Institute and Faculty of Actuaries

Semi-parametric Extensions of the Cairns-Blake-Dowd Model: A One-dimensional Kernel Smoothing Approach

Han Li

Colin O'Hare

16 August 2016

Institute and Faculty of Actuaries

Reasons why mortality modeling is IMPORTANT:

- During the past two decades: life expectancy - improving at approximately 3 years per decade.
- Mortality and longevity risk: significant risks faced by governments, insurance companies, pension providers and individuals.
- Accurate mortality forecast is of fundamental importance.

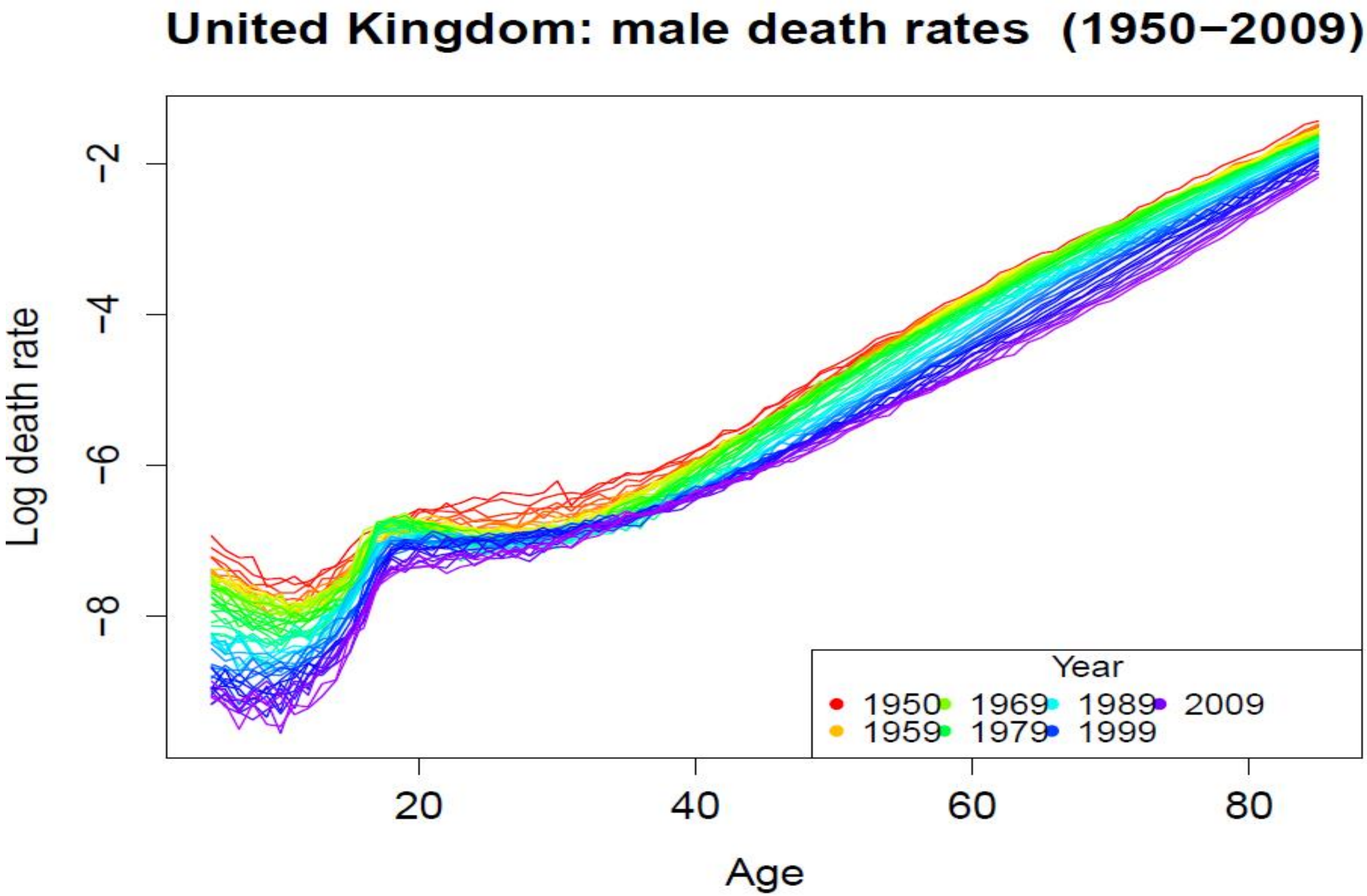
Institute and Faculty of Actuaries

Example: Adequate pricing of life annuities relies on the accuracy of future mortality projection.

- Quote from Sense and Sensibility (1811):

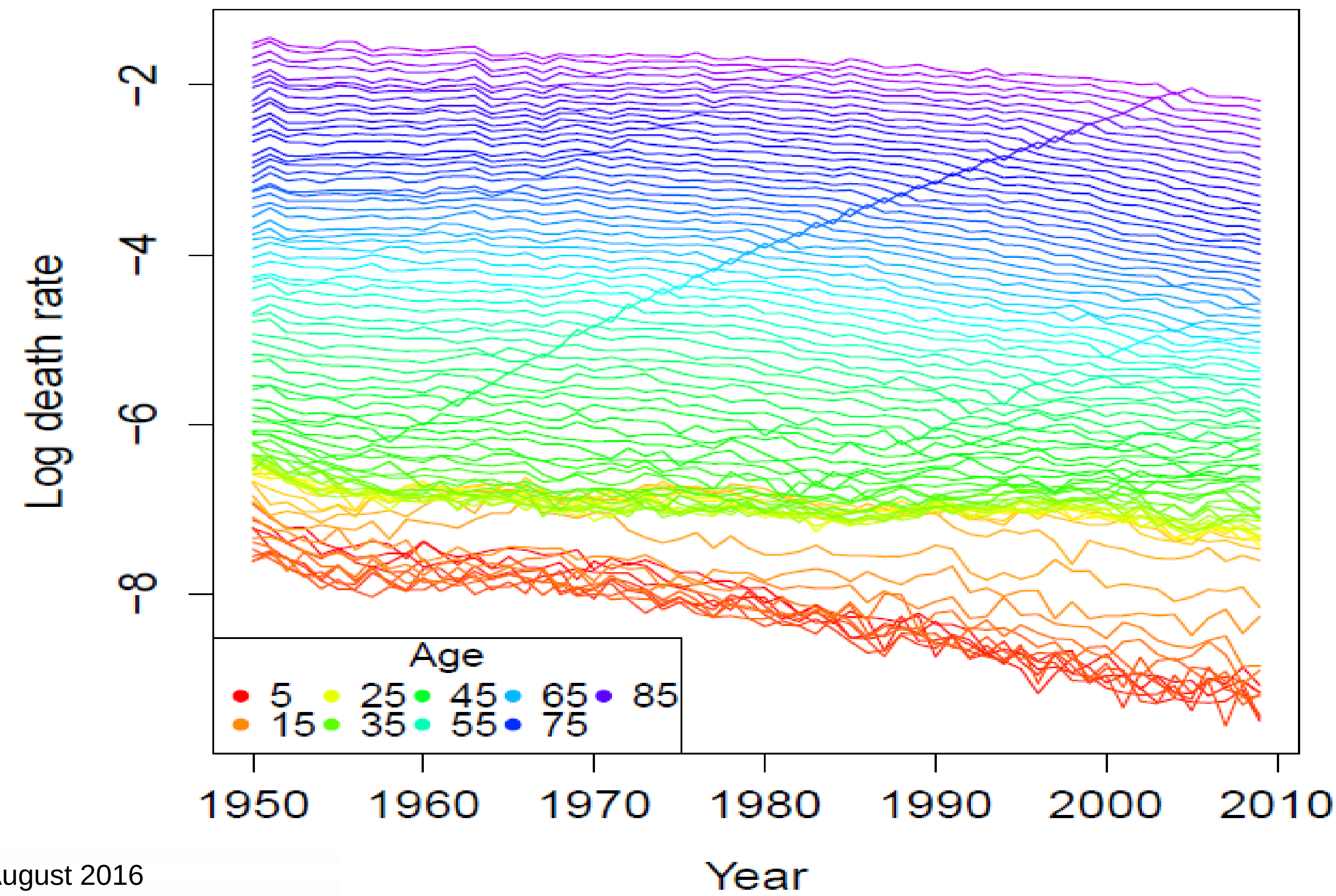
  "If you observe, people always live forever when there is an annuity to be paid them. An annuity is a very serious business; it comes over and over every year, and there is no getting rid of it. You are not aware of what you are doing. I have known a great deal of the trouble of annuities..."

- Unanticipated improvements in longevity have caused life offices and pension plan sponsors to incur losses on life annuity business as they are paying out for MUCH longer than was anticipated.


United Kingdom: male death rates (1950−2009)
Log death rate
-2
-4
-6
-8
20
40
60
80
Age
Year
1950
1959
1969
1979
1989
1999
2009

Institute and Faculty of Actuaries

United Kingdom: male death rates (1950−2009)



Institute and Faculty of Actuaries

Outline of the talk:

1. Literature review
2. A one-dimensional kernel smoothing approach
3. Empirical results and analysis
    - Data
    - Fitting Performance
    - Residual Checks
    - Forecasting Performance
4. Conclusions

Institute and Faculty of Actuaries

Existing mortality models

- Lee-Carter model (1992):

$$\log(m_{x,t}) = a_x + b_x\kappa_t \tag{1}$$

- Cairns-Blake-Dowd model (2006):

$$\text{logit}(q_{x,t}) = \kappa_t^1 + \kappa_t^2(x - \bar{x}) \tag{2}$$

- Generalizations of the CBD model:

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

$$\text{M7: } \text{logit}(q_{x,t}) = \kappa_t^1 + \kappa_t^2(x - \bar{x}) + \kappa_t^3[(x - \bar{x})^2 - \hat{\sigma}_x^2] + \gamma_{t-x}, \tag{4}$$

- Plat model (2009):

$$\log(m_{x,t}) = a_x + \kappa_t^1 + \kappa_t^2(\bar{x} - x) + \kappa_t^3(\bar{x} - x)^+ + \gamma_{t-x} \tag{5}$$

Motivations:

- Gains from the quadratic age-time effect:
  1. A precise mortality model: "clean" residual plots
  2. The quadratic age-time effect in M7 has significantly increased the randomness in residual plots (Cairns *et al.*, 2009).
  3. We decide to follow this path and further extend the CBD model by adding higher order age-time effects into the model.
- A semi-parametric panel approach to mortality modeling:
  1. Poisson assumption on number of deaths may not be valid.
  2. Li *et al.* (2015) proposed a local linear kernel smoothing (LLKS) approach: no assumption on number of deaths and better forecasting results.
  3. We decide to use the LLKS method to calibrate the proposed model and give local information more weights in the forecasting process.

Institute and Faculty of Actuaries

A time-varying coefficient mortality model

The time-varying coefficient (TVC) model is given as:

$$\text{logit}(q_{x,t}) = \sum_{i=1}^{r} \kappa_t^i [(x - \bar{x})^{i-1} - \sigma_x^{i-1}] \qquad (6)$$

where σ_x^n is the mean of $(x - \bar{x})^n$ and we define σ_x^0 to be 0. r is a non-negative integer number.

Institute and Faculty of Actuaries

LLKS Approach - intuition

- Having re expressed our mortality model as a panel model $Y_{it} = X_i'\beta_t$
- Assume that β_t is a linear function of time within a neighbourhood of t
- For each t estimate β_t by fitting a straight line based on local information
- The amount of local information to use is determined by the bandwidth h and the kernel smoothing function K
- Details in Li *et al (2015)*

A reminder of Li et al (2015) re-expressing the CBD model

In the Li *et al.*'s (2015) study, for $x \in [a+1, a+N]$ and $t \in [1, T]$, we re expressed the CBD model as a semi-parametric time-varying coefficient model in the following form:

- $Y_{it} = \text{logit}(q_{x,t})$, where $i = x - a$ and a is a non-negative integer.
- $X_i = \begin{pmatrix} 1 \\ x - \bar{x} \end{pmatrix}$.
- $\beta_t = \begin{pmatrix} \kappa_t^1 \\ \kappa_t^2 \end{pmatrix}$, where $\{\kappa_t^1, \kappa_t^2\}$ were smooth functions of time.
- The model can then be re-expressed as:

$$Y_{it} = \text{logit}(q_{x,t}) = X_i' \beta_t. \quad (7)$$

Following Li *et al.*'s (2015) study, for $x \in [a+1, a+N]$ and $t \in [1, T]$, we define:

- $Y_{it} = \text{logit}(q_{x,t})$, where $i = x - a$ and a is a non-negative integer.
- $X_i = \begin{pmatrix} 1 \\ x - \bar{x} \\ \vdots \\ (x - \bar{x})^{r-1} - \sigma_x^{r-1} \end{pmatrix}$.
- $\beta_t = \begin{pmatrix} \kappa_t^1 \\ \kappa_t^2 \\ \vdots \\ \kappa_t^r \end{pmatrix}$, where $\{\kappa_t^1, \kappa_t^2, \ldots, \kappa_t^r\}$ are smooth functions of time.

Institute and Faculty of Actuaries

The model can then be re-expressed as:

$$Y_{it} = \text{logit}(q_{x,t}) = \sum_{i=1}^{r} \kappa_t^i [(x - \bar{x})^{i-1} - \sigma_x^{i-1}] = X_i' \beta_t. \quad (8)$$

For $t \in [1, T]$, we define $\beta_t = \beta(\tau)$, where $\tau = t/T$. Thus the model can be approximated using results from Taylor expansion, for any given $\tau_0 \in [0, 1]$, we have:

$$Y_{it} = X_i' \beta(\tau) \approx X_i' [\beta(\tau_0) + \beta^{(1)}(\tau_0)(\tau - \tau_0)], \quad (9)$$

where $\beta^{(1)}(\tau_0)$ is the first order derivative of $\beta(\tau_0)$.

The local linear estimator of $\beta(\tau_0)$ can be obtained by minimizing the following weighted sum of squares with respect to $\left(\beta(\tau_0), \beta^{(1)}(\tau_0)\right)$:

$$\sum_{i=1}^{N}\sum_{t=1}^{T}\left\{Y_{it} - X_i'[\beta(\tau_0) + \beta^{(1)}(\tau_0)(\tau - \tau_0)]\right\}^2 K_h(\tau - \tau_0), \quad (10)$$

where $K_h(u) = h^{-1}K(u/h)$. h controls the amount of smoothing. We use "leave-one-out" cross-validation to select h and adopt the Epanechnikov kernel function as follows:

$$K(u) = 0.75(1 - u^2)I(|u| \leq 1). \quad (11)$$

- Model selection: Based on out-of-sample forecasting performance.

$$r_{\text{opt}} = \arg\min_{r} \frac{1}{Nn} \sum_{i=1}^{N} \sum_{t=T-n+1}^{T} (Y_{it} - X_i' \hat{\beta}_t)^2. \qquad (12)$$

- Different countries: different choice of r.
- Trade-off between bias and variance: needs to be considered.

Institute and Faculty of Actuaries

Data

The deaths and exposures data used to calculate central mortality rates are downloaded from the Human Mortality Database (HMD).

- Range of countries: Great Britain (GB), the United States (US), Australia (AUS), Netherlands (NL), Japan (JAP), France (FR) and Spain (SP).
- Investigation Period: 1950-2009 (post-war).
- Age range: 50-89 (older age).

Institute and Faculty of Actuaries

Statistical measures of performance

Define the following notation of statistical measures:

1. The average error:

$$E1 = \frac{1}{NT}\sum_x \sum_t \frac{\hat{m}_{x,t} - m_{x,t}}{m_{x,t}}. \tag{13}$$

2. The absolute average error:

$$E2 = \frac{1}{NT}\sum_x \sum_t \frac{|\hat{m}_{x,t} - m_{x,t}|}{m_{x,t}}. \tag{14}$$

3. The standard deviation of error:

$$E3 = \sqrt{\frac{1}{NT}\sum_x \sum_t (\frac{\hat{m}_{x,t} - m_{x,t}}{m_{x,t}})^2}. \tag{15}$$

Institute and Faculty of Actuaries

Fitting performance

		TVC			CBD: LLKS			M7		
	r	$E1$	$E2$	$E3$	$E1$	$E2$	$E3$	$E1$	$E2$	$E3$
GB	4	-0.20	2.45	3.18	-0.18	4.24	5.28	3.66	4.03	5.45
US	4	0.02	1.85	2.39	0.11	3.33	4.45	3.42	4.64	6.16
AUS	3	0.07	3.77	4.93	0.13	4.71	6.08	2.11	4.50	5.24
NL	5	-0.05	2.63	3.38	0.07	4.29	5.34	3.26	4.14	5.36
JAP	3	0.12	3.28	4.13	0.26	4.88	6.20	3.41	3.86	5.38
FR	6	0.01	2.42	3.25	0.28	6.53	8.25	3.41	3.82	5.29
SP	5	0.06	3.58	5.41	0.17	5.17	6.87	3.37	4.42	6.30
SWIT	4	-0.04	3.99	5.07	0.03	4.98	6.39	3.40	5.10	6.53
SWE	5	0.03	3.23	4.18	0.10	4.09	5.27	3.23	4.67	5.94
PORT	5	-0.02	4.49	5.94	0.19	6.43	8.21	3.83	5.22	6.98

Table: Fitting results (%) for male mortality rates from 1950-2011, ages 50-89.

Institute
and Faculty
of Actuaries

Residual plots

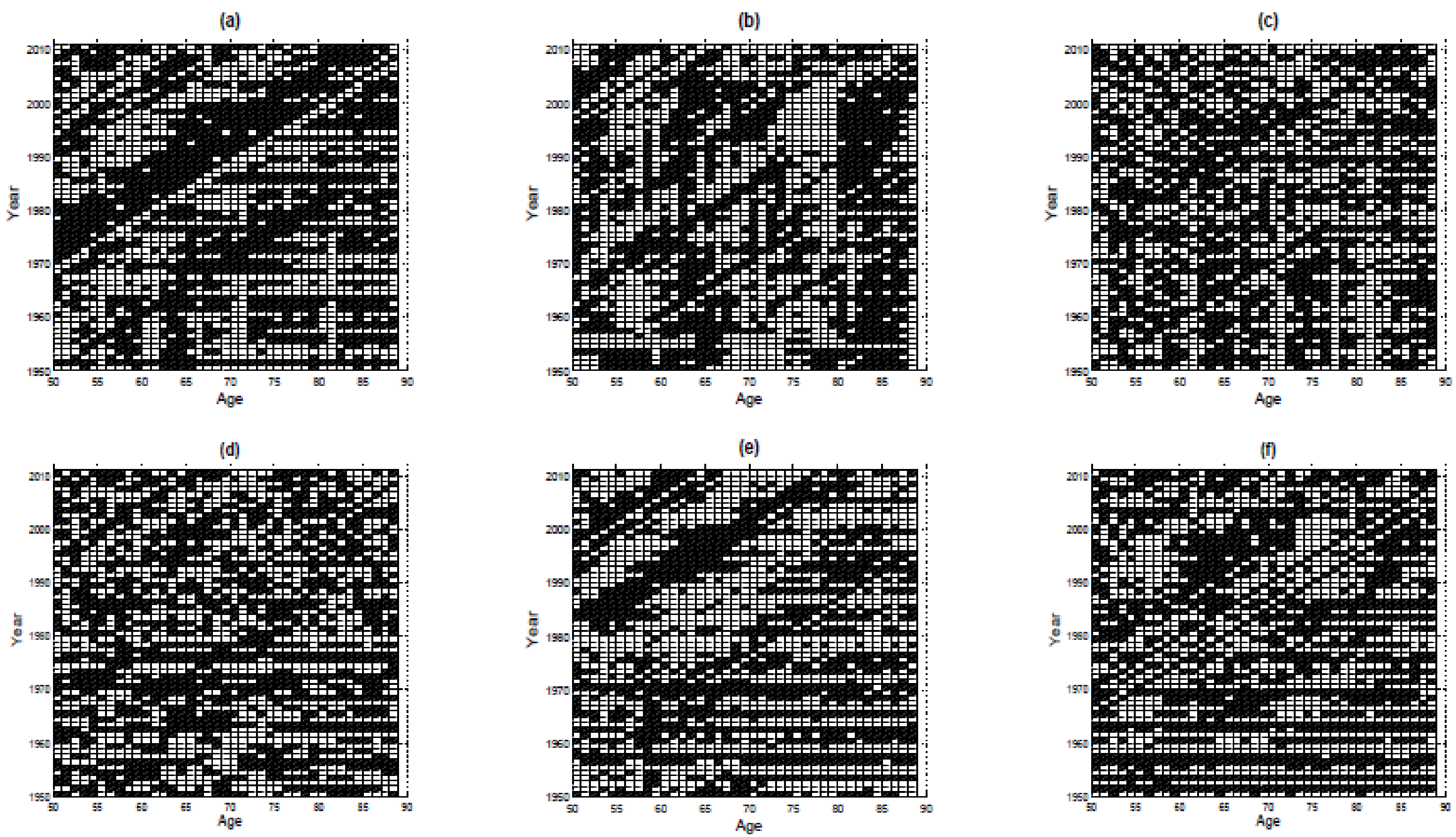
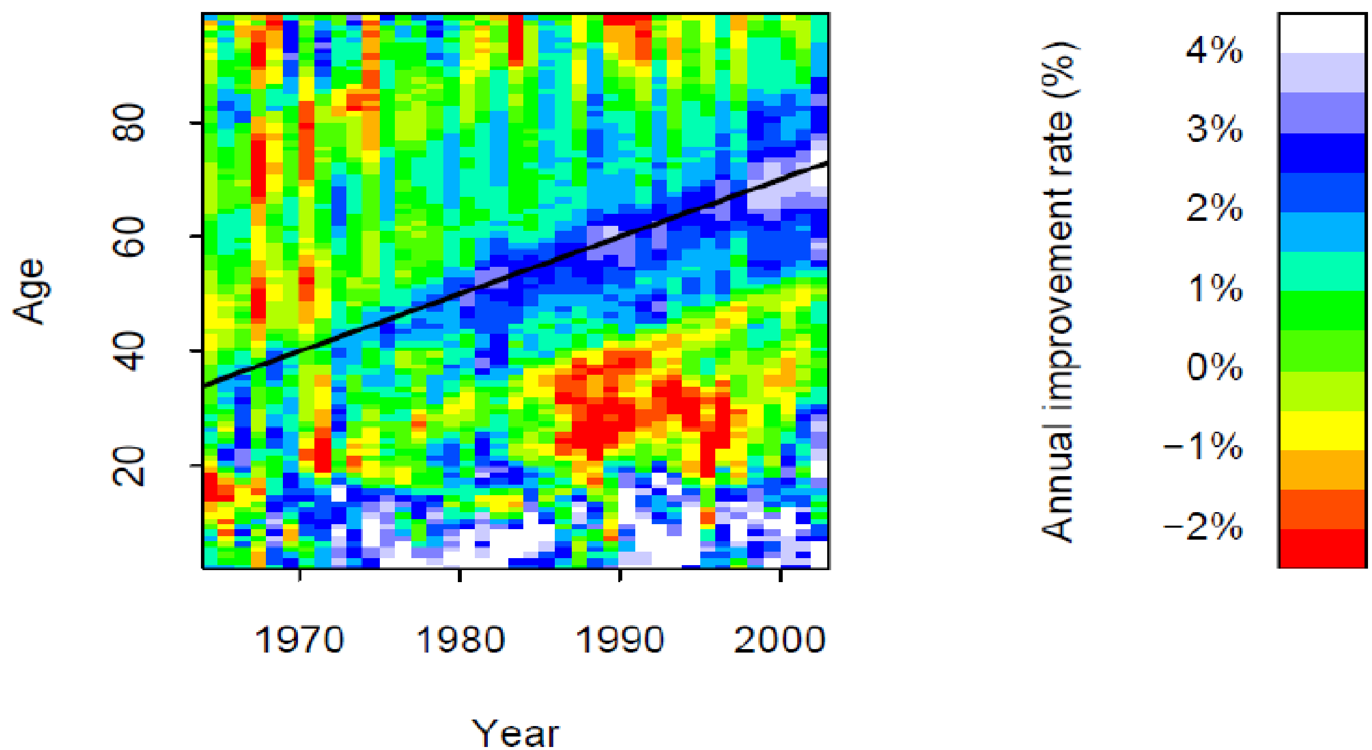


Figure: Residual plots of the 2-D LOP models for (a) GB, (b) US, (c) AUS, (d) NL, (e) JAP, (f) FR

Recall: cohort effect

Mortality improvement rates for England & Wales.



Generations born between 1925 and 1945 have experienced a more rapid improvement in mortality rates compared to other generations.

Institute and Faculty of Actuaries

Relation of age, time and cohort

In the literature of mortality modeling, age and time effects have been identified as the most important factors that would affect mortality rates.

The relationship between age, time and cohort:

$$i = t - x \tag{16}$$

where i represents cohort group.

Among the three variables, only two of them can be controlled at the same time as clearly cohort can be expressed as a function of age and time. Therefore, technically speaking, the interactions of t and x should be able to capture any underlying patterns in the mortality surface.

Institute and Faculty of Actuaries

Question: So why are there still diagonal patterns in some of the residual plots?

- Any non-differentiable part of the mortality surface would not be adequately captured by the model since systematically higher or lower mortality rates are outside the domain of smooth functions.
- Cohort effects: dependency of mortality experience for a group of individuals born in the same year.
- When a historical event significantly affected the mortality experience of certain generations, these patterns will continue in the future if there is strong a dependency on mortality experience within each cohort group.

Institute and Faculty of Actuaries

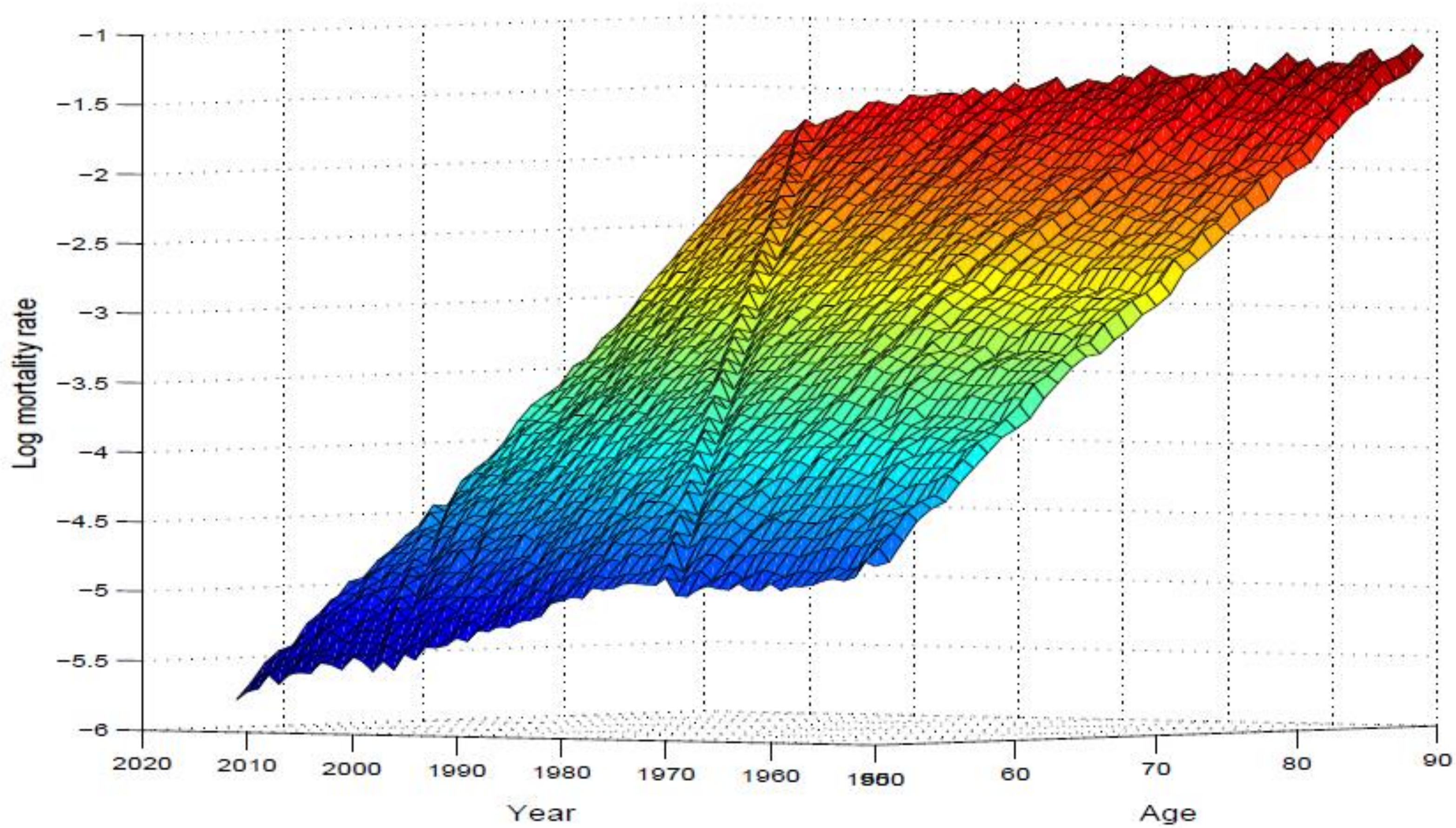


Figure: Log mortality rates for GB from 1950–2009, ages 50–89.

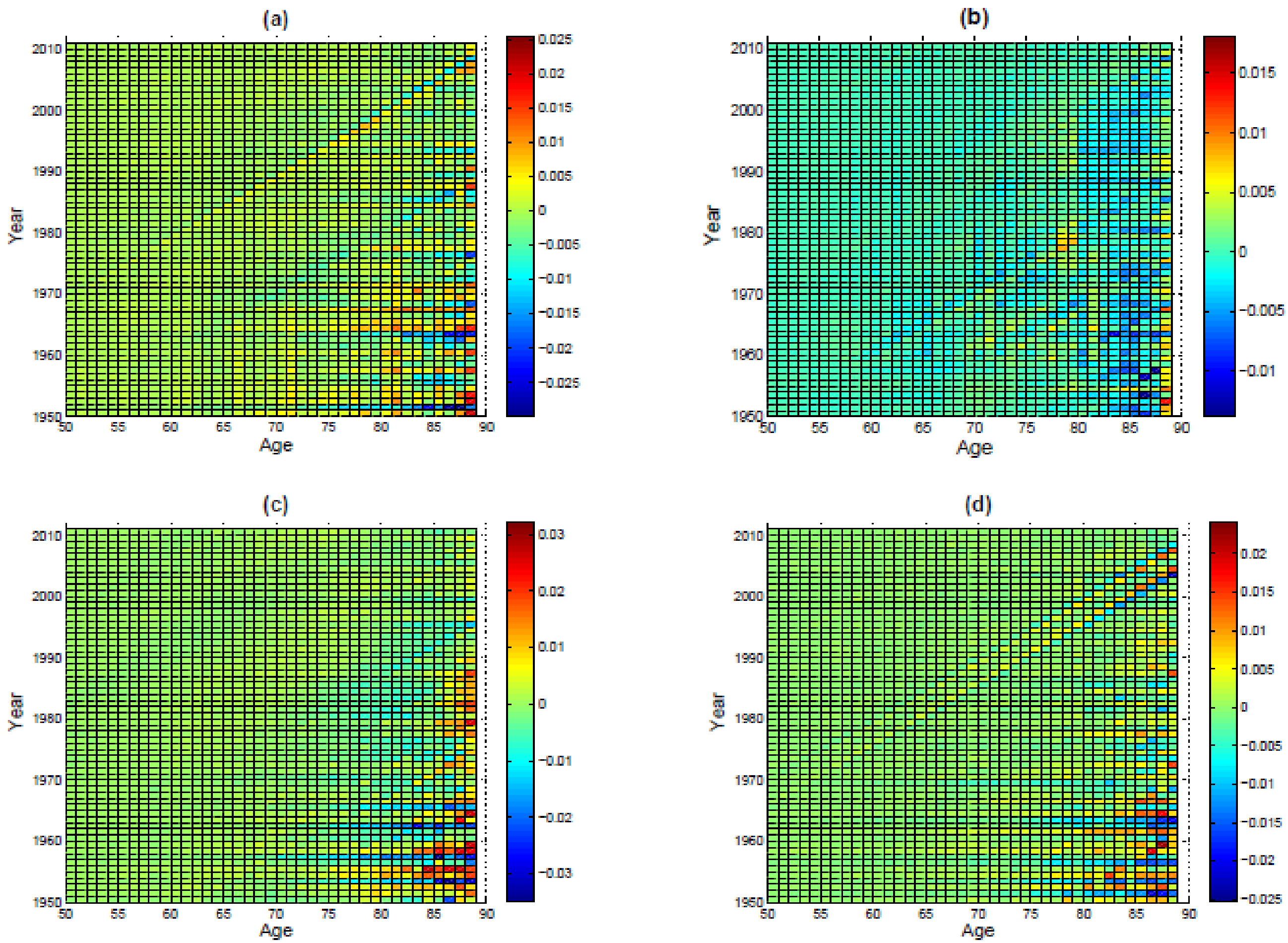


Figure: Residual plots of the TVC model for (a) GB, (b) US, and (d) FR.

Institute and Faculty of Actuaries

Forecasting performance

	TVC			CBD: LLKS			M7		
	$E1$	$E2$	$E3$	$E1$	$E2$	$E3$	$E1$	$E2$	$E3$
GB	-0.76	2.60	3.47	-0.49	5.23	6.10	-1.49	3.16	3.82
US	-1.30	2.76	3.50	-0.91	7.28	8.53	-6.59	8.22	10.86
AUS	-5.27	5.48	6.24	-4.84	7.95	10.00	-5.22	5.57	7.33
NL	-3.21	4.49	5.91	-3.05	5.92	7.43	1.80	4.03	5.22
JAP	0.29	3.52	4.46	0.62	7.55	8.76	1.96	6.04	7.07
FR	-2.14	3.94	5.12	-1.46	11.72	13.49	-3.28	5.32	7.26
SP	1.55	2.95	3.84	2.03	8.43	9.79	2.73	5.79	6.81
SWIT	-2.37	5.14	6.25	-1.90	8.56	10.40	-5.77	7.41	10.34
SWE	1.06	3.73	5.30	1.38	7.37	8.83	2.76	5.69	7.17
PORT	-0.08	3.66	4.86	0.94	11.92	13.37	11.00	15.20	19.19

Table: The 5-year-ahead male mortality forecasting results for ages 50–89, from 2007–2011.

Conclusions

- We apply a one-dimensional kernel smoothing approach via the introduction of TVC mortality models: data-driven models and allow us to have specific model design for each country's mortality experience.

- We argue that underlying mortality patterns can be sufficiently captured and the level of randomness in the residual plots is for satisfactory.

- The proposed TVC model fits historical mortality data well and achieves much better forecasting performance.
- The semi-parametric approach to mortality modeling is very attractive.

Questions?

Comments?

Thank you!

Institute and Faculty of Actuaries

Institute and Faculty of Actuaries

Research Gate page:

http://www.researchgate.net/profile/Han_Li51

- Li, H., O'Hare, C., Zhang, X., 2015a. A semiparametric panel approach to mortality modeling. Insurance: Mathematics and Economics 61, 264–270.
- Li, H., O'Hare, C., 2015b. Mortality forecast: Global or Local? Presented to the Actuaries Institute ASTIN, AFIR/ERM and IACA Colloquia, 2015, Sydney.
- Li, H., O'Hare, C., Vahid, F., 2016a. A flexible functional form approach to mortality modeling. Forthcoming in Journal of Forecasting. DOI: 10.1002/for.2437.
- Li, H., O'Hare, C., Vahid, F., 2016b. Two-dimensional kernel smoothing of mortality surface: an evaluation of cohort strength. Forthcoming in Journal of Forecasting. DOI: 10.1002/for.2399.