Variations of the linear logarithm hazard transform for modelling cohort mortality rates

Cary Tsai and Qian Wang

Department of Statistics and Actuarial Science
Simon Fraser University, Canada

The 49th Actuarial Research Conference
University of Califonia - Santa Barbara
July 14-16, 2014

Motivations of developing effective mortality models

- Modeling the changes and dynamics of mortality rates is critical to the solvency of life insurers and social benefit programs
- Mortality is also one of the key factors in pricing and reserving of life insurance and annuity products
- Failing in capturing the downward trends in mortality rates would under-price/-reserve annuity products and then expose annuity providers to the risk of financial insolvency

Mortality rates and sequences

- Three ways of expressing mortality rates
  - Central death rate $m_{x,t}$
  - Death probability $q_{x,t}$
  - Force of mortality $\mu_{x,t}$
- A period mortality sequence $u_{x_0,t_0,n}$ of length n starting at age x_0 in year t_0 is defined as

$$u_{x_0,t_0,n} = \{u_{x_0,t_0}, u_{x_0+1,t_0}, \ldots, u_{x_0+n-1,t_0}\}, u \in \{m, q, \mu\}.$$

- A cohort mortality sequence $u^c_{x_0,t_0,n}$ of length n starting at age x_0 in year t_0 is defined as

$$u^c_{x_0,t_0,n} = \{u_{x_0,t_0}, u_{x_0+1,t_0+1}, \ldots, u_{x_0+n-1,t_0+n-1}\}, u \in \{m, q, \mu\}.$$

Mortality rate sequences

Figure: Illustration of $q_{x_0,t_0,n}$, $q^c_{x_0,t_0+1,n}$ and $q^c_{x_0,t_0+m,n}$

Age\Year	t_0	t_0+1	$\cdots$	t_0+m	t_0+m+1	$\cdots$	t_0+n	$\cdots$	$t_0+m+n-1$
x_0	q_{x_0,t_0}	q_{x_0,t_0+1}	$\cdots$	q_{x_0,t_0+m}					
x_0+1	q_{x_0+1,t_0}		$\ddots$		$\ddots$				
x_0+2	q_{x_0+2,t_0}			$\ddots$		$\ddots$			
$\vdots$					$\ddots$			$\ddots$	
x_0+n-1	q_{x_0+n-1,t_0}						q_{x_0+n-1,t_0+n}		$q_{x_0+n-1,t_0+m+n-1}$

$\boxed{q_{x_0,t_0,n}}$

$\boxed{q^c_{x_0,t_0+1,n}}$

$\boxed{q^c_{x_0,t_0+m,n}}$

The Lee-Carter model

$$\ln(m_{x,t}) = a_x + b_x k_t + \epsilon_{x,t}$$

- a_x and b_x are the age-specific constants, $x = x_0, \ldots, x_0 + n - 1$;
- k_t is the time-varying index, $t = t_0 + 1, \ldots, t_0 + m$;
- for each x, error terms $\epsilon_{x,t} \overset{i.i.d}{\sim} N\left(0, \sigma^2_{\epsilon_x}\right)$, $t = t_0 + 1, \ldots, t_0 + m$;
- two constraints: $\sum_t k_t = 0$ and $\sum_x b_x = 1$.

The parameters are estimated by SVD (singular value decomposition); or alternatively,

- $\sum_t k_t = 0 \Rightarrow \hat{a}_x = \frac{1}{m} \sum_{t=t_0+1}^{t_0+m} \ln(m_{x,t})$,
- $\sum_x b_x = 1 \Rightarrow \hat{k}_t = \sum_{x=x_0}^{x_0+n-1} [\ln(m_{x,t}) - \hat{a}_x]$,
- $\hat{b}_x$ can be computed by regressing $[\ln(m_{x,t}) - \hat{a}_x]$ on $\hat{k}_t$ without the constant term for each age x.

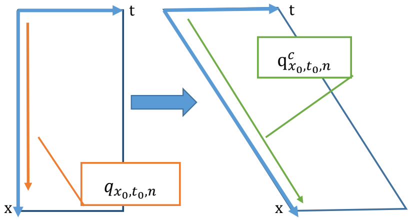
The CBD model

$$logit(q_{x,t}) = ln(\frac{q_{x,t}}{1-q_{x,t}}) = \kappa_t^1 + \kappa_t^2(x-\bar{x}) + \epsilon_{x,t}$$

- $\bar{x} = \frac{1}{n}\sum_{x=x_0}^{x_0+n-1} x$;
- κ_t^1 and κ_t^2 are the time-varying parameters, $t = t_0+1,\ldots,t_0+m$, and can be obtained by the least square method;
- for each x, error terms $\epsilon_{x,t} \overset{i.i.d}{\sim} N(0, \sigma^2_{\epsilon_x})$, $t = t_0+1,\ldots,t_0+m$.

Data area for the Lee-Carter and CBD models

Figure



- data area: rectangle $\Rightarrow$ parallelogram
- m cohort mortality sequences $q^c_{x_0,t_0+1,n}, \ldots, q^c_{x_0,t_0+m,n}$ are used to fit the Lee-Carter and CBD models and estimate parameters.

The Lee-Carter model

- Modelling the time factor k_t with a random walk with drift process, $\hat{k}_t = \hat{k}_{t-1} + c + e_t$, where c is estimated by

$$\hat{c} = \frac{1}{m-1} \sum_{t=t_0+1}^{t_0+m-1} (\hat{k}_{t+1} - \hat{k}_t).$$

- The projected $m_{x,t_0+m+\tau}$ is obtained by

$$\ln(\hat{m}_{x,t_0+m+\tau}) = \hat{a}_x + \hat{b}_x \times \tilde{\hat{k}}_{t_0+m+\tau},$$

where

$$\tilde{\hat{k}}_{t_0+m+\tau} = \hat{k}_{t_0+m} + \tau \times \hat{c}.$$

The CBD model

- Assuming random walk with drift for κ_t, $\hat{\kappa}_t = \hat{\kappa}_{t-1} + c + V \cdot Z_t$
  - $\hat{\kappa}_t = (\hat{\kappa}_t^1, \hat{\kappa}_t^2)'$ and $c = (c_1, c_2)'$;
  - V is the variance-covariance matrix for κ;
  - Z_t is the two-dimension standard normal distribution.
- Similar to the Lee-Carter model, c is estimated as

$$\hat{c} = \frac{1}{m-1} \sum_{t=t_0+1}^{t_0+m-1} (\hat{\kappa}_{t+1} - \hat{\kappa}_t).$$

- The projected $q_{x,t_0+m+\tau}$ is obtained by

$$logit(\hat{q}_{x,t_0+m+\tau}) = \tilde{\hat{\kappa}}^1_{t_0+m+\tau} + \tilde{\hat{\kappa}}^2_{t_0+m+\tau}(x - \bar{x}),$$

where

$$\tilde{\hat{\kappa}}^i_{t_0+m+\tau} = \hat{\kappa}^i_{t_0+m} + \tau \times \hat{c}_i, \quad i = 1, 2.$$

A variation of the LLHT model

For any year $t + t_0$, the LLHT model is expressed as follows:

$$\ln(\mu^{c}_{x_0,t_0+t,n}) = \alpha_t \times \ln(\mu_{x_0,t_0,n}) + \beta_t + \epsilon_t.$$

Regressing $\ln(\mu^{c}_{x_0,t_0+t,n})$ on $\ln(\mu_{x_0,t_0,n})$ for each $t = 1, \ldots, m$ yields $\{(\hat{\alpha}_t, \hat{\beta}_t), t = 1, \ldots, m\}$ where $\epsilon_t = \{\epsilon_{x_0,t}, \ldots, \epsilon_{x_0+n-1,t}\} \overset{i.i.d}{\sim} N(0, \sigma^2_{\epsilon_t})$.

Under the constant force of mortality assumption,

$$q_{x,t} = 1 - e^{-\int_0^1 \mu_{x,t}(s)ds} = 1 - e^{-\mu_{x,t}}$$

and

$$m_{x,t} = \frac{q_{x,t}}{\int_0^1 {}_s p_{x,t} ds} = \frac{q_{x,t}}{\int_0^1 p^s_{x,t} ds} = -\ln(p_{x,t}) = \mu_{x,t}.$$

Mortality rate sequence

Figure: Fitting $\ln(\mu^c_{x_0,t_0+t,n})$ with $\ln(\mu_{x_0,t_0,n})$, $t = 1\ldots,m$

Age\Year	t_0	t_0+1	$\cdots$	t_0+m	t_0+m+1	$\cdots$	t_0+n	$\cdots$	$t_0+m+n-1$
x_0	q_{x_0,t_0}	q_{x_0,t_0+1}	$\cdots$	q_{x_0,t_0+m}					
x_0+1	q_{x_0+1,t_0}		$\ddots$		$\ddots$				
x_0+2	q_{x_0+2,t_0}			$\ddots$		$\ddots$			
$\vdots$					$\ddots$			$\ddots$	
x_0+n-1	q_{x_0+n-1,t_0}						q_{x_0+n-1,t_0+n}		$q_{x_0+n-1,t_0+m+n-1}$

$\boxed{q^c_{x_0,t_0+1,n}}$

$\boxed{q^c_{x_0,t_0+m,n}}$

$\boxed{q_{x_0,t_0,n}}$

Figure: The fitted values $\hat{\beta}_t$ for the UK males from regressing $ln(\mu^c_{45,1956,35}), \ldots, ln(\mu^c_{45,1975,35})$ on $ln(\mu_{45,1955,35})$

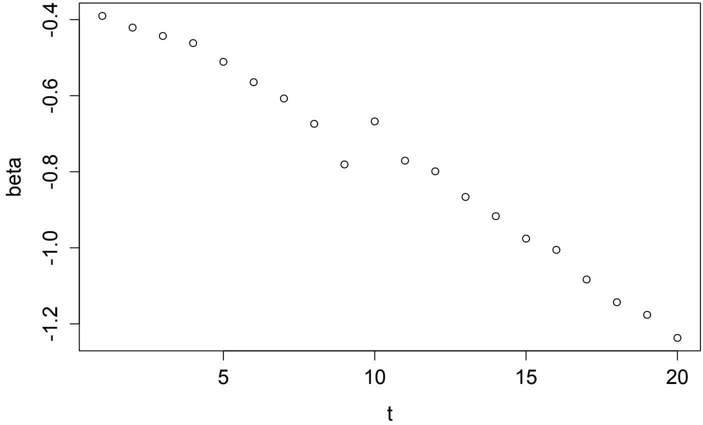
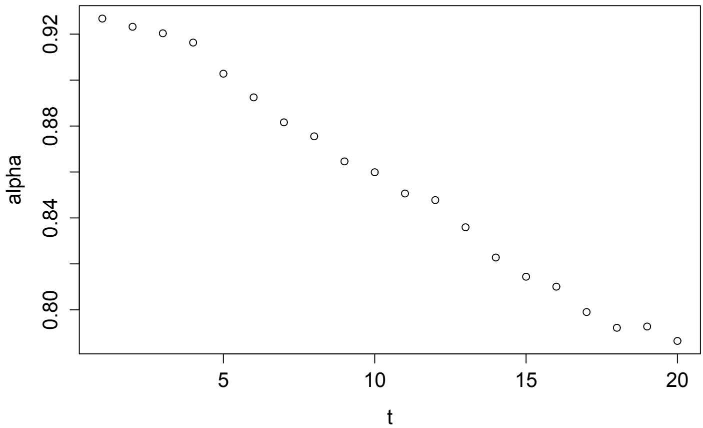


Figure: The fitted values $\hat{\alpha}_t$ for the UK males from regressing $ln(\mu^c_{45,1956,35}), \ldots, ln(\mu^c_{45,1975,35})$ on $ln(\mu_{45,1955,35})$



The LLHT-LR model

Assume that

$$\hat{\alpha}_t = a \times t + b + e_{\alpha,t}, t = 1, \ldots, m$$

and

$$\hat{\beta}_t = c \times t + d + e_{\beta,t}, t = 1, \ldots, m.$$

- $e_{\gamma,t}$, $t = 1, \ldots, m$, are error terms, and $e_{\gamma,t} \overset{i.i.d}{\sim} N\left(0, \sigma^2_{e_\gamma}\right)$, $\gamma = \{\alpha, \beta\}$.
- a, b, c and d are all constants and can be estimated by least square error method

Future mortality rate can be predicted by

$$ln(\hat{\mu}^c_{x,t_0+m+\tau}) = \tilde{\hat{\alpha}}_{m+\tau} \times \ln(\mu_{x,t_0}) + \tilde{\hat{\beta}}_{m+\tau}.$$

where $\tilde{\hat{\alpha}}_{m+\tau} = \hat{a} \times (m + \tau) + \hat{b}$ and $\tilde{\hat{\beta}}_{m+\tau} = \hat{c} \times (m + \tau) + \hat{d}$

The LLHT-RW model

Assume that

$$\hat{\alpha}_t = \hat{\alpha}_{t-1} + f + e_{\alpha,t}$$

and

$$\hat{\beta}_t = \hat{\beta}_{t-1} + g + e_{\beta,t}.$$

- f and g are all constants, and is estimated as parameters by random walk with drift,
- $e_{\gamma,t}$, $t = 1, \ldots, m$ are error terms, and $e_{\gamma,t} \overset{i.i.d}{\sim} N(0, \sigma^2_{e_\gamma})$, $\gamma = \{\alpha, \beta\}$.

Future mortality rate can be predicted by

$$\ln(\hat{\mu}^c_{x,t_0+m+\tau}) = \tilde{\hat{\alpha}}_{m+\tau} \times \ln(\mu_{x,t_0}) + \tilde{\hat{\beta}}_{m+\tau}.$$

The Lee-Carter model

The variance comes from two sources:

- e_t from k_t

$$\hat{\sigma}_e^2 = \frac{1}{m-2} \sum_{t=t_0+1}^{t_0+m-1} [\hat{k}_{t+1} - \hat{k}_t - \hat{c}]^2$$

- $\epsilon_{x,t}$ from the whole model, for each age x

$$\hat{\sigma}_{\epsilon_x}^2 = \frac{1}{m-2} \sum_{t=t_0+1}^{t_0+m} [ln(m_{x,t}) - \hat{a}_x - \hat{b}_x \times \hat{k}_t]^2$$

The estimated variance of the projected $ln(m_{x,t})$ is obtained as follows,

$$\hat{\sigma}^2(ln(\hat{m}_{x,t_0+m+\tau})) = \hat{b}_x^2 \times \hat{\sigma}^2_{\hat{\tilde{k}}_{t_0+m+\tau}} + \hat{\sigma}_{\epsilon_x}^2 = \tau \times \hat{b}_x^2 \times \hat{\sigma}_e^2 + \hat{\sigma}_{\epsilon_x}^2,$$

The CBD model

The variance comes from two sources:

- errors from κ

$$\hat{\sigma}^2_{\kappa_t^{(i)}} = \frac{1}{m-2} \sum_{t=t_0+1}^{t_0+m-1} [\hat{\kappa}_{t+1}^{(i)} - \hat{\kappa}_t^{(i)} - \hat{c}_{(i)}]^2, i = \{1, 2\}$$

- $\epsilon_{x,t}$ from the whole model, for each age x

$$\hat{\sigma}^2_{\epsilon_x} = \frac{1}{m-2} \sum_{t=t_0+1}^{t_0+m} [logit(q_{x,t}) - \hat{\kappa}_t^1 - \hat{\kappa}_t^2 \times (x - \bar{x})]^2$$

The variance of the projected $logit(q_{x,t})$ is obtained as follows,

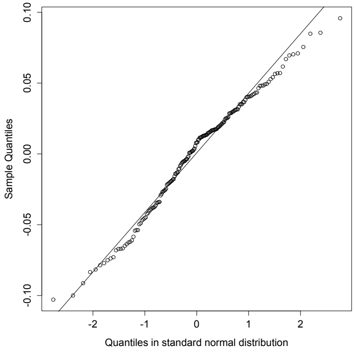
$$\sigma^2(logit(\hat{q}_{x,t})) = \tau \times \sigma^2_{\kappa_t^1} + \tau \times \sigma^2_{\kappa_t^2} \times (x - \bar{x})^2 + \sigma^2_{\epsilon_x}.$$

The LLHT-based models

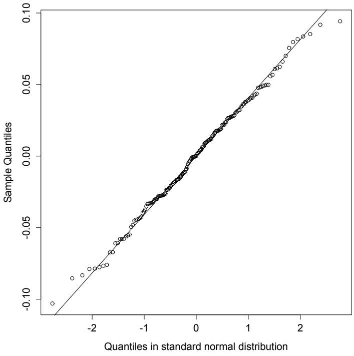
- The variance comes from two sources:
  - $\epsilon_{x,t}$, the error term from the whole model
  - errors from estimated intercept and slope parameters
- For $\epsilon_{x,t}$,
  - at first we assume for each t $\epsilon_{x,t} \overset{i.i.d}{\sim} N(0, \sigma^2_{\epsilon_t}), x = x_0, \ldots, x_0 + n - 1.$
  - From empirical data, $\sigma^2_{\epsilon_t}$ does not increase in t, so we assume $\epsilon_{x,t} \overset{i.i.d}{\sim} N(0, \sigma^2_{\epsilon}), x = x_0, \ldots, x_0 + n - 1, t = t_0 + 1, \ldots, t_0 + m.$
  - $\hat{\sigma}^2_{\epsilon}$ can be estimated by

$$\hat{\sigma}^2_{\epsilon} = \frac{\sum_{t=1}^{m} \hat{\sigma}^2_{\epsilon_t}}{m} = \frac{\sum_{t=1}^{m} \sum_{k=1}^{n} (ln(\mu_{x_0+k-1,t_0+t}) - \hat{\alpha}_t \times ln(\mu_{x_0+k-1,t_0}) - \hat{\beta}_t)^2}{m \times (n-2)}.$$

Figure: Q-Q plots of all fitted errors by regressing $ln(\mu^c_{1956,45,35}), \ldots, ln(\mu^c_{1960,45,35})$ on $ln(\mu_{1955,45,35})$ for the USA population



(a) Male



(b) Female

The LLHT-LR/LLHT-RW models

For the LLHT-LR model, the estimate of standard deviation of the predicted coefficients is obtained by

$$\hat{\sigma}(\hat{\gamma}_{m+\tau}) = \hat{\sigma}_{e_\gamma}\sqrt{1+\frac{1}{m}+\frac{(m+\tau-\bar{t})^2}{\sum_{t=1}^{m}(t-\bar{t})^2}}, \gamma \in \{\alpha, \beta\}$$

where $\hat{\sigma}^2_{e_\alpha}$ and $\hat{\sigma}^2_{e_\beta}$ are obtained from the residules.
For the LLHT-RW model, the estimates of variance of the predicted coefficient are

$$\hat{\sigma}^2(\hat{\alpha}_{m+\tau}) = \hat{\sigma}^2_{e_\alpha} \times \tau, \quad \hat{\sigma}^2(\hat{\beta}_{m+\tau}) = \hat{\sigma}^2_{e_\beta} \times \tau,$$

where

$$\hat{\sigma}^2_{e_\alpha} = \frac{1}{m-2}\sum_{t=1}^{m-1}[\hat{\alpha}_{t+1}-\hat{\alpha}_t-\hat{f}]^2, \quad \hat{\sigma}^2_{e_\beta} = \frac{1}{m-2}\sum_{t=1}^{m-1}[\hat{\beta}_{t+1}-\hat{\beta}_t-\hat{g}]^2.$$

Combining the two sources of variances, the total variance is

$$\hat{\sigma}^2(ln(\hat{\mu}^c_{x,t_0+m+\tau})) = \hat{\sigma}^2(\hat{\alpha}_{m+\tau}) \times [ln(\mu_{x,t_0})]^2 + \hat{\sigma}^2(\hat{\beta}_{m+\tau}) + \hat{\sigma}^2_\epsilon.$$

Data for illustration

- For the Lee-Carter and CBD models, 5 cohort mortality sequences $q^c_{x_0,1956,35}, \ldots, q^c_{x_0,1960,35}$ are used for fitting.
- For the LLHT models, the period mortality sequence $q_{x_0,1955,35}$ is used to fit the 5 cohort mortality sequences $q^c_{x_0,1956,35}, \ldots, q^c_{x_0,1960,35}$, respectively.
- The 15 cohort mortality sequences $q^c_{x_0,1961,35}, \ldots, q^c_{x_0,1975,35}$ are predicted for all models.
- The USA, the UK and Japan population are selected from Human Mortality Database with $x_0 \in [40, 50]$.

Errors for evaluating accuracy of projection

The MAE_t (mean absolute error), $RMSE_t$ (root mean square error) and $MAPE_t$ (mean absolute percentage error) for $q^c_{x_0,t,n}$ over the age span $[x_0, x_0 + n - 1]$ in year t are given by

$$MAE_t = \frac{1}{n}\sum_{k=1}^{n} \left|\hat{q}_{x_0+k-1,t+k-1} - q_{x_0+k-1,t+k-1}\right|,$$

$$RMSE_t = \sqrt{\frac{1}{n}\sum_{k=1}^{n}[\hat{q}_{x_0+k-1,t+k-1} - q_{x_0+k-1,t+k-1}]^2}$$

and

$$MAPE_t = \frac{1}{n}\sum_{k=1}^{n}\left|\frac{\hat{q}_{x_0+k-1,t+k-1} - q_{x_0+k-1,t+k-1}}{q_{x_0+k-1,t+k-1}}\right| \times 100\%.$$

The overall *MAE*, *RMSE* and *MAPE* over $[x_0, x_0 + n - 1]$ and $[t_0 + m + 1, t_0 + m + M]$ are calculated for comparison.

$$MAE^{[x_0, x_0+n-1]}_{[t_0+m+1, t_0+m+M]} = \frac{1}{M} \sum_{t=t_0+m+1}^{t_0+m+M} MAE_t,$$

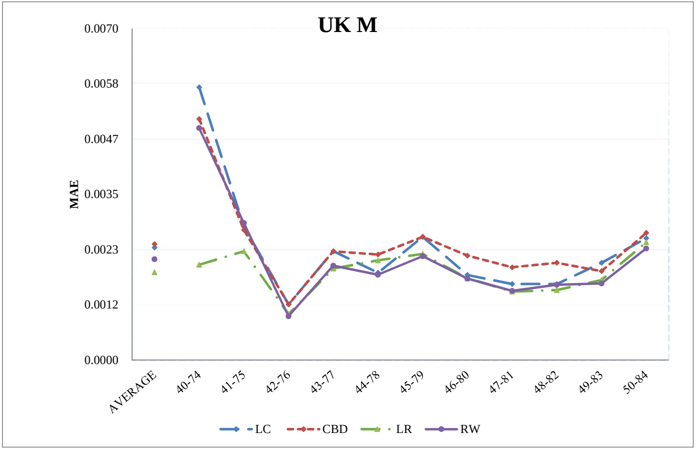
$$RMSE^{[x_0, x_0+n-1]}_{[t_0+m+1, t_0+m+M]} = \frac{1}{M} \sum_{t=t_0+m+1}^{t_0+m+M} RMSE_t$$

and

$$MAPE^{[x_0, x_0+n-1]}_{[t_0+m+1, t_0+m+M]} = \frac{1}{M} \sum_{t=t_0+m+1}^{t_0+m+M} MAPE_t.$$

Accuracy of projection

Figure: Projected errors *MAE* against x_0 for the UK males



Accuracy of projection

Figure: Projected errors *RMSE* against x_0 for the UK males

Accuracy of projection

Figure: Projected errors *MAPE* against x_0 for the UK males

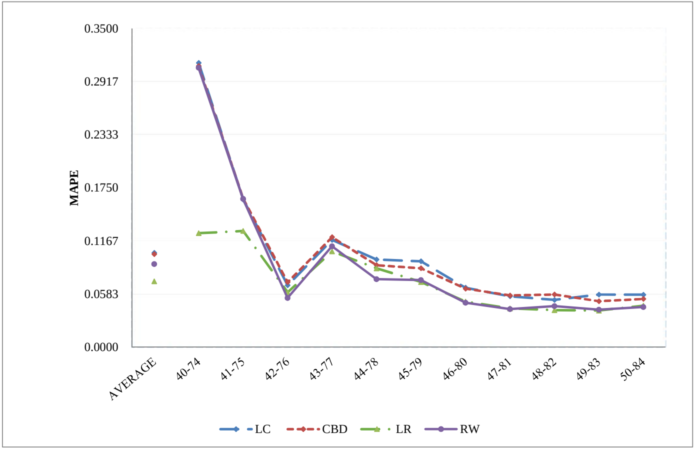


Table: Overall average projection errors over 11 age groups

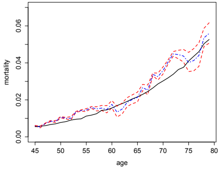
error	method	USA M	USA F	UK M	UK F	JAP M	JAP F	Average
MAE	LC	18.93	8.98	23.79	13.62	17.57	7.57	15.08
	CBD	13.97	10.44	24.51	10.83	18.81	10.74	14.88
	LR	13.97	8.14	18.58	10.23	19.76	8.21	13.15
	RW	14.27	8.42	21.34	11.89	19.63	8.24	13.97
RMSE	LC	24.87	12.13	33.51	19.29	23.82	10.78	20.73
	CBD	17.69	17.13	36.71	14.82	27.98	17.85	22.03
	LR	18.10	13.42	27.43	14.94	30.51	13.98	19.73
	RW	18.31	13.44	30.67	16.90	30.27	13.98	20.59
MAPE	LC	8.47	7.22	10.37	10.38	9.78	8.27	9.08
	CBD	6.28	6.52	10.22	8.54	9.10	8.98	8.27
	LR	6.15	5.42	7.24	7.36	9.30	6.65	7.02
	RW	6.34	5.69	9.13	9.17	9.13	6.63	7.68

Note that *RMSE* and *MAE* are scaled to ($\times 10^{-4}$) and *MAPE* is a percentage.

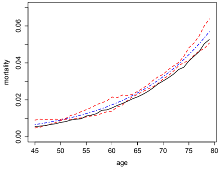
Here, the overall average error is defined as $E = \frac{1}{11}\sum_{x_0=40}^{50} E^{[x_0,x_0+n-1]}_{[t_0+m+1,t_0+m+M]}$, $E \in \{MAE, RMSE, MAPE\}$.

Confidence interval

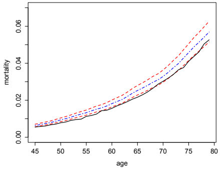
Figure: 95% confidence intervals on $q^{c}_{45,1975,35}$ for the USA males



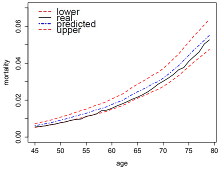
(a) LC



(b) CBD



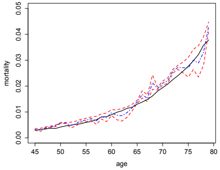
(c) LR



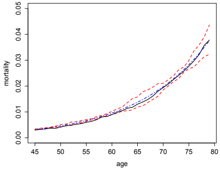
(d) RW

Confidence interval

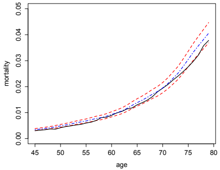
Figure: 95% confidence intervals on $q^c_{45,1975,35}$ for the USA females



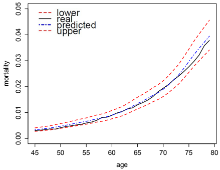
(a) LC



(b) CBD

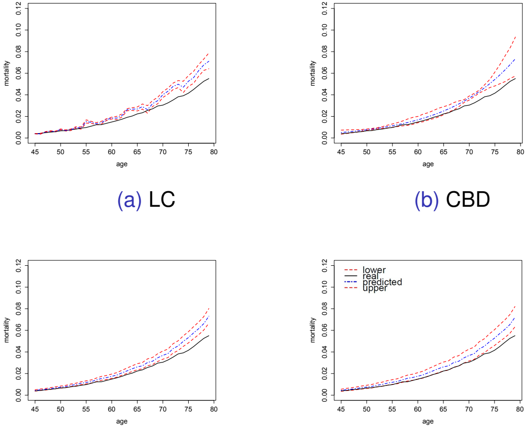


(c) LR



(d) RW

Figure: 95% confidence intervals on $q^c_{45,1975,35}$ for the UK males

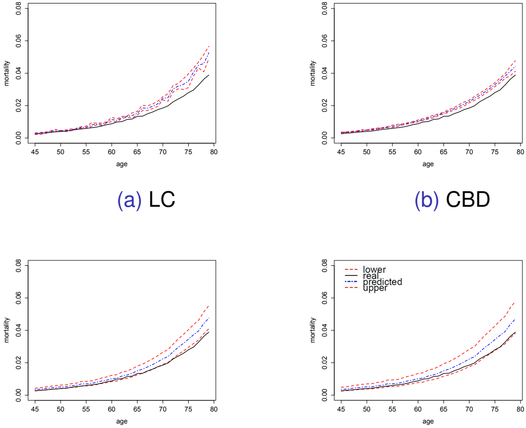


(a) LC (b) CBD

(c) LR (d) RW

Confidence interval

Figure: 95% confidence intervals on $q^c_{45,1975,35}$ for the UK females

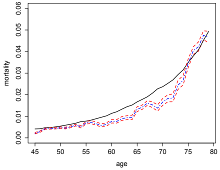


(a) LC (b) CBD

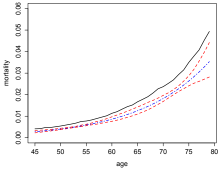
(c) LR (d) RW

Confidence interval

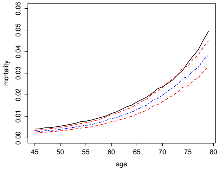
Figure: 95% confidence intervals on $q^c_{45,1975,35}$ for the Japan males



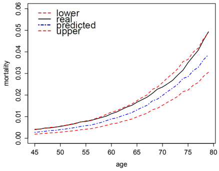
(a) LC



(b) CBD



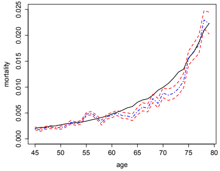
(c) LR



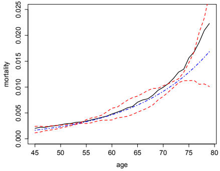
(d) RW

Confidence interval

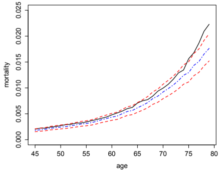
Figure: 95% confidence intervals on $q^{c}_{45,1975,35}$ for the Japan females



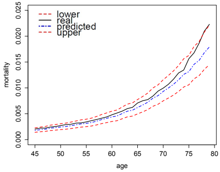
(a) LC



(b) CBD



(c) LR



(d) RW

Errors in pricing

- $A^1_{x_0:\overline{n}|}$, $\ddot{a}_{x_0:\overline{n}|}$ and $P^1_{x_0:\overline{n}|}$ are calculated based on the real and predicted $q^c_{x_0,t,35}$ for $x_0 \in [40, 50]$ and $t \in [1961, 1975]$.
- For each x_0 and t, the relative error is

$$\frac{\hat{X}_{x_0:\overline{n}|}}{X_{x_0:\overline{n}|}} - 1, X = \{A^1, \ddot{a}, P^1\}.$$

- For each year t, the average relative error of each year t over all x_0s (11 age groups) are calculated.
- a poorer forecast on a cohort mortality sequence does not necessarily lead to a higher relative error on premium since ${}_np_{x_0,t} = p_{x_0,t} \cdots p_{x_0+n-1,t+n-1}$.

Figure: Average relative errors on $A^{1}_{x_0:\overline{35}|}$ for the USA males

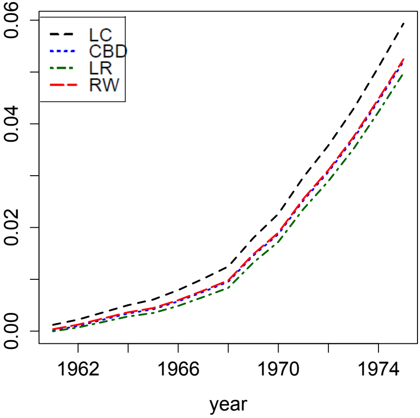


Figure: Average relative errors on $\ddot{a}_{x_0:\overline{35}|}$ for the USA males

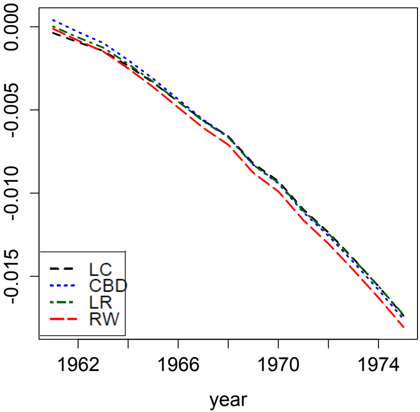
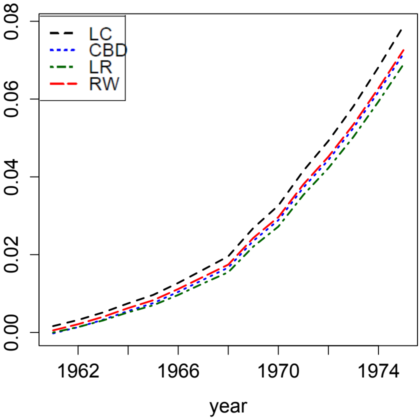


Figure: Average relative errors on $P^{1}_{x_0:\overline{35}|}$ for the USA males



Thank you!