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

Han Li* and Colin O'Hare†

*ARC Centre of Excellence in Population Ageing Research (CEPAR), UNSW Sydney
†Department of Econometrics and Business Statistics, Monash University

13th International Longevity Risk and Capital Markets Solutions Conference

21–22 September 2017

Background

Reasons why mortality modeling is IMPORTANT:

Background

Reasons why mortality modeling is IMPORTANT:

- During the past two decades: life expectancy - improving at approximately 3 years per decade.

Background

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.

Background

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.

Background

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..."

Background

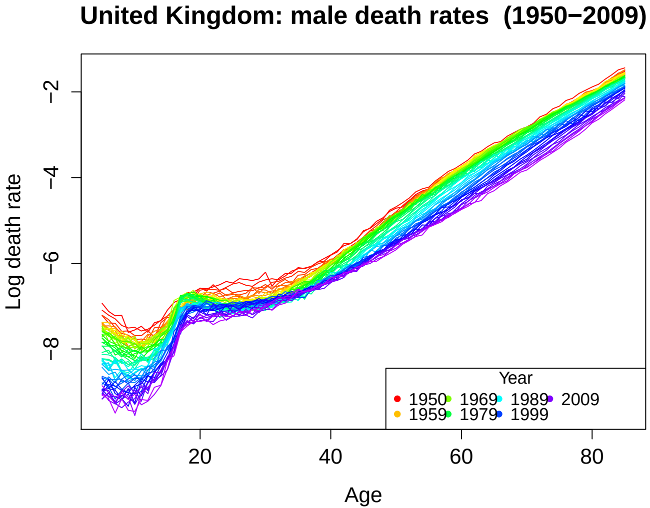
Example: Adequate pricing of life annuities relies on the accuracy of future mortality projection.

- 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..."

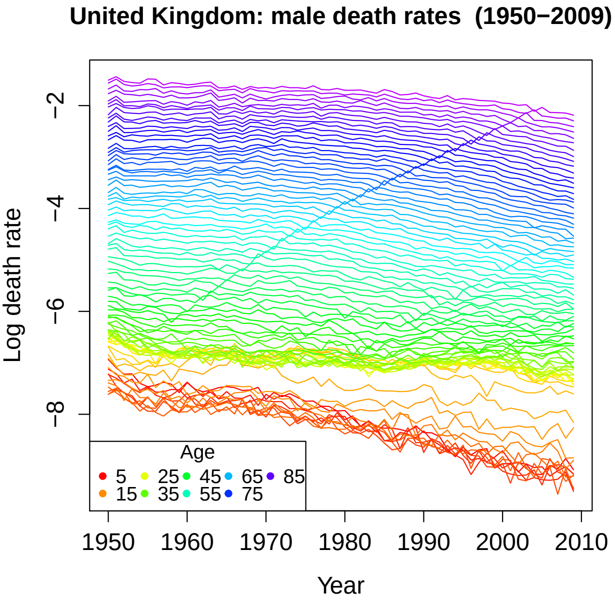
- Unanticipated improvements in longevity have caused life offices and pension plan sponsors to incur losses on life annuity business as they are paying out for MUCH longer than was anticipated.

Background



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

Background



United Kingdom: male death rates (1950−2009)
Log death rate
−2
−4
−6
−8
Age
5
15
25
35
45
55
65
75
85
1950
1960
1970
1980
1990
2000
2010
Year

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

Existing mortality models

- Lee-Carter model (1992):

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

- Cairns-Blake-Dowd model (2006):

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

- Generalizations of the CBD model (2009):

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

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

- Hyndman and Ullah model (2007):

$$\log(m_{x,t}) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) \quad (5)$$

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. Li *et al.* (2015) proposed a local linear kernel smoothing (LLKS) approach: simultaneously estimate and project parameters in the model and achieve improved forecasting results.
  2. We decide to use the LLKS method to calibrate the proposed model and give local information more weights in the forecasting process.

A time-varying coefficient mortality model

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

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

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.

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:

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

A time-varying coefficient mortality model

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.

A time-varying coefficient mortality model

The model can then be re-expressed as:

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

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:

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

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

A time-varying coefficient mortality model

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

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

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:

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

ARC CENTRE OF EXCELLENCE IN POPULATION AGEING RESEARCH

A time-varying coefficient mortality model

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

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

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

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 ages).

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:

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

Fitting performance

	TVC				CBD: LLKS			CBD(r):MLE			HU(r)		
	r	$E1$	$E2$	$E3$	$E1$	$E2$	$E3$	$E1$	$E2$	$E3$	$E1$	$E2$	$E3$
GB	4	-0.20	2.45	3.18	-0.18	4.24	5.28	0.98	3.98	5.39	**0.04**	**1.82**	**2.49**
US	4	**0.02**	1.85	2.39	0.11	3.33	4.45	-0.31	3.25	4.41	0.04	**1.43**	**1.87**
AUS	3	**0.07**	3.77	4.93	0.13	4.71	6.08	0.47	4.43	5.95	0.10	**2.99**	**3.89**
NL	5	-0.05	2.63	3.38	0.07	4.29	5.34	3.86	4.12	5.39	**0.04**	**2.15**	**2.81**
JAP	3	0.12	3.28	4.13	0.26	4.88	6.20	-0.42	4.41	5.93	**0.07**	**2.51**	**3.28**
FR	6	**0.01**	2.42	3.25	0.28	6.53	8.25	-0.98	6.33	8.09	0.06	**1.76**	**2.55**
SP	5	0.06	3.58	5.41	0.17	5.17	6.87	-0.50	4.63	6.32	**0.04**	**2.06**	**2.76**

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

Residual plots

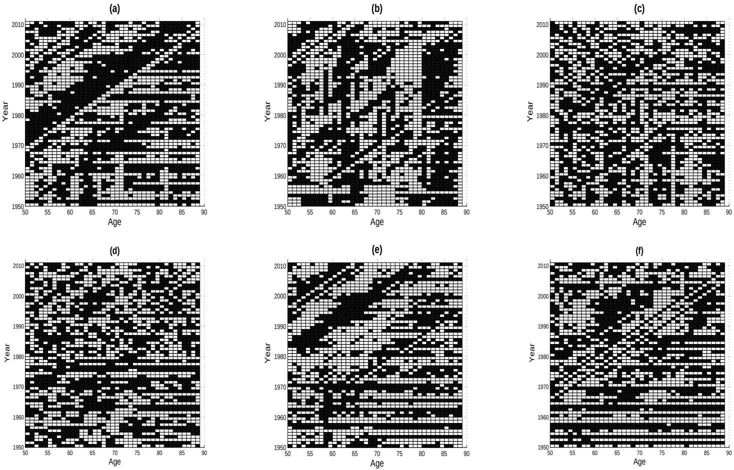
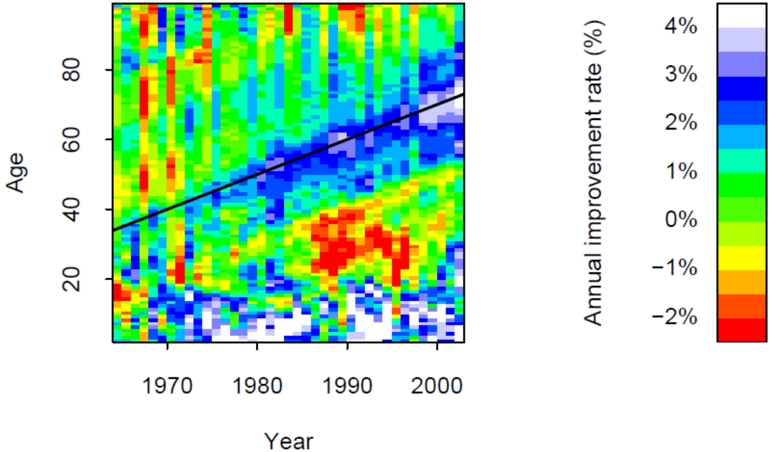


Figure: Residual plots of the 2-D LOP models for (a) GB, (b) US, (c) AUS, (d) NL, (e) JAP, and (f) FR

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.

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.

Relation of age, time and cohort

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.

GB male mortality surface

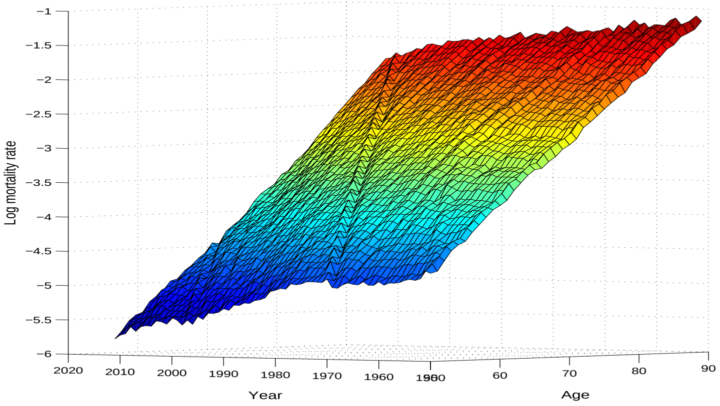


Figure: Log mortality rates for GB from 1950–2009, ages 50–89.

Possible reasons for the diagonal patterns: survivors of World War II period having higher mortality rates relative to other generations!

Residual plots

However, the magnitude of these diagonal patterns is negligible!

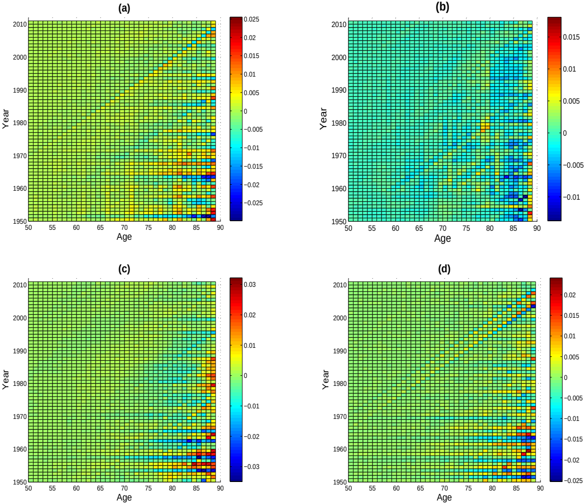


Figure: Residual plots of the TVC model for (a) GB, (b) US, (c) JAP and (d) FR.

Forecasting performance

	TVC				CBD: LLKS			CBD(r): MLE			HU(r)		
	r	$E1$	$E2$	$E3$	$E1$	$E2$	$E3$	$E1$	$E2$	$E3$	$E1$	$E2$	$E3$
GB	4	-0.76	**2.60**	**3.47**	**-0.49**	5.23	6.10	2.99	6.60	8.05	1.03	4.98	5.89
US	4	-1.30	**2.76**	**3.50**	-0.91	7.28	8.53	**0.85**	8.28	9.45	1.32	4.31	5.74
AUS	3	-5.27	5.48	6.24	-4.84	7.95	10.00	**-1.93**	8.17	10.44	-2.86	**5.40**	**7.66**
NL	5	-3.21	**4.49**	**5.91**	-3.05	5.92	7.43	6.14	8.24	10.13	**1.53**	5.63	6.69
JAP	3	**0.29**	3.52	4.46	0.62	7.55	8.76	-3.06	8.15	9.17	0.62	**3.00**	**3.87**
FR	6	-2.14	**3.94**	5.12	-1.46	11.72	13.49	**-0.13**	12.43	14.10	2.96	4.20	**5.10**
SP	5	1.55	**2.95**	**3.84**	2.03	8.43	9.79	**-0.22**	8.65	9.95	6.13	6.29	7.35
SWIT	4	-2.37	**5.14**	**6.25**	-1.90	8.56	10.40	**-0.10**	9.41	11.40	0.69	5.42	6.62
SWE	5	**1.06**	**3.73**	**5.30**	1.38	7.37	8.83	2.33	8.11	9.90	2.25	4.65	5.71
PORT	5	**-0.08**	**3.66**	**4.86**	0.94	11.92	13.37	0.58	11.70	14.06	4.89	7.16	8.89

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

Forecasting performance

		TVC			CBD: LLKS			CBD(r): MLE			HU(r)		
	r	*E*1	*E*2	*E*3	*E*1	*E*2	*E*3	*E*1	*E*2	*E*3	*E*1	*E*2	*E*3
GB	3	**6.14**	**8.57**	**10.95**	6.15	8.64	11.05	16.76	16.83	20.40	7.70	10.74	13.82
US	2	**1.43**	10.73	13.86	**1.43**	10.73	13.86	6.34	9.90	12.50	4.45	**8.62**	**10.54**
AUS	2	7.02	11.96	14.50	7.02	11.96	14.50	16.15	16.67	19.83	**5.98**	**9.56**	**11.59**
NL	2	**12.54**	**13.21**	**17.16**	**12.54**	**13.21**	**17.16**	22.61	22.65	27.83	14.70	15.45	20.54
JAP	2	**-0.32**	**6.39**	**7.81**	**-0.32**	**6.39**	**7.81**	-2.96	7.11	8.98	1.52	7.43	8.64
FR	2	**6.07**	12.23	14.93	**6.07**	12.23	14.93	6.92	12.97	16.05	8.15	**8.82**	**11.07**
SP	3	**1.56**	**3.78**	**4.97**	1.88	7.25	8.96	2.87	7.91	10.10	5.55	6.56	8.60
SWIT	2	**0.63**	**7.50**	**9.28**	**0.63**	**7.50**	**9.28**	8.74	11.54	14.98	6.21	9.14	11.73
SWE	3	-0.59	**4.58**	**5.90**	**0.14**	6.27	7.96	10.44	11.62	14.76	-0.49	4.94	6.33
PORT	3	**4.84**	**7.76**	**10.10**	5.18	10.01	12.88	11.79	14.36	18.29	12.14	12.67	16.10

Table: The 15-year-ahead male mortality forecasting results for ages 50–89, from 1997–2011.

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.
- The study also imply that a simpler model may perform better for long-term mortality forecasts.

End of presentation

Thank you!

Any questions/ comments/ suggestions?

Contact email: han.li@unsw.edu.au