DISCUSSION PAPER PI-2002

CBDX: A Workhorse Mortality Model from the Cairns-Blake-Dowd Family

Kevin Dowd, Andrew J.G. Cairns & David Blake

January 2020

ISSN 1367-580X

The Pensions Institute
Cass Business School
City, University of London
106 Bunhill Row
London EC1Y 8TZ
UNITED KINGDOM
http://www.pensions-institute.org/

CBDX: A Workhorse Mortality Model from the Cairns-Blake-Dowd Family

Kevin Dowd,[*] Andrew J.G. Cairns[♣] and David Blake[♦]

This draft: 5 January 2020

Abstract

The purpose of this paper is to identify a workhorse mortality model for the adult age range (i.e., excluding the accident hump and younger ages). It applies the "general procedure" (GP) of Hunt and Blake (2014) to identify an age-period model that fits the data well before adding in a cohort effect that captures the residual year-of-birth effects arising in the original age-period model. The resulting model is intended to be suitable for a variety of populations, but economises on the number of period effects in comparison with a full implementation of the GP. We estimate the model using two different iterative Maximum Likelihood (ML) approaches – one Partial ML and the other Full ML – that avoid the need to specify identifiability constraints.

Key Words: mortality rates, Cairns-Blake-Dowd mortality model, CBDX mortality model

JEL codes: G220, G230, J110

1. Introduction

Broadly speaking, there are two main families of stochastic mortality models. The first are descendants of the Lee-Carter model (Lee and Carter, 1992). The original Lee-Carter model had two age effects and one period effect, with variants adding a second period component and cohort effects (e.g., Renshaw and Haberman, 2003, 2006). The second is the Cairns-Blake-Dowd family (Cairns *et al.*, 2006, 2009). The original CBD model (later reparameterised as M5 using the classification of Cairns *et al*., 2009) had two period effects and no age or cohort effects, but a later extension, M6, added a cohort effect and a further extension, M7, added a third period effect. An important (near) generalization of model M6 is the Plat model (Plat, 2009), which has two or three period effects, a cohort effect and an age effect. See Cairns (2014) for a review of these models.

This article sets out a new mortality model, denoted CBDX,[1] which is, loosely speaking, an extension of CBD/M5 with age and cohort effects and up to three period effects. The aim is to create a workhorse model for use by practising actuaries who wish to model the

[*] Durham University Business School, Mill Hill Lane, Durham DHL 3LB, United Kingdom. Corresponding author: kevin.dowd@durham.ac.uk.

[♣] Maxwell Institute for Mathematical Sciences and Department of Actuarial Mathematics and Statistics, Heriot-Watt University, Edinburgh, EH14 4AS, United Kingdom.

[♦] Pensions Institute, Cass Business School, City University of London, 106 Bunhill Row, London, EC1Y 8TZ, United Kingdom.

[1] The term CBDX was first used in Hunt and Blake (2020a,b).

mortality of groups of adults over a wider range of ages than is normally advisable for the standard CBD models (e.g., M5 and M7 in Cairns et al., 2009). The motivation comes from the "general procedure" (GP) for constructing mortality models outlined in Hunt and Blake (2014). The GP "identifies every significant demographic feature in the data in a sequence, beginning with the most important. For each demographic feature, we need to apply expert judgment to choose a particular parametric form to represent it. By following the GP, it is possible to construct mortality models with sufficient terms to capture accurately all the significant information present in the age, period, and cohort dimensions of the data. In particular, the GP prevents structure in the data that is genuinely associated with a period effect being wrongly allocated to a cohort effect (p. 116)".

Using mortality data for the UK over the full age range, Hunt and Blake (2014, Figure 10) identify seven age-period terms and one cohort term. However, the majority of the improvement in goodness of fit comes from the first three period terms. An analysis of US mortality data shows that five period terms can be identified, but again most of the improvement in goodness of fit comes from the first three period terms. Using England & Wales male mortality data, our results lead us to conclude that the best workhorse model for wider age ranges, that start as low as 40 and go up to 89, is CBDX3, the CBDX model with three period effects.

The paper is organised as follows. Section 2 sets out the CBDX model. Section 3 compares this model to related models. Section 4 discusses the estimation procedure for the CBDX model. Section 5 describes the data and results from estimating the model. Section 6 reports the results of a BIC model comparison, in which the CBDX model is compared with well-known related models. Section 7 provides an illustrative financial application. Section 8 proposes a novel method for forecasting that blends statistical analysis of the historical period effects with expert judgement about central forecasts. Section 9 discusses some features and potential extensions of the CBDX model. Section 10 concludes.

2. The CBDX Model

Consider the following mortality model:

$$\log m(t,x) = \alpha(x) + \sum_{i=1}^{K} \beta_i(x)\kappa_i(t) + \gamma(c) \quad (1)$$

where $m(t,x)$ is a death rate, t refers to the time period, x refers to age and $c = t - x$ refers to the year of birth. The $\alpha(x)$, $\kappa(t)$ and $\gamma(c)$ are the age-related, period-related and cohort-related effects, and the parameters $\beta_1 = 1$, $\beta_2 = (x - \bar{x})$, $\beta_3 = (x - \bar{x})^2 - \sigma_x^2$ are fixed throughout, and where $\bar{x}$ and σ_x^2 are the mean and variance of the ages in our sample age range. As discussed below, this model is similar in spirit to Plat (2009) in adding an age effect, $\alpha(x)$, to the original CBD models.

This model comes in three versions:

- $K = 1$, i.e., the case with only one $\kappa(t)$ variable – denoted model CBDX1

- $K = 2$, i.e., the case with two $\kappa(t)$ variables – denoted model CBDX2
- $K = 3$, i.e., the case with three $\kappa(t)$ variables – denoted model CBDX3.

This CBDX family of models is similar to the CBD family except for the following:

- It uses the log of the death rate, $\log m(t,x)$, instead of the logit of the mortality rate, logit $q(t,x)$, as the dependent variable
- It includes (as in Plat, 2009) an additional age-related variable, $\alpha(x)$.
- Whereas CBD treat $\gamma(c)$ process as a trendless mean-reverting AR(1) process, we treat $\gamma(c)$ as a residual that should be trendless and mean-reverting if the model is well-specified.
- Model fitting differs from previous CBD family models by using a sequential maximum likelihood algorithm which is not dependent on any explicit identifiability constraints.[2]

The Lee-Carter model showed the advantages of including a non-parametric "base mortality table," $\alpha(x)$, in Equation (1). The advantages of our treatment of the $\gamma(c)$ process, of using a sequential algorithm, and of not relying on explicit identifiability constraints are discussed below.

3. The CBDX Model and Related Models

The CBDX model is related to a number of well-known mortality models, in particular Lee-Carter, M3, M5, M6, M7 and the Plat model.

The Lee-Carter model, referred to by Cairns *et al*., (2009) as M1, is:

$$\log m(t,x) = \alpha(x) + \beta_2(x)\kappa_2(t) \tag{2}$$

where $\alpha(x)$ is a non-parametric "base mortality table," $\beta_2(x)$ is an age-dependent parameter, and n_a is the number of ages in the sample data age range. This model satisfies two identifiability conditions set out, e.g., in Cairns *et alia* (2009, p. 7).

Model M3 (Cairns et al., 2009) is the standard Age-Period-Cohort model, and is identical to what we label here as CBDX1, and so is not considered further.

Model M5 (Cairns et al., 2009) is a reparameterised version of the CBD two-factor mortality model (et al. 2006) and postulates that the mortality rate $q(t,x)$ satisfies:

$$\text{logit } q(t,x) = \kappa_1(t) + (x - \bar{x})\kappa_2(t) \tag{3}$$

[2] However, identifiability constraints are implicit in the model fitting process as a result of the sequential approach to parameter estimation. This means that once the sequential model fitting process has produced a final set of parameter estimates, the imposition of explicit identifiability constraints at this stage might change the parameter estimates themselves, but would not change the estimated mortality rates or the value of the log-likelihood.

where we assume that $q(t,x) = 1 - e^{-m(t,x)}$, $\bar{x}$ is the average of the ages used in the sample age range and $\boldsymbol{\kappa}_t = [\kappa_1(t), \kappa_2(t)]$ follows a two-dimensional random walk with drift:

$$\boldsymbol{\kappa}_t = \boldsymbol{\kappa}_{t-1} + \boldsymbol{\mu} + \boldsymbol{C}Z_t \tag{4}$$

Model M5 does not have any identifiability issues.

Model M6 is a generalised version of M5 with a cohort effect:

$$\text{logit}\, q(t,x) = \kappa_1(t) + (x - \bar{x})\kappa_2(t) + \gamma(c) \tag{5}$$

where the $\boldsymbol{\kappa}_t$ process follows (4) and $\gamma(c)$ is the cohort effect. This model satisfies the identifiability constraints:

$$\sum_{c=1}^{n_c} \gamma(c) = 0 \quad \text{and} \quad \sum_{c=1}^{n_c} c\gamma(c) = 0 \tag{6}$$

where n_c is the number of years of birth in the sample data.

Model M7 is a second generalised version of M5 with a cohort effect:

$$\text{logit}\, q(t,x) = \kappa_1(t) + (x - \bar{x})\kappa_2(t) + [(x - \bar{x})^2 - \sigma_x^2]\kappa_3(t) + \gamma(c) \tag{7}$$

where the period effects $\boldsymbol{\kappa}_t$ now follow a three-dimensional random walk with drift and $\gamma(c)$ is a cohort effect modelled as an AR(1) process. This model satisfies the identifiability constraints:

$$\sum_{c=1}^{n_c} \gamma(c) = 0, \sum_{c=1}^{n_c} c\gamma(c) = 0 \text{ and } \sum_{c=1}^{n_c} c^2\gamma(c) = 0. \tag{8}$$

More details on the models can be found in Cairns *et al.* (2009).

CBDX2 and CBDX3 are similar to models M6 and M7, respectively, except in so far as they: (a) use $\log m(t,x)$ rather than $\text{logit}\, q(t,x)$ as the dependent variable; (b) include an additional age effect, $\alpha(x)$; (c) use a sequential estimation algorithm and (d) (as a result of (c)) do not require any explicit identifiability constraints. CBDX2 is the same as the two period effect version of the Plat (2009) model and CBDX3 adds a third period effect to CBDX2.[3]

The age, period and cohort effects for the different models are summarised in Table 1.

[3] CBDX3 requires seven identifiability constraints: $\sum_{t=1}^{n_y} \kappa_1(t) = 0, \sum_{t=1}^{n_y} \kappa_2(t) = 0, \sum_{t=1}^{n_y} \kappa_3(t) = 0, \sum_{c=1}^{n_c} \gamma(c) = 0, \sum_{c=1}^{n_c} c\gamma(c) = 0, \sum_{c=1}^{n_c} c^2\gamma(c) = 0$ and $\sum_{c=1}^{n_c} c^3\gamma(c) = 0$. CBDX2 requires 5 constraints because it drops the last of the κ and γ constraints. CBDX1 requires 3 constraints because it drops the last two of the κ and γ constraints.

Table 1: Dependent Variables and Age-, Period- and Cohort-Effects for Different Mortality Models

Model	Dependent Variable	Age-Related	Period-Related			Cohort Related
CBDX1	$\log m(t,x)$	$\alpha(x)$	$\kappa_1(t)$			$\gamma(c)$
CBDX2	$\log m(t,x)$	$\alpha(x)$	$\kappa_1(t)$	$\kappa_2(t)$		$\gamma(c)$
CBDX3	$\log m(t,x)$	$\alpha(x)$	$\kappa_1(t)$	$\kappa_2(t)$	$\kappa_3(t)$	$\gamma(c)$
M1	$\log m(t,x)$	$\alpha(x), \beta_2(x)$		$\kappa_2(t)$		
M5	logit $q(t,x)$		$\kappa_1(t)$	$\kappa_2(t)$		
M6	logit $q(t,x)$		$\kappa_1(t)$	$\kappa_2(t)$		$\gamma(c)$
M7	logit $q(t,x)$		$\kappa_1(t)$	$\kappa_2(t)$	$\kappa_3(t)$	$\gamma(c)$

Note: x = age, t = period, c = birth year.

It is also helpful to set out which models incorporate others as special cases of others. The two core special cases are M3 and M5. Ignoring differences in algorithms and identifiability conditions for present purposes:

- CBDX1 is a special case of CBDX2.
- CBDX2 is a special case of CBDX3.
- M5 is a special case of M6.
- M6 is a special case of M7.

4. Estimation of the CBDX Model

4.1 The sequential estimation algorithm

Let $\Theta = \{\alpha, \kappa, \gamma\}$ be the parameter set to be estimated, where κ corresponds to $\kappa_1(t)$ for CBDX1, $(\kappa_1(t), \kappa_2(t))$ for CBDX2 and $(\kappa_1(t), \kappa_2(t), \kappa_3(t))$ for CBDX3, for $t = 1, \dots, n_y$. The Poisson log-likelihood is then

$$(9)\quad l(\alpha,\kappa,\gamma) = \sum_{t,x}[D(t,x)\log m(t,x) - E(t,x)m(t,x)] + constant$$
$$= \sum_{t,x} D(t,x)\big(\alpha(x) + \kappa_1(t) + \beta_2(x)\kappa_2(t) + \beta_3(x)\kappa_3(t) + \gamma(t-x)\big)$$
$$- \sum_{t,x} E\big(t,x_j\big)\exp\big(\alpha(x) + \kappa_1(t) + \beta_2(x)\kappa_2(t) + \beta_3(x)\kappa_3(t) + \gamma(t-x)\big)$$
$$+ constant$$

where $\beta_2(x) = (x - \bar{x})$ and $\beta_3(x) = [(x-\bar{x})^2 - \sigma_x^2]$, $D(t,x)$ is the number of deaths in year t aged x last birthday, and $E(t,x)$ is the average number of individuals in year t aged x last birthday.

We propose the following sequential estimation algorithm.

Partial ML (PML)

Stage P1 (Alpha, Kappa) involves the joint estimation of the $\alpha(x)$ and $\kappa(t)$ parameters. Stage P2 (Gamma) uses $\hat{\alpha}(x)$ and $\hat{\kappa}(t)$ to estimate the $\gamma(c)$ parameters, $\hat{\gamma}(c)$. This two-stage procedure is very much in the spirit of the general procedure of Hunt and Blake (2014), in which the model structure identifies sufficient age and period effects first before adding a cohort effect to capture small residual patterns in the data by year of birth. At this point PML is complete. Although PML does not require any explicit identifiability constraints, they are implicit in the fact that the age and period effects are kept fixed while we estimate the cohort effects.

Full ML (ML)

We start with the three stages of PML. The Stage F1 (Alpha, Kappa) revision uses the latest $\hat{\gamma}(c)$ to re-estimate the $\alpha(x)$ and $\kappa(t)$. The Stage F2 (Gamma) revision then uses the latest $\hat{\alpha}(x)$ and $\hat{\kappa}(t)$ to re-estimate the $\gamma(c)$ state variables. We then repeat stages F1 and F2 until all parameter estimates and the likelihood converge. These estimates are those that maximise the likelihood function.

4.2 Illustrating the sequential estimation algorithm using CBDX3

Consider CBDX3. In this case, $K = 3$, $\beta_1 = 1$, $\beta_2(x) = (x - \bar{x})$ and $\beta_3(x) = (x - \bar{x})^2 - \sigma_x^2$, and (1) becomes:

$$\log m(t,x) = \alpha(x) + \kappa_1(t) + (x - \bar{x})\kappa_2(t) + [(x - \bar{x})^2 - \sigma_x^2]\kappa_3(t) + \gamma(c) \quad (10)$$

Stage P1A (Alpha)

We now take (10) and set the $\kappa(t)$ and $\gamma(c)$ to zero to obtain:

$$log\ m(t,x) = \alpha(x) \quad (11)$$

$$m(t,x) = e^{\alpha(x)} \quad (12)$$

Implicitly, we are assuming that that the starting values of the $\kappa(t)$ and $\gamma(c)$ are zero for the purposes of estimating $\alpha(x)$.

We then substitute (11) and (12) into (9) to obtain the log-likelihood function expressed in terms of $\alpha(x)$:

$$l(\alpha) = \sum_{j=1}^{n_a}\sum_{t=1}^{n_y} D(t,x_j)\alpha(x_j) - \sum_{j=1}^{n_a}\sum_{t=1}^{n_y} E(t,x_j)e^{\alpha(x_j)} + \text{constant} \quad (13)$$

where n_y is the number of sample years.

The ML estimator for $\alpha(x_i)$ is obtained by differentiating (13) with respect to $\alpha(x_j)$ and rearranging the first-order condition to give:

(14) $$\alpha(x_j) = log\left\{\frac{\sum_{t=1}^{n_y} D(t,x_j)}{\sum_{t=1}^{n_y} E(t,x_j)}\right\}$$

for $j = 1, \dots, n_a$.

Stage P1K (Kappa)

In this stage, we input the estimate $\hat{\alpha}(x)$ obtained from Stage P1A into (10):

(15) $$\log m(t,x) = \hat{\alpha}(x) + \kappa_1(t) + (x-\bar{x})\kappa_2(t) + [(x-\bar{x})^2 - \sigma_x^2]\kappa_3(t)$$

We then obtain the log-likelihood function

(16) $$l(\kappa_1, \kappa_2, \kappa_3) =$$
$$\sum_{t=1}^{n_y}\sum_{j=1}^{n_a} D(t,x_j)\left(\hat{\alpha}(x_j) + \kappa_1(t) + (x_j-\bar{x})\kappa_2(t) + \left[(x_j-\bar{x})^2 - \sigma_x^2\right]\kappa_3(t)\right) -$$
$$\sum_{t=1}^{n_y}\sum_{j=1}^{n_a} E(t,x_j) e^{\hat{\alpha}(x_j)} e^{\kappa_1(t)} e^{(x_j-\bar{x})\kappa_2(t)}\, e^{[(x_j-\bar{x})^2-\sigma_x^2]\kappa_3(t)}$$
$$+ \text{constant}$$

We obtain estimates of $\kappa_1(t)$, $\kappa_2(t)$ and $\kappa_3(t)$ that maximise (16) using a Newton-Raphson iterative algorithm.

Stage P1C

Repeat stages P1A and P1K until the estimates and the log-likelihood have converged.

Stage P2G (Gamma)

In this stage, we input the estimates $\hat{\alpha}(x)$, $\hat{\kappa}_1(t)$, $\hat{\kappa}_2(t)$ and $\hat{\kappa}_3(t)$ into (10) to obtain:

(17) $$\log m(t,x) = \hat{\alpha}(x) + \hat{\kappa}_1(t) + (x-\bar{x})\hat{\kappa}_2(t) + [(x-\bar{x})^2 - \sigma_x^2]\hat{\kappa}_3(t) + \gamma(c)$$

We then derive the log-likelihood function in terms of γ:

(18) $$l(\gamma) =$$
$$\sum_{t=1}^{n_y}\sum_{j=1}^{n_a} D(t,x_j)(\hat{\alpha}(x_j) + \hat{\kappa}_1(t) + (x_j-\bar{x})\hat{\kappa}_2(t) + [(x_j-\bar{x})^2 - \sigma_x^2]\hat{\kappa}_3(t) + \gamma(t-x_j))$$
$$-\sum_{t=1}^{n_y}\sum_{j=1}^{n_a} E(t,x_j) e^{\hat{\alpha}(x_j)} e^{\hat{\kappa}_1(t)} e^{(x_j-\bar{x})\hat{\kappa}_2(t)} e^{[(x_j-\bar{x})^2-\sigma_x^2]\hat{\kappa}_3(t)} e^{\gamma(t-x_j)}$$
$$+ \text{constant}$$

We then differentiate the log-likelihood function (18) with respect to $\gamma(c)$ and rearrange the first-order conditions to obtain the MLE for $\gamma(c)$:

(19) $$\gamma(c) = log\left\{\frac{\sum_{t=1}^{n_y}\sum_{j=1}^{n_a} w_c(t,x_j)D(t,x_j)}{\sum_{t=1}^{n_y}\sum_{j=1}^{n_a} w_c(t,x_j)E(t,x_j)e^{\hat{\alpha}(x_j)}e^{\hat{\kappa}_1(t)}e^{(x_j-\bar{x})\hat{\kappa}_2(t)}e^{[(x_j-\bar{x})^2-\sigma_{\bar{x}}^2]\hat{\kappa}_3(t)}}\right\}$$

for $c = 1, \ldots, n_c$, and where the weighting function $w_c(t,x)$ satisfies

(20) $$w_c(t,x) = \begin{Bmatrix}1\\0\end{Bmatrix} \text{ if } \begin{cases}t - x = c\\ t - x \neq c.\end{cases}$$

Full maximum likelihood

As explained above under *Full ML (ML).*

4.3. Estimation of CBDX1 and CBDX2

The estimation of CBDX1 is similar to that of CBDX3 except we keep $\kappa_2(t)$ and $\kappa_3(t)$ set to zero, and the estimation of CBDX2 is similar to that of CBDX3 except we keep $\kappa_3(t)$ set to zero.

5. Data and Results from Estimating the Model

5.1. Data

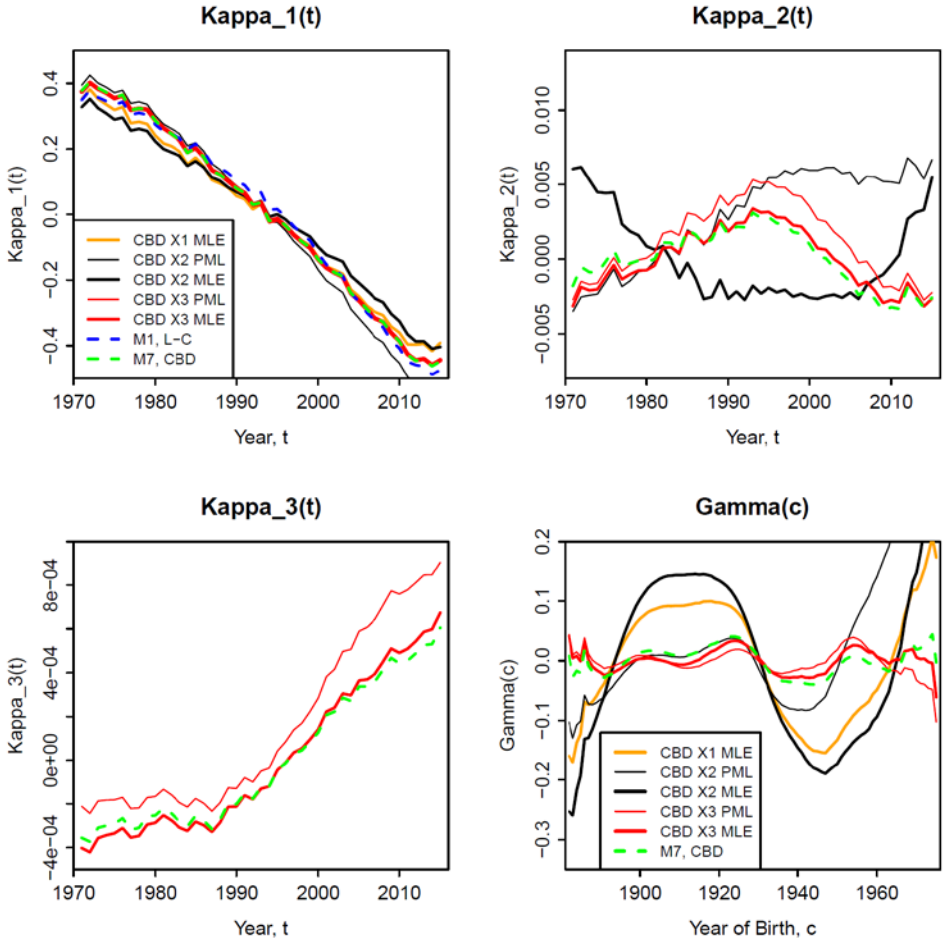
The model is estimated using England & Wales (E&W) male death rate data spanning the years 1971–2015 and ages 40–89. The data came from the Life & Longevity Markets Association. The data have had anomalies removed using the methodology of Cairns et al. (2016). The number of cohorts is $n_c = n_y + n_a - 1$ and let x be the ages vector.

Using the above data set, the key parameters of the model are $n_y = 45, n_a = 50, n_c = 94$, and birth years 1882 – 1975.

5.2. Results

We compare results for the following models: Lee-Carter (M1, full MLE), the original CBD model (M5, MLE), second-generation CBD models M6 and M7 (MLE), and CBDX1, CBDX2 and CBDX3 (both PML and MLE in each case). Figure 1 compares fitting results for a selection of these models for the fitted period and cohort effects.

Figure 1: Estimated Period and Cohort Effects for Selected Models



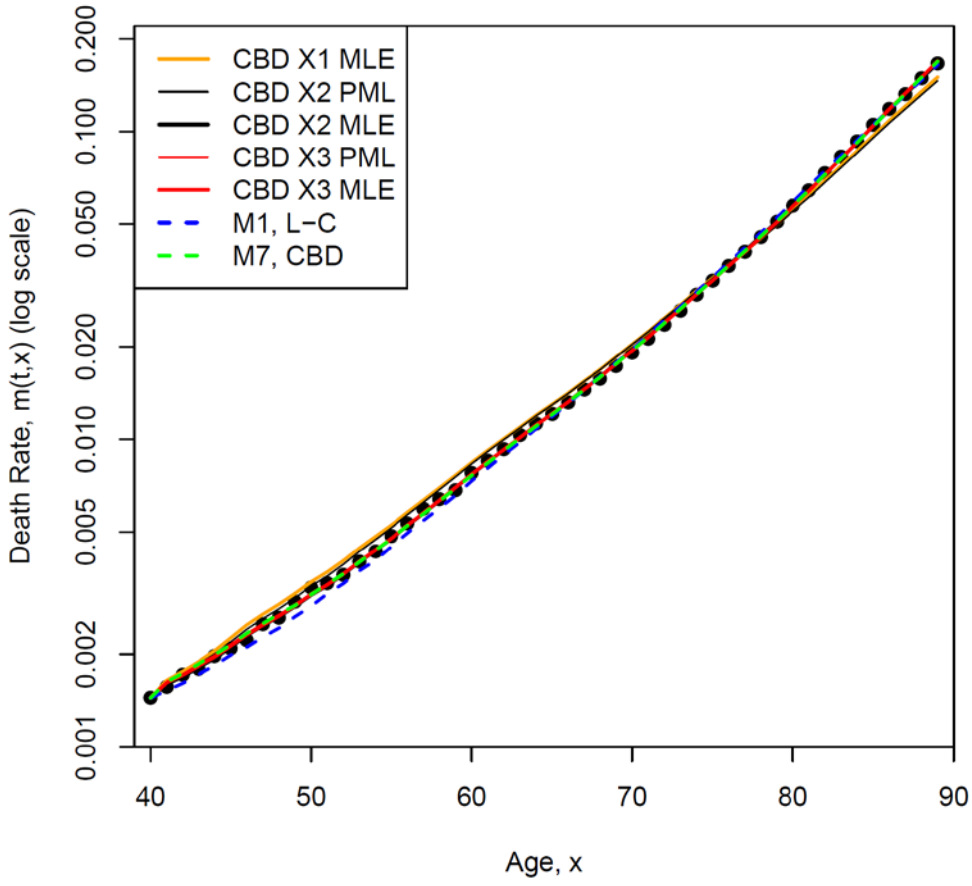
For these selected models, we see that estimates for $\kappa_1(t)$ are quite consistent (top left panel)[4] reinforcing the notion that $\kappa_1(t)$ captures the headline mortality improvements at all levels. There is a close correspondence between M7 and CBDX3 (PML and MLE) in all four plots indicating that there is robustness in the addition of $\alpha(x)$ and in the switch from logit $q(t,x)$ to $\log m(t,x)$ as dependent variable. On the other hand, differences in the fitted cohort effects between the six models plotted are quite striking. Specifically, cohort effects for CBDX1 and CBDX2 (MLE) are much larger and have much more structure than for M7 and CBDX3, while for CBDX2, the PML and MLE cohort effects are quite different, indicating a lack of robustness.

This is a clear indication that CBDX1 and CBDX2 lack sufficient structure in the main age and period effects and that the cohort effect is being used as a proxy for capturing unmodelled age or period effects. This point can be seen by comparing the results for CBDX2 for the PML and MLE cases. The PML fits the age and period effects first, and the cohort effect second. As a rule of thumb, residuals by year of birth for the age-period case should approximately match the fitted cohort in the MLE. However, this is not the case for CBDX2, and is a further indication that, for this dataset, the age and period effects in CBDX2 lack sufficient structure.

[4] To aid comparison, the period effect for the Lee-Carter model has been divided by n_a and period effects for M7 have been shifted to be centred around zero.

In Figure 2, as a snapshot of how well the models fit the data, we show actual death rates (dots) versus fitted death rates for various models. We can see that all of the models shown give a reasonable fit, but some models clearly fit much more closely than others. Models that fit less well are M1, CBDX1-MLE and CBDX2-PML, while the others are much closer and difficult to distinguish. Again, we see a significant difference between CBDX2-PML and CBDX2-MLE, indicating that the two age-period effects are not sufficient to capture the dynamics over the 45-year period. In contrast, the difference between CBDX3-PML and CBDX3-MLE is negligible.

Figure 2: Crude and Fitted Death Rates in 2015 for Selected Models



6. Bayesian Information Criteria Model Comparisons

To assess how well the models compare, Table 2 shows the Bayesian Information Criterion (BIC) results for M1, M5, M6, M7, and CBDX1, CBDX2, CBDX3 under their PML and MLE variants.[5]

[5] Here the Bayesian Information Criterion is defined as $BIC = -2l + k \log n_{obs}$ where l is the maximum log-likelihood, k is the number of degrees of freedom in the model and n_{obs} is the number of observations (i.e. years times ages). Models with a smaller BIC are considered to be better.

Table 2: Log-likelihoods and Bayesian Information Criteria Results

Model	Log-likelihood	Parameters	Constraints	Degrees of Freedom	BIC
M1	-16090.70	145	2	143	33285.18
M5	-28628.74	90	0	90	57952.16
M6	-13162.64	184	2	182	27730.07
M7	-12014.76	229	3	226	**25773.95**
CBDX1-PML	-17162.86	189	3 (*)	186	35761.40
CBDX1-MLE	-14903.29	289	3	186	31242.26
CBDX2-PML	-17443.27	234	5 (*)	229	36654.11
CBDX2-MLE	-12338.60	234	5	229	26444.78
CBDX3-PML	-12085.20	279	7 (*)	272	26269.88
CBDX3-MLE	-11919.35	279	7	272	**25938.18**

Note: Constraints marked with a (*) are implicit rather than explicitly applied in the model fitting process.

Based on the BIC results, M7 should be marginally preferred to CBDX3. We also note that M1 and M5 both perform relatively poorly, presumably because of their simple structure and their lack of a cohort effect. In addition, we see that there is a substantial difference between the CBDX2 PML and MLE variants.

Figure 3 presents Pearson standardised residuals for M7 and CBDX3. These are quite similar across the models. However, M7 residuals by age (top left panel) exhibit a modest wave pattern that CBDX3 eliminates through its use of its additional age effect, $\alpha(x)$. Otherwise, the plots in Figure 3 indicate that both models fit the data very well. But the top left panel does suggest that CBDX3 might be preferred over M7, notwithstanding the BIC scores.

The results reported above are based on sample age range of 40-89 and years 1971-2015. As a robustness check, Table 3 reports the BIC rankings for a variety of age ranges and years.

We see that models M7, CBDX2-MLE and CBDX3-MLE consistently perform well across different sample age and year ranges and generate the same average ranks.

We also see that the best performing models on the wider age ranges (see first two columns) are M7 and CBDX3-MLE: they both rank 1 or 2. From these results, there is nothing to choose between them. But then introduce the scatterplots in Figure 3 and the balance swings in favour of CBDX3-MLE. We would therefore argue that the best performing model on the wider age range is CBDX3-MLE.

On the other hand, if one wants to use a narrower (i.e., older) age range (see third and fourth columns), then the two equal best performers are M6 and CBDX2-MLE and there is nothing to choose between them either.

Finally, we also observe that the PML versions of the CBDX models always rank worse than their MLE counterparts, and the difference in ranking is especially marked for CBDX1 and CBDX2.

Figure 3: Scatterplots of Pearson Standardised Residuals for M7 and CBDX3 against Age (top), Calendar year (middle) and Year of Birth (bottom)

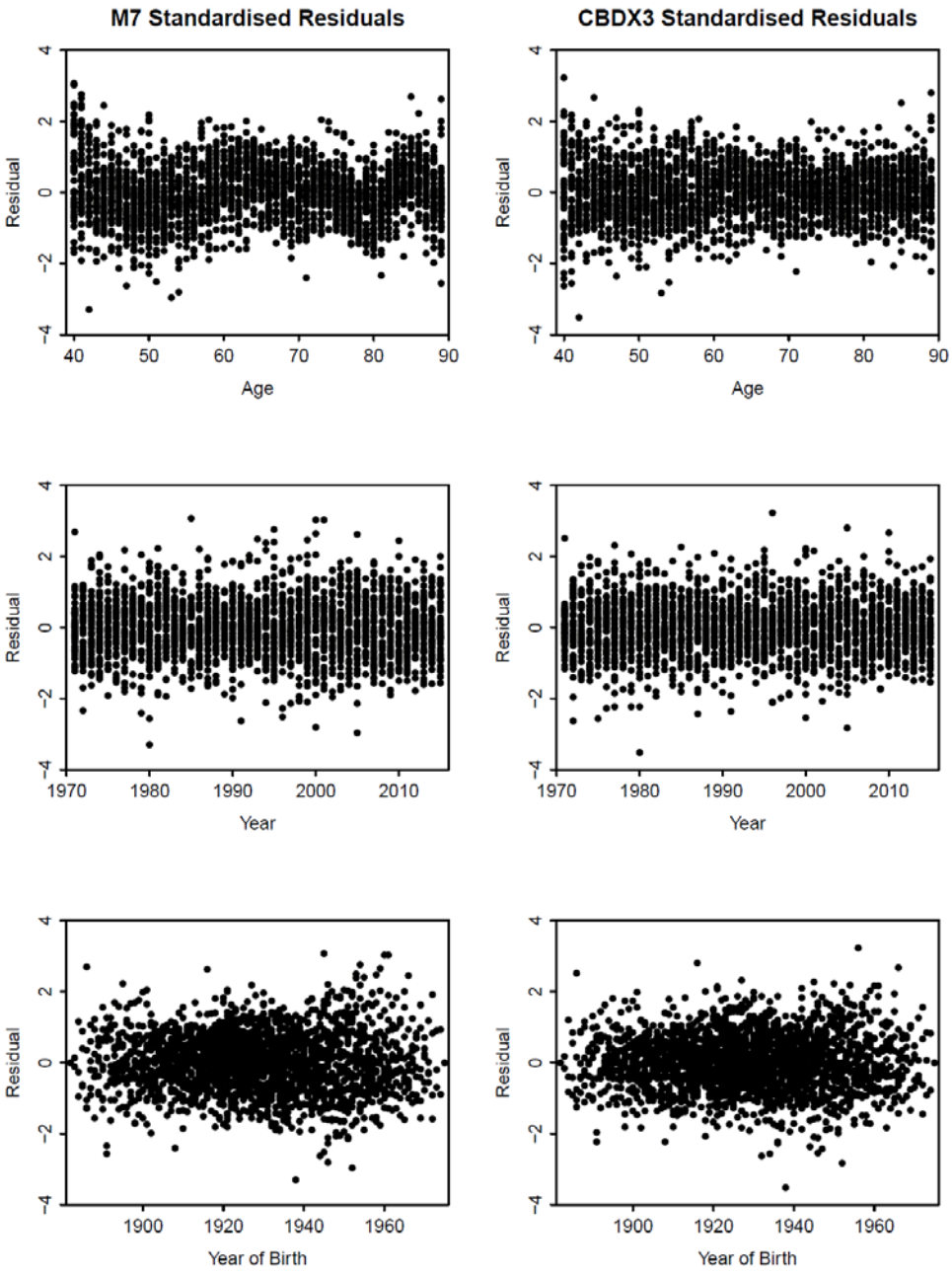


Table 3: BIC-based Model Rankings for Alternative Age Ranges and Ranges of Years

Model	Ranking				Average
Ages:	40-89	40-89	60-94	60-94	Rank
Years:	1971-2015	1961-2015	1971-2015	1961-2015	
M1	7	7	6	7	6.75
M5	10	10	9	9	9.5
M6	5	5	1	2	3.25
M7	1	2	3	4	2.5
CBDX1-PML	8	9	10	10	9.25
CBDX1-MLE	6	6	8	8	7
CBDX2-PML	9	8	7	6	7.5
CBDX2-MLE	4	3	2	1	2.5
CBDX3-PML	3	4	5	5	4.25
CBDX3-MLE	2	1	4	3	2.5

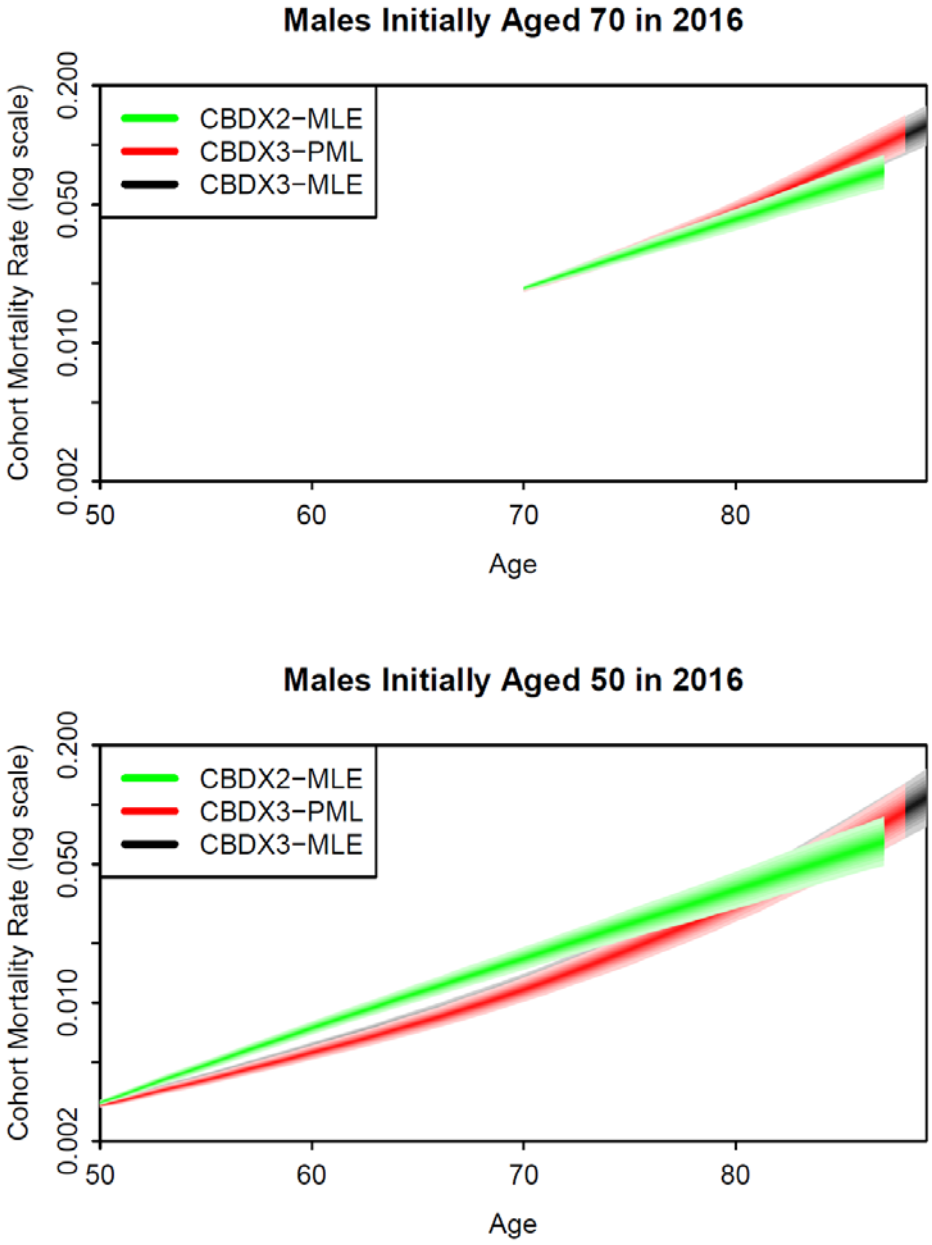
7. A Financial Application

We now consider a financial application of the models which focus on annuity prices for two cohorts of males: one aged exactly 70, and one aged 50 at the start of 2016.

Figure 4 shows fan charts (covering the 5% to the 95% quantiles) of the underlying death rates for the two cohorts up to age 89. In both plots, the two variants of CBDX3 give quite consistent results due to the similarity of the historical period and cohort effects. CBDX2 initially tracks CBDX3-MLE fairly closely in the upper plot, but then deviates quite significantly. In the lower plot, CBDX2-MLE follows quite a different path. There are three reasons for these differences. First, the single values of the cohort effects for these cohorts (born in 1946 and 1966 respectively) are quite different (Figure 1, bottom right panel). Second, the upwards drift in $\kappa_3(t)$ (Figure 1, bottom left panel; an effect not present in CBDX2) causes mortality rates at high ages to improve at a slower rate in CBDX3. Third, in Figure 4, we can see that fitted death rates in 2015 are higher for CBDX2-MLE than either CBDX3 over the ages 50-70.

Tables 4 and 5 show the prices, per £1 per annum, of term annuities payable to males aged exactly 70 and 50, respectively. For the 70-year-old, payments commence immediately and are payable until age 90, and for the 50-year-old, the annuity is deferred to age 65 and ceases at age 90. For the 70-year-old, we can see some variation in prices, with M6, CBDX1 and CBDX2 well above the other models. This can be attributed to the significant differences in the fitted cohort effect at age 70 (year of birth 1946 in Figure 1, bottom right panel). For the 50-year-old, price differences between models are reversed and again mainly reflect differences in the fitted cohort effects for year of birth 1966. These differences emphasize the importance of (a) getting a good estimate of the cohort effect, and (b) ensuring that the model has adequate age and period effects to avoid overfitting with the cohort effect. Figure 4 also helps to illustrate how these price differences emerge in relation to differences in the underlying mortality rates. For example, for the cohort aged 50, the higher mortality rates up to age 80 in the CBDX2 fan lead to lower survival probabilities and lower deferred annuity values.

Figure 4: Fan Charts for CBDX3-PML and -MLE and CBDX2-MLE Cohort q Rates



Notes: Based on E&W data for years 1971:2015 and ages 40:89.

Tables 4 and 5 also include in parentheses the additional capital required to cover runoff of the liability at a 95% level of confidence (additional 95% VaR).[6] Models M1 and CBDX1 have less uncertainty in simulated mortality rates at high ages than other models and this results in lower VaRs for these models. Other models have similar VaR levels. At age 70, VaRs are larger than the variation in prices between models. At age 50, VaRs are much higher as payments are more distant and more risky. We also see that variations in prices are larger than the VaRs.

[6] Simulations here do not include parameter uncertainty.

Table 4: Prices of a Term Annuity Function from Age 70

Model	Term Annuity Price (additional 95% VaR)
M1	12.63 (0.24)
M5	12.56 (0.29)
M6	12.89 (0.30)
M7	12.73 (0.32)
CBDX1-MLE	12.92 (0.24)
CBDX2-MLE	12.94 (0.29)
CBDX3-PML	12.68 (0.31)
CBDX3-MLE	12.70 (0.31)

Notes: Based on E&W data for years 1971:2015 and ages 40:89. The term annuity presumes a continuously compounding risk-free interest rate of 1.5% p.a.

Table 5: Prices of a Deferred Term Annuity Function from Age 50 with Payments Commencing at Age 65

Model	Deferred Term Annuity Price (additional 95% VaR)
M1	12.59 (0.40)
M5	12.44 (0.44)
M6	12.03 (0.44)
M7	12.52 (0.45)
CBDX1-MLE	11.74 (0.44)
CBDX2-MLE	11.83 (0.44)
CBDX3-PML	12.63 (0.44)
CBDX3-MLE	12.50 (0.44)

Notes: Based on E&W data for years 1971:2015 and ages 40:89. The deferred term annuity presumes a continuously compounding risk-free interest rate of 1.5% p.a.

8. Recalibrating Forecasts

In this section, we propose a novel method for forecasting that blends statistical analysis of the historical period effects with expert judgement about central forecasts.

Current practice in the UK life insurance industry following guidance from the UK's Prudential Regulatory Authority (PRA, 2016) can be summarised as follows. Establish a best estimate (BE) for future mortality using, e.g., expert judgement (EJ) or statistical modelling. As a separate exercise, identify distinct stochastic mortality models from four common families of model (e.g. CBD, Lee-Carter) calibrated using standard statistical methods for their volatilities. For each model, calculate best estimate annuity liabilities and the associated Solvency Capital Requirement (SCR) as a percentage of the same model's best estimate. Use expert judgement to establish a single percentage SCR based on results for the four models. Finally, apply this percentage SCR to the BE-EJ best estimate based on expert judgement. While this approach on the face of it seems reasonable, there is an uncomfortable disconnect between the model used to establish

the best estimate, and the models used to establish how much uncertainty there is around that best estimate. The approach we propose here resolves this issue.

Our new model for forecasting the period effects is generalised to a multivariate random walk with time dependent drifts:

$$(21a) \qquad K_1(t)=K_1(t-1)+\mu_1(t)+Z_1(t)$$
$$(21b) \qquad K_2(t)=K_2(t-1)+\mu_2(t)+Z_2(t)$$
$$(21c) \qquad K_3(t)=K_3(t-1)+\mu_3(t)+Z_3(t)$$

where $(Z_1(t), Z_2(t), Z_3(t))$ are i.i.d. multivariate normal vectors with mean 0 and $Cov(Z_i(t), Z_j(t)) = \sigma_{ij}$ are estimates from the historical observations of $(K_1(t), K_2(t), K_3(t))$.

The standard model has $\mu_j(t) = \mu_j$, i.e., constant drift, estimated from the historical data:

$$(22) \qquad \hat{\mu}_j=(K_j(t_1) - K_j(t_0)/(t_1 - t_0).$$

The proposed model assumes

$$(23) \qquad \mu_j(t) = \left(\mu_{j0} - \mu_{j1}\right)e^{-\beta(t-t_1)} + \mu_{j1}.$$

So for each $K_j(t)$, we require:

(a) $K_j(t_1)$= last observation (already known),
(b) μ_{j0}= short term drift,
(c) μ_{j1}= long term drift, and
(d) β= rate at which we move from μ_{j0} to μ_{j1}

with (b), (c), and (d) to be specified by the user. Other models for $\mu_j(t)$ are, of course, possible (e.g., as in CMI, 2018).

One procedure for calibration is: (1) choose μ_{j0} = an estimate of the *recent* trend in $K_j(t)$ up to t_1; (2) calibrate $(\mu_{11}, \mu_{21}, \mu_{31})$, so that the prices of 3 annuities for ages 60, 65, 70 match stated prices (best-estimate liabilities) P_{60}, P_{65}, P_{70} that might have been established using a given best estimate projection of mortality, e.g., using the CMI mortality projections model (CMI, 2018).[7]

By calibrating the μ_{jk}'s in this way, we have a single stochastic model that matches best estimate liabilities (e.g., as output from a different model such as CMI, 2018, calibrated using expert judgement) and gives us a consistent estimate of the uncertainty around that liability. Having a single model that delivers both characteristics within a coherent setting gives us more confidence that the value-at-risk associated with the best estimate is reliable.

[7] Note that CMI (2018) proposes that previously established cohort effects also decay gradually to zero in projections. We do not propose that here, but it could be incorporated into our model if desired or if necessary to get a satisfactory calibration.

By way of example, we considered two variants alongside the baseline documented in the previous section:

- Variant 1: μ_0= mean drift (vector) over the period 2011-2015; μ_1= long-term historical drift (1971-2015).
- Variant 2: μ_0= mean drift (vector) over the period 2011-2015; μ_1= (-0.015,0,0) (i.e., in the long run, the improvement rate will be 1.5% per annum at all ages from 50 to 89, consistent with recent actuarial practice).

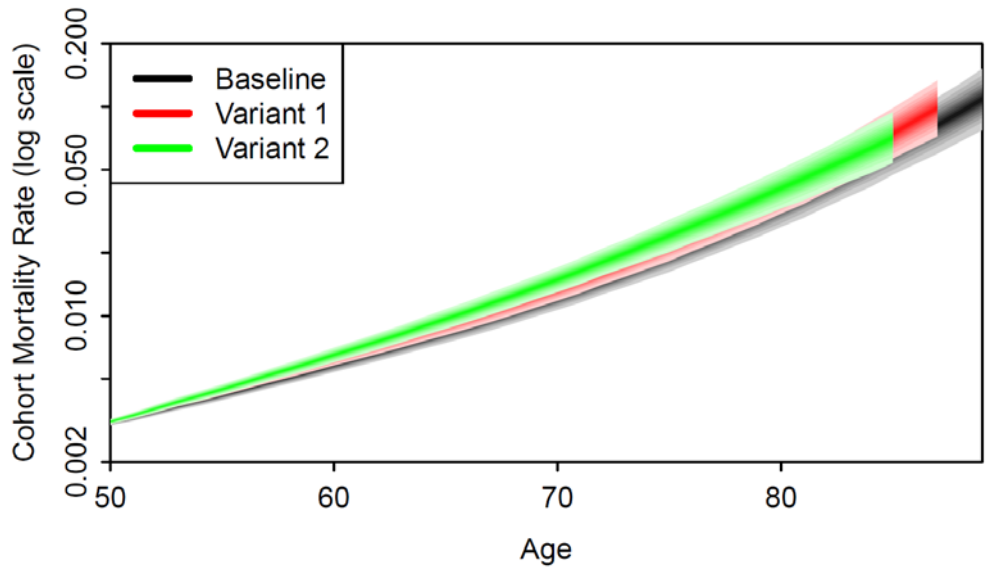
Table 6 gives prices and VaRs for the deferred annuity from age 50, and the immediate annuity from age 70, while Figure 5 shows fan charts for the cohort mortality rates from age 50 for the three variants.

Table 6: Deferred and Immediate Annuity Prices for Three Calibrations of the Time-Dependent Drift of the Multivariate Random Walk

	Age 50 Price (additional 95% VaR)	Age 70 Price (additional 95% VaR)
Baseline	12.50 (0.44)	12.70 (0.31)
Variant 1	12.13 (0.47)	12.42 (0.33)
Variant 2	11.95 (0.49)	12.41 (0.33)

Figure 5 helps to illustrate how differences in prices emerge under the three variants. Variant 1 deviates initially from the baseline because of the lower initial improvement rate, but, thereafter, tracks the trend of the baseline fan. The overall higher mortality then results in a lower price in Table 6. Variant 2 tracks Variant 1 initially, having the same initial rate of improvement, but then deviates from Variant 2 because of the different long-term improvement rates.

Figure 5: Cohort Mortality Fan Charts for Three Variants with Different Parameters for the Random Walk Drifts: Males Aged 50 in 2016



9. Features and Extensions of the CBDX Model

9.1 Features

The CBDX model has a number of useful features.

The first point to note, touched on a little earlier, is that CBDX3 provides a satisfactory γ process. Ideally, the γ plot should be zero-trended, devoid of any particular features (e.g., spikes) and 'small' in value (i.e., so we are not relying on the γ process, which is only a residual in the model, to do any heavy lifting). We see that the CBDX3 γ plot in Figure 1 fits these criteria well, in contrast to the γ plots provided by other models (apart from M7) or by other versions of CBDX. The fact that the γ process is a model residual also means that the γ plot gives us another goodness-of-fit indicator, since any misspecifications or other problems in the model would be carried through to the residual. Our approach to the treatment of γ thus allows for more helpful diagnostics of the model's goodness of fit.

The difference between the CBDX2 and CBDX3 γ plots provides a further confirmation of the importance of the third period effect. Hunt and Blake (2014) stress that, if a significant improvement in fit can be achieved through the addition of a further period effect, then this should be preferred over the immediate introduction of a cohort effect. A cohort effect should only be introduced when (a) residuals by year of birth in the age-period-only version of the model do not exhibit significant structure other than random waves, and (b) the addition of a cohort effect does not result in a material change in the estimated period effects compared to the age-period-only model (other than those due to identifiability constraints).

A second point is that our sequential iterative algorithm (PML) can provide another useful model-building diagnostic. Our fitting approach is to start by fitting the age-period version of the model (the $\alpha(x)$ and the $\kappa_i(t)$) and then the cohort effect $\gamma(c)$. For the full MLE, we then revise the fits iteratively, refitting each in turn, until we have an ML fit for the CBDX1/2/3 model. If the overall fit is not a good one (e.g., CBDX2 in Figure 1), then we will get a γ plot that is visibly not small, featureless or zero-trended, and that γ plot would be a signal that the model is inadequate in some way (e.g., in the sense of Hunt and Blake, 2014)

Third, our iterative fitting process allows us to fit age, period and cohort effects without explicitly imposing identifiability constraints. Not having to specify those constraints is helpful because there has been a growth in the number and importance of such conditions in recent mortality modelling literature, and this growth has led to both robustness problems and to difficulties in making projections of future mortality rates (see, e.g., Hunt and Blake, 2020a,b).

9.2 Extensions

We can consider a number of possible extensions:

- Apply the model to other countries or sub-populations to confirm that CBDX3 is a suitable workhorse model.

- Multi-population (e.g., gravity) modelling: the CBDX3 model can be used as the base model for a multi-population (or gravity-type) mortality model, which could then be used for hedging analysis (see, e.g., Dowd *et al*., 2011, 2018; Cairns *et al*., 2011a).

- Parameter uncertainty*:* A useful feature of the model is that, like its CBD predecessors, the simulation model can be easily adapted to allow for parameter uncertainty along the lines shown in Cairns *et al*. (2006, 695-698, 2011b) or Dowd *et al.* (2010 pp. 296-297).

10. Conclusions

The purpose of this paper is to identify a workhorse model suitable for modelling the mortality of the mature adult population without the need for the large number of age effects that are usually required to model the full human age range (i.e., infant mortality and the accident hump). We first used Hunt and Blake (2014)'s general procedure to identify an age-period model that fits the data well before adding in a cohort effect that then captures the residual year-of-birth effects. Based on a variety of diagnostics, our preferred model for wider age ranges is the CBDX model with three period effects, CBDX3.

Acknowledgements

The authors gratefully acknowledge funding from the Actuarial Research Centre of the Institute and Faculty of Actuaries through the "Modelling Measurement and Management of Longevity and Morbidity Risk" research programme.

References

Cairns, A.J.G. (2014) "Modeling and Management of Longevity Risk", in P. B. Hammond, R. Maurer, and O. S. Mitchell, eds., *Recreating Sustainable Retirement: Resilience, Solvency, and Tail Risk.* Oxford, U.K.: Oxford University Press, pp. 71-88.

Cairns, A.J.G, D. Blake, K. Dowd (2006) "A Two-Factor Model for Stochastic Mortality with Parameter Uncertainty: Theory and Calibration." *Journal of Risk and Insurance* 73: 687–718.

Cairns, A.J.G., D. Blake, K. Dowd, G. D. Coughlan, D. Epstein, A. Ong and I. Balevich (2009) "A Quantitative Comparison of Stochastic Mortality Models Using Data from England and Wales and the United States." *North American Actuarial Journal* 13(1): 1-35.

Cairns, A.J.G., D. Blake, K. Dowd, G. D. Coughlan, and M. Khalaf-Allah (2011a) "Bayesian Stochastic Mortality Modelling for Two Populations." *ASTIN Bulletin* 41(1): 29-59.

Cairns, A. J.G., D. Blake, K. Dowd, G. D. Coughlan, and M. Khalaf-Allah (2011b) "Mortality Mortality Density Forecasts: An Analysis of Six Stochastic Mortality Models." *Insurance: Mathematics and Economics* 48: 355-367.

Cairns, A. J.G., D. Blake, K. Dowd, and A.R. Kessler (2016) "Phantoms Never Die: Living with Unreliable Population Data." *Journal of the Royal Statistical Society A*, 179, Part 4: 975-1005.

CMI (2018) "CMI Mortality Projections Model: CMI_2017." CMI Working Paper 105, Institute and Faculty of Actuaries.

Dowd, K., D. Blake and A.J.G. Cairns (2018) "Hedging Annuity Risks with the Age-Period-Cohort Two-Population Gravity Model." *North American Actuarial Journal*. https://www.tandfonline.com/doi/full/10.1080/10920277.2019.1652102.

Dowd, K., A. J. G. Cairns, D. Blake, G. D. Coughlan, D. Epstein and M. Khalaf-Allah (2010) "Backtesting Stochastic Mortality Models: An Ex-Post Evaluation of Multi-Period-Ahead Density Forecasts." *North American Actuarial Journal* 14(4): 281-298.

Dowd, K., A. J.G. Cairns, D. Blake, G. D. Coughlan, D. Epstein, and M. Khalaf-Allah (2011) "A Gravity Model of Mortality Rates for Two Related Population." *North American Actuarial Journal* 15(2): 334-356.

Hunt, A., and D. Blake (2014) "A General Procedure for Constructing Mortality Models", *North American Actuarial Journal* 18: 116-138.

Hunt, A., and D. Blake (2020a) "Identifiability in Age/Period Mortality Models." *Annals of Actuarial Science*, forthcoming.

Hunt, A., and D. Blake (2020b) "Identifiability in Age/Period/Cohort Mortality Models." *Annals of Actuarial Science*, forthcoming.

Lee, R. D., and L. R. Carter (1992) "Modeling and Forecasting U.S. Mortality." *Journal of the American Statistical Association* 87: 659–75.

Plat, R. (2009) "On Stochastic Mortality Modeling." *Insurance: Mathematics and Economics* 45: 393-404.

PRA (2016) "Solvency II: consolidation of Directors' letters." Consultation Paper CP20/16, Prudential Regulation Authority, Bank of England.

Renshaw, A.E., and S. Haberman (2003) "Lee-Carter Mortality Forecasting with Age-Specific Enhancement." *Insurance: Mathematics and Economics* 33: 255-272.

Renshaw, A.E., and S. Haberman (2006) "A Cohort-Based Extension to the Lee Carter Model for Mortality Reduction Factors." *Insurance: Mathematics and Economics* 38: 556-70.