk and Insurance 72 (2), 227–252.

Michaelson, A., Mulholland, J., 2014. Strategy for increasing the global capacity for longevity risk transfer: Developing transactions that attract capital markets investors. Journal of Alternative Investments 17 (1), 18–27.

Mitchell, D., Brockett, P. L., Mendoza-Arriaga, R., Muthuraman, K., 2013. Modeling and forecasting mortality rates. Insurance: Mathematics and Economics 52 (2), 275–285.

Murphy, M., 2009. The "golden generations" in historical context. British Actuarial Journal 15 (S1), 151–184.

Murphy, M., 2010. Re-examining the dominance of birth cohort effects on mortality. Population and Development Review 36 (2), 365–90.

Plat, R., 2009. On stochastic mortality modeling. Insurance: Mathematics and Economics 45 (3), 393–404.

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

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

Salhi, Y., Loisel, S., 2009. Longevity basis risk modeling: A co-integration based approach. Tech. rep., University of Lyon.

Standard and Poor's, 2010. Presale information: Kortis Capital Ltd. Tech. rep., Standard and Poors.

Villegas, A. M., Haberman, S., 2014. On the modeling and forecasting of socioeconomic mortality differentials: An application to deprivation and mortality in England. North American Actuarial Journal 18 (1), 168–193.

Wang, C.-W., Huang, H., Liu, I.-C., 2011. A quantitative comparison of the Lee-Carter model under different types of non-Gaussian innovations. The Geneva Papers on Risk and Insurance Issues and Practice 36 (4), 675–696.

Wang, J. L., Huang, H., Yang, S. S., Tsai, J. T., 2009. An optimal product mix for hedging longevity risk in life insurance companies: The immunization theory approach. Journal of Risk and Insurance 77 (2), 473–497.

Willets, R., 1999. Mortality in the next millennium. Staple Inn Actuarial Society.

Willets, R., 2004. The cohort effect: Insights and explanations. British Actuarial Journal 10 (4), 833–877.

Yang, S. S., Wang, C.-W., 2013. Pricing and securitization of multi-country longevity risk with mortality dependence. Insurance: Mathematics and Economics 52 (2), 157–169.

Zhou, R., Wang, Y., Kaufhold, K., Li, J. S.-H., Tan, K. S., 2014. Modeling period effects in multi-population mortality models: Applications to Solvency II. North American Actuarial Journal 18 (1), 150–167.