Mortality Modelling with R

Andrés M. Villegas

School of Risk and Actuarial Studies, UNSW Sydney

15 January 2020, Living to 100 Symposium
Orlando, Florida

Aims

Gain familiarity with use of the R packages **demography**, **StMoMo** and **lifecontingencies** to:

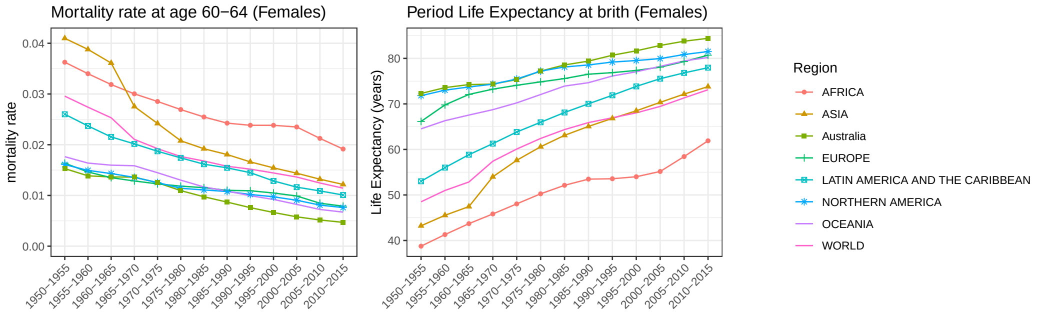
- Obtain mortality data from Human Mortality Database
- Fit stochastic mortality models (including Lee-Carter model)
- Compare goodness of fit of mortality models
- Forecast future mortality rates (deterministic and using Monte Carlo simulations)
- Compute demographic and actuarial quantities

Agenda

- Motivation
- Review of mortality modelling metodologies
- Mortality modelling with R
  - Downloading mortality data with **demography**
  - Fitting, comparing and projecting mortality model with **StMoMo**
  - Turning mortality projections into demographic calculations with **lifecontingencies**
- Conclusions and Summary
- Preview of recent **StMoMo** developments (if time permits!)

Motivation and Review of Mortality Modelling

Recent trends in mortality and life expectancy



Source: United Nations World Population Prospects 2017

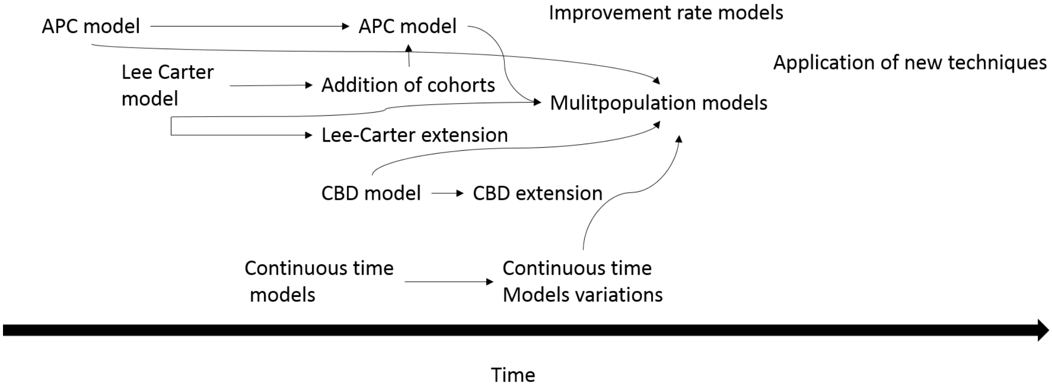
- Good-news!
- Important social and financial implications for governments, insurers, individuals.
- Need to model and project these trends

Mortality forecasting methodologies

A good overview of methodologies is given in the review papers by Booth and Tickle (2008), Wong-Fupuy and Haberman (2004), Pitacco (2004) and in the by Pitacco et al. (2009)

- Expert based
- Explanatory
  - Structural Modelling (Explanatory or Econometric).
  - Cause of death decomposition
- **Extrapolation**
  - Trend modelling

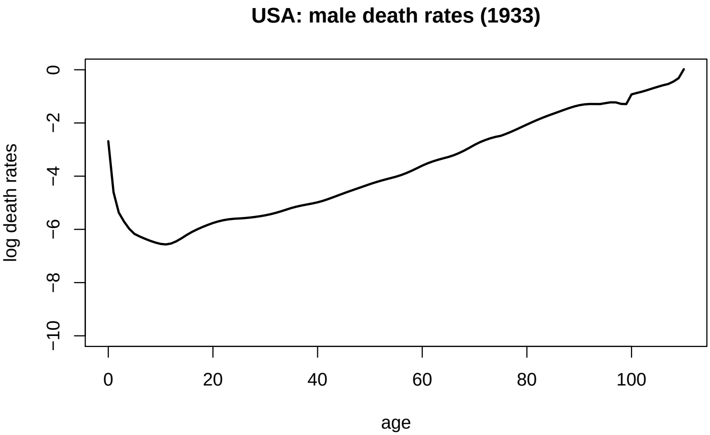
A timeline of "recent" mortality modelling methodologies



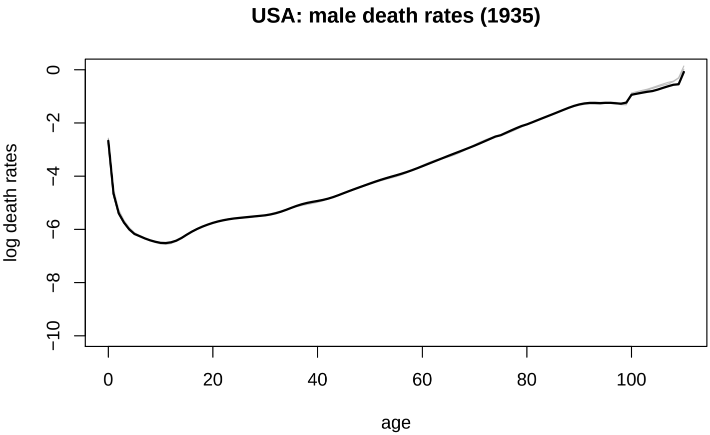
Advances in single population mortality modelling

- **Lee-Carter model** (Lee and Carter 1992)
  - Add more bilinear age-period components (Renshaw and Haberman 2003)
  - Add a cohort effect (Renshaw and Haberman 2006)
- Two factor **CBD model** (Cairns, Blake, and Dowd 2006)
  - Add cohort effect, quadratic age term (Cairns et al. 2009)
  - Combine with features of the Lee-Carter (Plat 2009)
- **Many more models** proposed in the literature (e.g. Aro and Pennanen (2011), O'Hare and Li (2012), Börger, Fleischer, and Kuksin (2013), Alai and Sherris (2014))

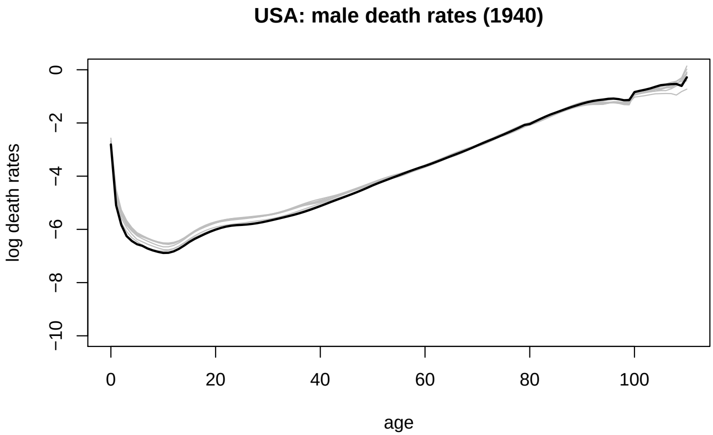
Lee-Carter model



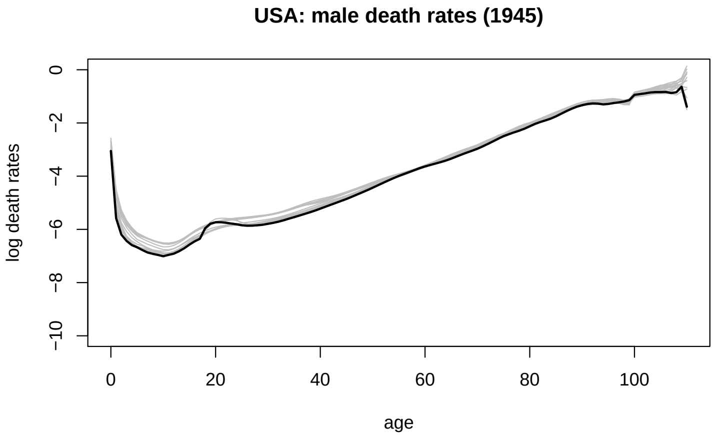
Lee-Carter model



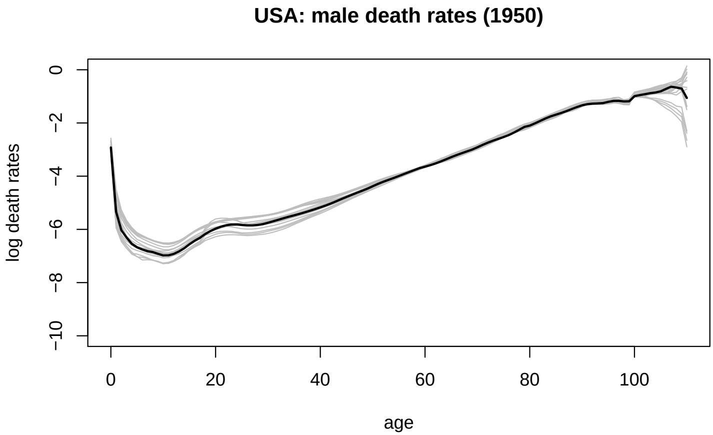
Lee-Carter model



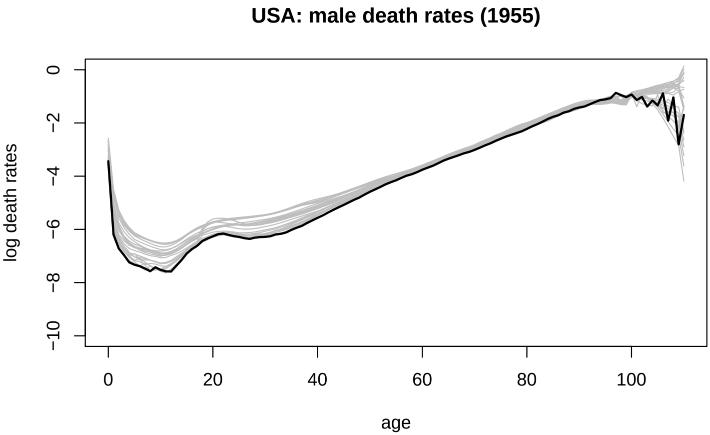
Lee-Carter model



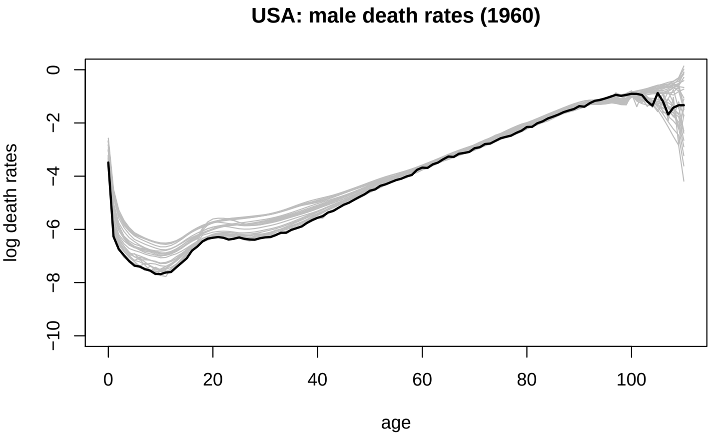
Lee-Carter model



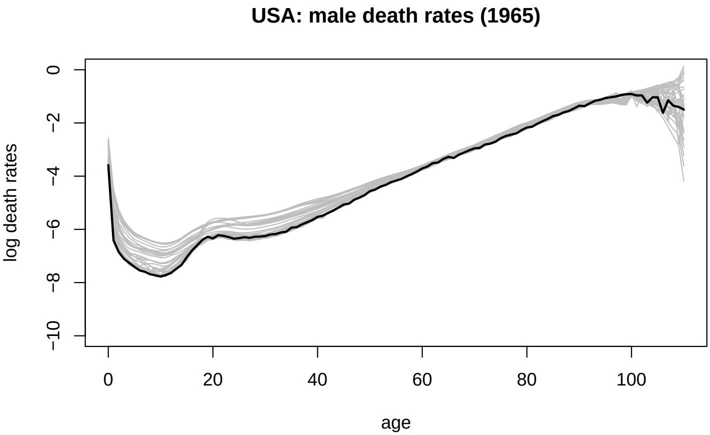
Lee-Carter model



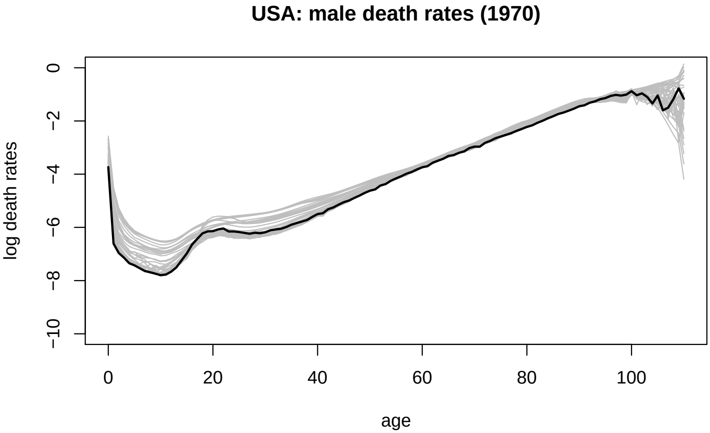
Lee-Carter model



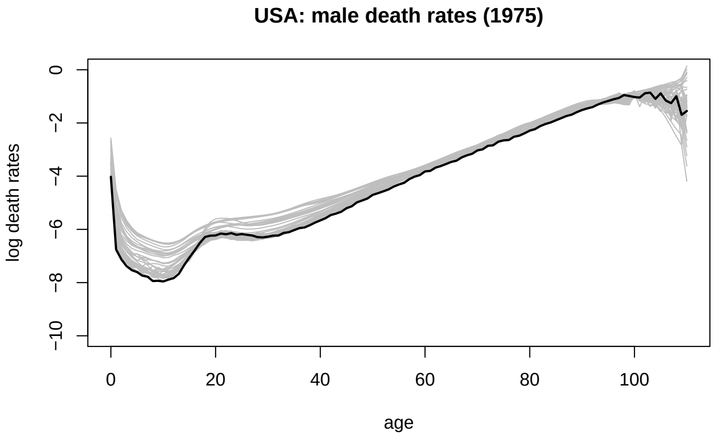
Lee-Carter model



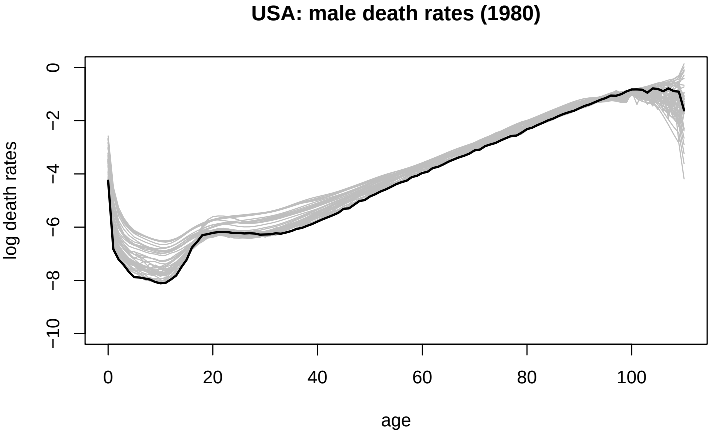
Lee-Carter model



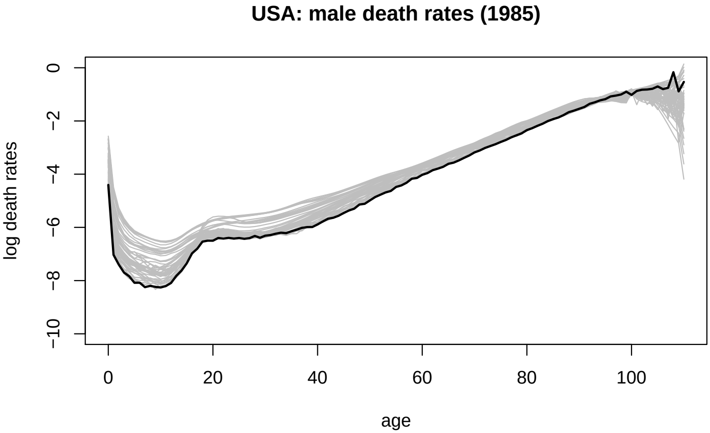
Lee-Carter model



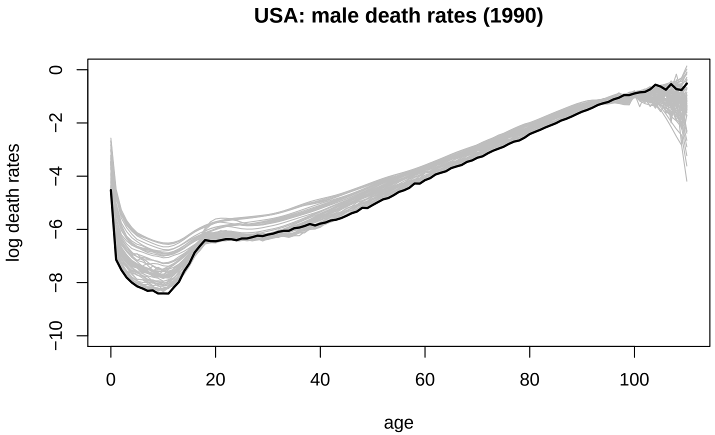
Lee-Carter model



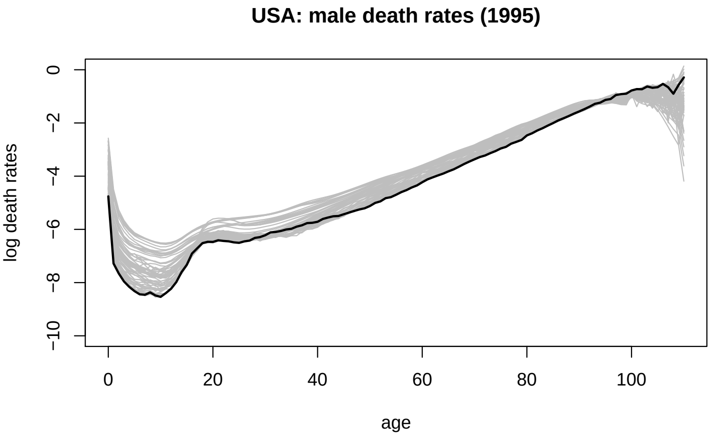
Lee-Carter model



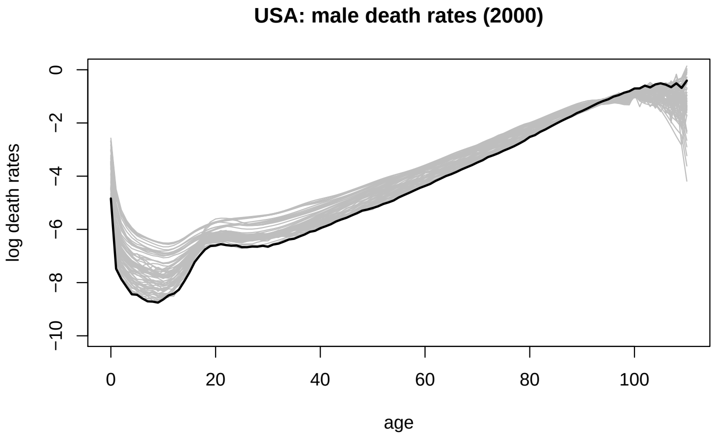
Lee-Carter model



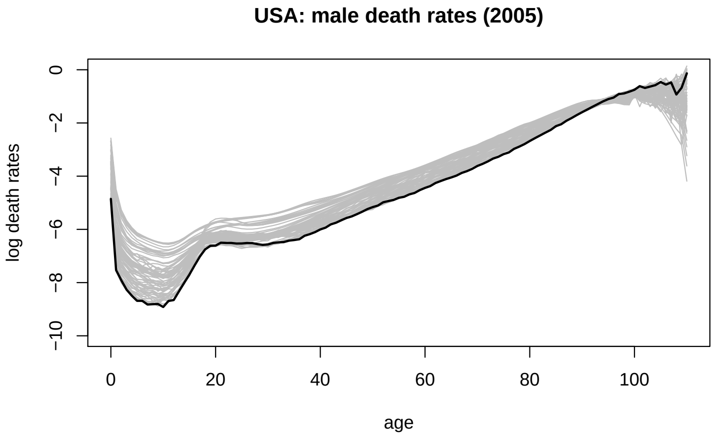
Lee-Carter model



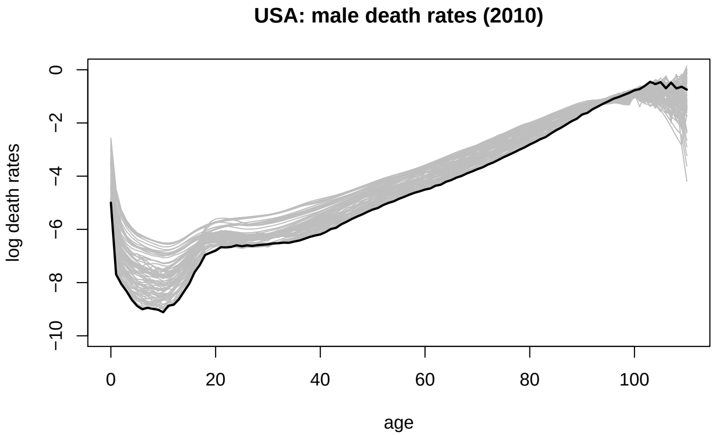
Lee-Carter model



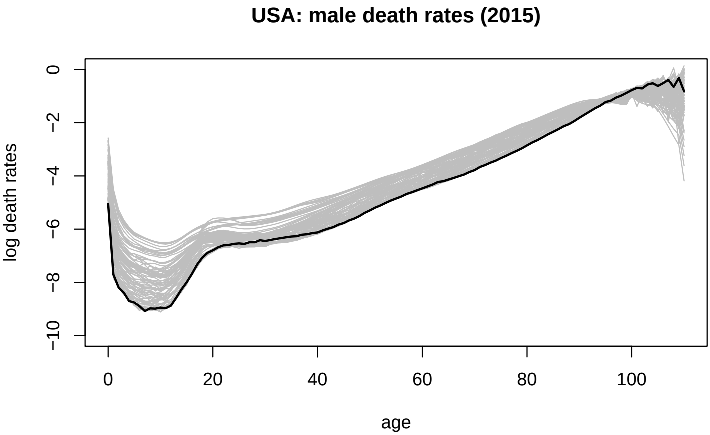
Lee-Carter model



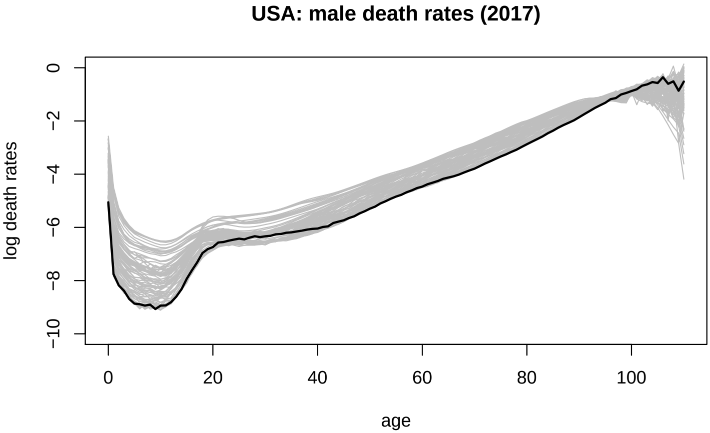
Lee-Carter model



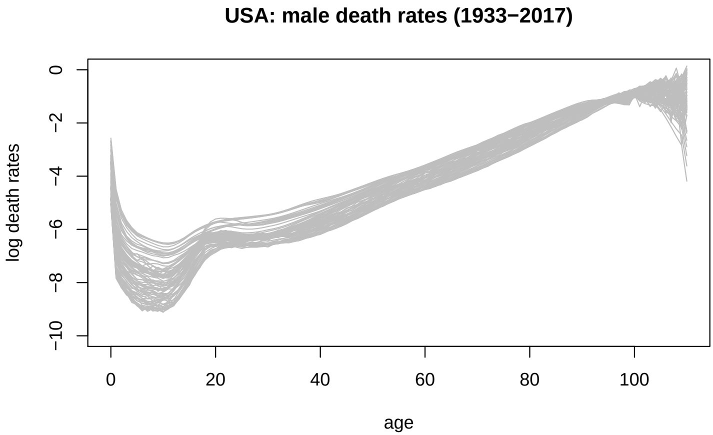
Lee-Carter model



Lee-Carter model

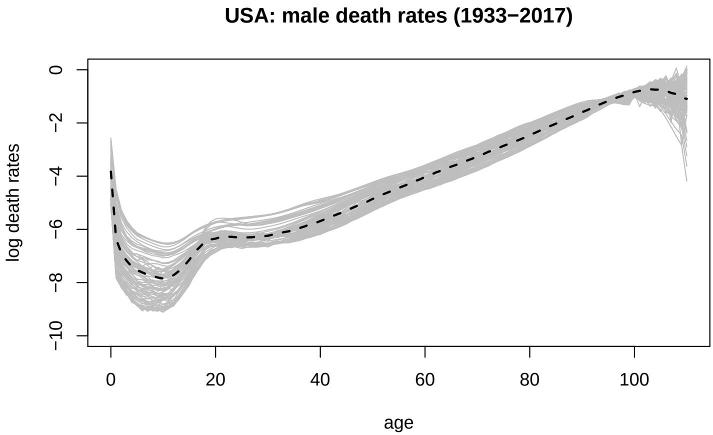


Lee-Carter model



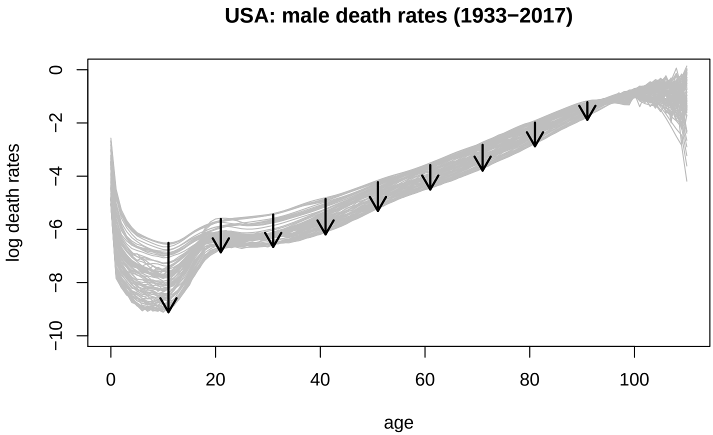
$$\log \mu_{xt} \quad =$$

Lee-Carter model



$$\log \mu_{xt} = \alpha_x$$

Lee-Carter model

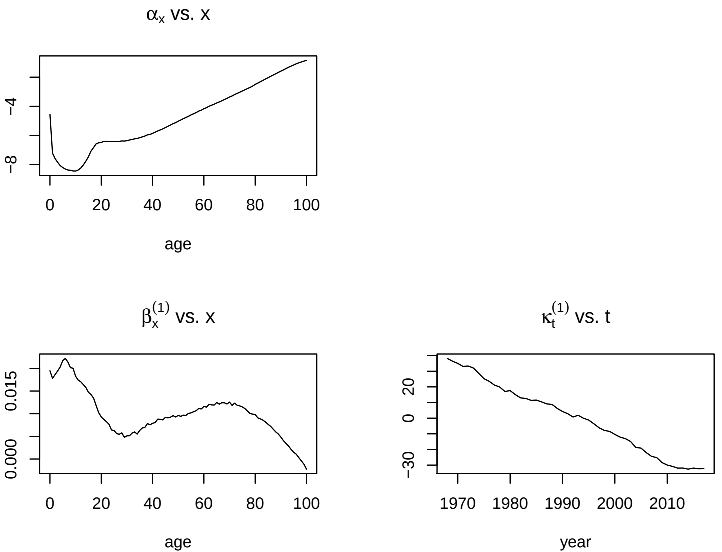


$$\log \mu_{xt} \quad = \quad \alpha_x \quad + \qquad\qquad \kappa_t$$

Lee-Carter model

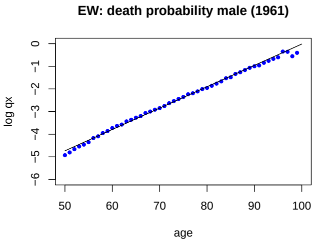
$$\log \mu_{xt} = \alpha_x + \beta_x \kappa_t$$

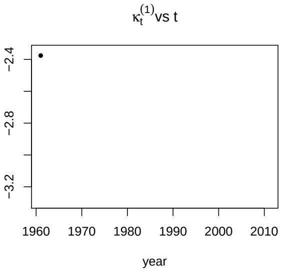
Lee-Carter model

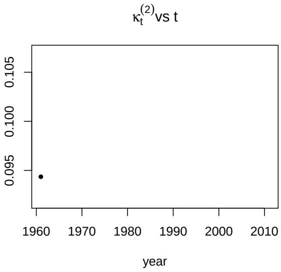


Cairns-Blake-Dowd model

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

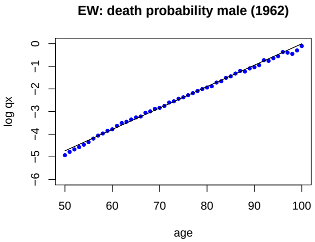


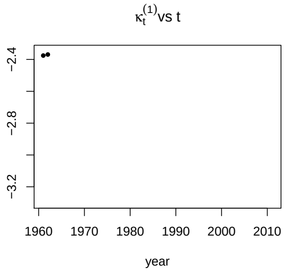


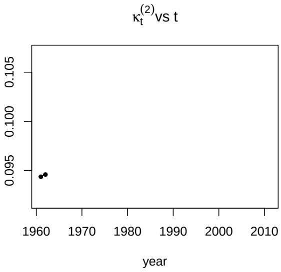


Cairns-Blake-Dowd model

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

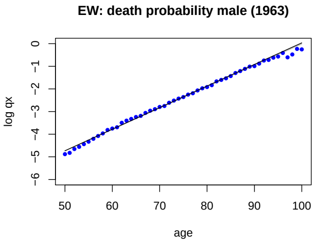


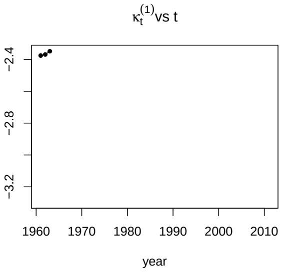


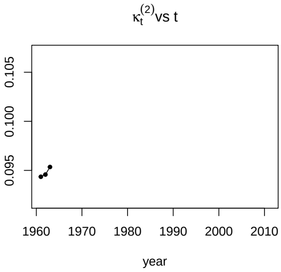


Cairns-Blake-Dowd model

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

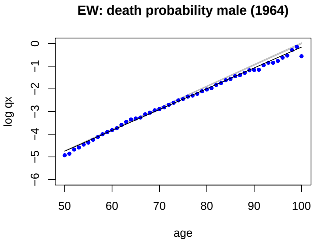


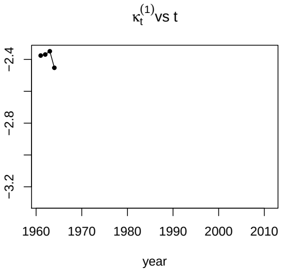


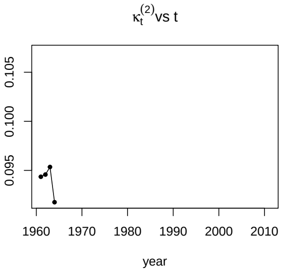


Cairns-Blake-Dowd model

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

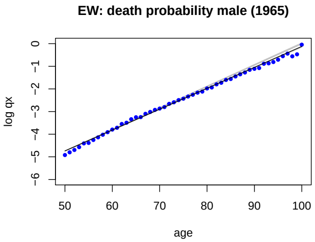


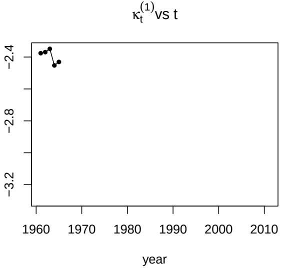


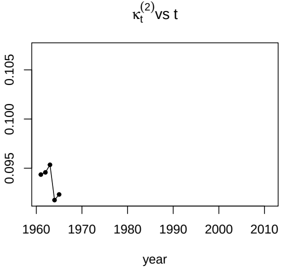


Cairns-Blake-Dowd model

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

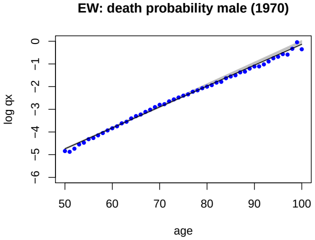


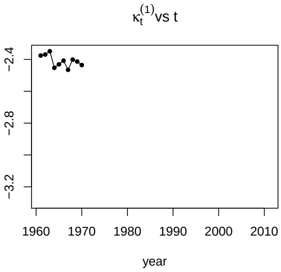


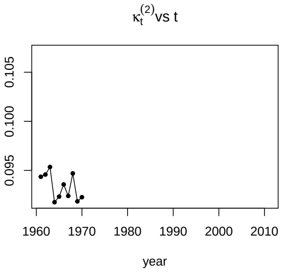


Cairns-Blake-Dowd model

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

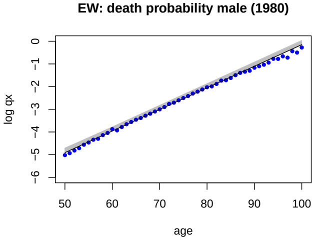


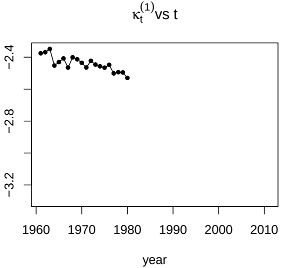


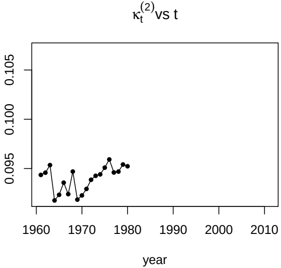


Cairns-Blake-Dowd model

$$\operatorname{logit} q_{xt} = \kappa_t^{(1)} + (x - \bar{x})\kappa_t^{(2)}$$

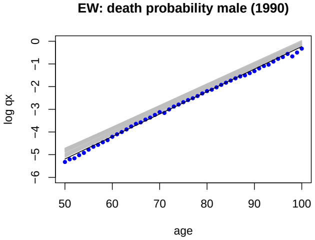


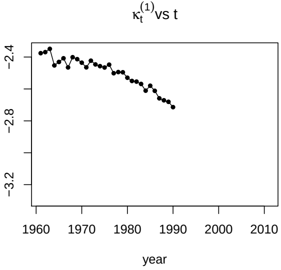


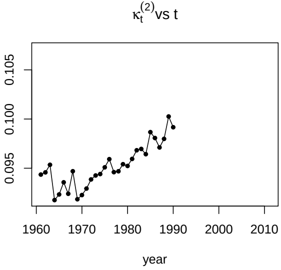


Cairns-Blake-Dowd model

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

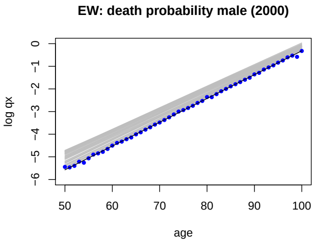


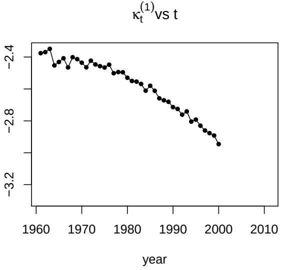


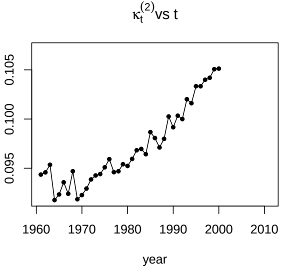


Cairns-Blake-Dowd model

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

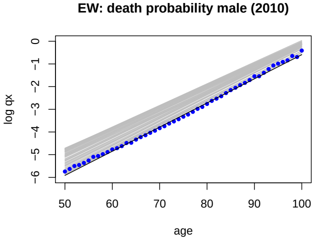


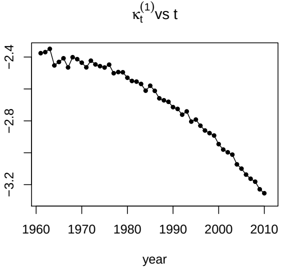


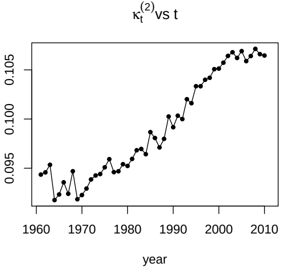


Cairns-Blake-Dowd model

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







Other Stochastic Mortality Models

Name	Form	Parameters
LC	$\alpha_x + \beta_x^{(1)}\kappa_t^{(1)}$	$2n_a + n_y$
LC2	$\alpha_x + \beta_x^{(1)}\kappa_t^{(1)} + \beta_x^{(2)}\kappa_t^{(2)}$	$3n_a + 2n_y$
RH	$\alpha_x + \beta_x^{(1)}\kappa_t^{(1)} + \beta_x^{(0)}\gamma_c$	$3n_a + n_y + n_c$
APC	$\alpha_x + \kappa_t^{(1)} + \gamma_c$	$n_a + n_y + n_c$
CBD	$\kappa_t^{(1)} + (x - \bar{x})\kappa_t^{(2)}$	$2n_y$
M7	$\kappa_t^{(1)} + (x - \bar{x})\kappa_t^{(2)} + ((x - \bar{x})^2 - \sigma_x^2)\kappa_t^{(3)} + \gamma_c$	$3n_y + n_c$
sPLAT	$\alpha_x + \kappa_t^{(1)} + (\bar{x} - x)\kappa_t^{(2)} + \gamma_c$	$n_a + 2n_y + n_c$
cPLAT	$\alpha_x + \kappa_t^{(1)} + (\bar{x} - x)\kappa_t^{(2)} + (\bar{x} - x)^+\kappa_t^{(3)} + \gamma_c$	$n_a + 3n_y + n_c$

Generalised Age-Period-Cohort stochastic mortality models

Recent research has proposed a unifying framework discrete stochastic mortality models

- General Age-Period-Cohort model structure (Hunt and Blake 2015)
- Generalised (non-)linear model (Currie 2014)
- R Implementation of GAPC models (Villegas, Millossovich, and Kaishev 2018)

Generalised Age-Period-Cohort stochastic mortality models

1. Random Component:

$$D_{xt} \sim \text{Poisson}(E^c_{xt}\mu_{xt}) \quad \text{or} \quad D_{xt} \sim \text{Binomial}(E_{xt}, q_{xt})$$

2. Systematic Component:

$$\eta_{xt} \quad = \quad \alpha_x \quad + \quad \sum_{i=1}^{N} \beta_x^{(i)} \kappa_t^{(i)} \quad + \quad \beta_x^{(0)} \gamma_{t-x}$$

- Lee-Carter type $\rightsquigarrow \beta_x^{(i)}$, non-parametric
- CBD type $\rightsquigarrow \beta_x^{(i)} \equiv f^{(i)}(x)$, pre-specified parametric function

3. Link Function:

$$g\left(\mathbb{E}\left(\frac{D_{xt}}{E_{xt}}\right)\right) = \eta_{xt}$$

- log-Poisson: $\eta_{xt} = \log \mu_{xt}$
- logit-Binomial: $\eta_{xt} = \text{logit}\, q_{xt}$

Generalised Age-Period-Cohort stochastic mortality models

4. Set of parameter constraints:
   - Need parameters constraints to ensure identifiability
5. Forecasting and simulation
   - Period indexes: Multivariate random walk with drift

$$\boldsymbol{\kappa}_t = \boldsymbol{\delta} + \boldsymbol{\kappa}_{t-1} + \boldsymbol{\xi}_t^{\kappa}, \qquad \boldsymbol{\kappa}_t = \begin{pmatrix} \kappa_t^{(1)} \\ \vdots \\ \kappa_t^{(N)} \end{pmatrix}, \qquad \boldsymbol{\xi}_t^{\kappa} \sim N(\mathbf{0}, \Sigma),$$

   - Cohort effect: ARIMA(p, q, d) with drift

$$\Delta^d \gamma_c = \delta_0 + \phi_1 \Delta^d \gamma_{c-1} + \cdots + \phi_p \Delta^d \gamma_{c-p} + \epsilon_c + \delta_1 \epsilon_{c-1} + \cdots + \delta_q \epsilon_{c-q}$$

Mortality Modelling with R

Mortality modelling in R

- **Demography** (Hyndman 2014)
  - Download data from the Human Mortality Database
  - Lee-Carter model and several of its variants
- **ilc** (Butt, Haberman, and Shang 2014)
  - Lee-Carter with cohorts and Lee-Carter under a Poisson framework
- **Lifemetrics** (http://www.macs.hw.ac.uk/~andrewc/lifemetrics/)
  - CBD and extensions
  - Lee-Carter with cohorts and Lee-Carter under a Poisson framework
- **StMoMo** (Villegas, Millossovich, and Kaishev 2018)
  - Implements Generalised Age-Period-Cohort Models

Installation of R Packages

- To install the pacakges we use:

```r
install.packages("demography")
install.packages("StMoMo")
install.packages("lifecontingencies")
```

- To load within R:

```r
library(demography)
library(StMoMo)
library(lifecontingencies)
```

Downloading Data from the Human Mortality Database

- Human Mortality Database (http://www.mortality.org/) contains population mortality data for 39 countries under consistent data protocol and in consistent format
  - Very useful for cross-country comparisons
  - Need to register to obtain username and password
- Subregional versions available for:
  - USA (https://usa.mortality.org/)
  - Japan (http://www.ipss.go.jp/p-toukei/JMD/index-en.asp)
  - Australia (http://demography.cass.anu.edu.au/research/australian-human-mortality-database)
  - Canada (http://www.bdlc.umontreal.ca/CHMD/)
- For this example, we will use male USA data from 1950 to 2014

Australian Data from Human Mortality Database

```
library(demography)
USdata <- hmd.mx(country = "USA", username = "username",
                    password = "password")
```

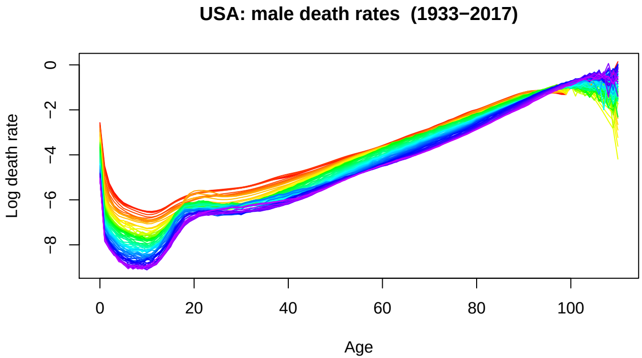
```
USdata
```

```
## Mortality data for AUS
##     Series: female male total
##     Years: 1921 - 2016
##     Ages:  0 - 110
```

USA Data from Human Mortality Database

```r
library(demography)
USdata <- hmd.mx(country = "USA", username = "username",
                 password = "password")
```

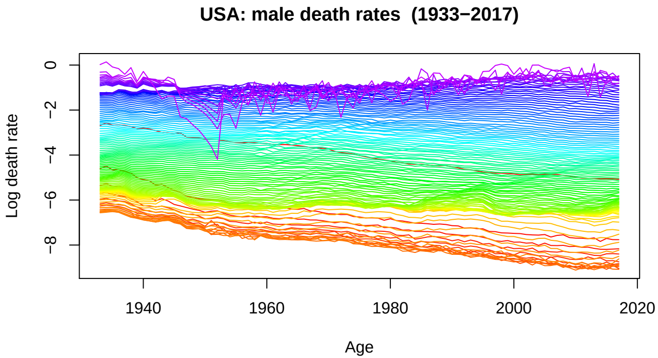
```r
plot(USdata, series = "male")
```



USA Data from Human Mortality Database

```r
library(demography)
USdata <- hmd.mx(country = "USA", username = "username",
                  password = "password")
```

```r
plot(USdata, series = "male", plot.type = "time")
```



StMoMo: **St**ochastic **Mo**rtality **Mo**delling

Who is MoMo?

StMoMo: Stochastic Mortality Modelling

Who is MoMo?

StMoMo: **St**ochastic **Mo**rtality **Mo**delling

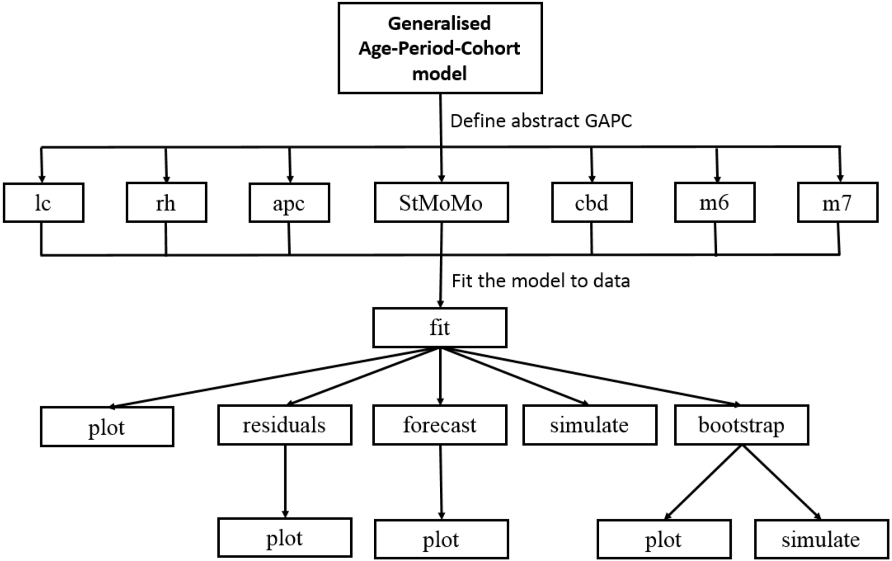
Who is MoMo?

Stochastic **Mo**rtality **Mo**delling

StMoMo: **St**ochastic **Mo**rtality **Mo**delling

Who is MoMo?

Overview of the structure of **StMoMo**



GAPC stochastic mortality models with **StMoMo**

Model	Predictor (η_{xt})
LC	$\alpha_x + \beta_x^{(1)}\kappa_t^{(1)}$
APC	$\alpha_x + \kappa_t^{(1)} + \gamma_{t-x}$
CBD	$\kappa_t^{(1)} + (x - \bar{x})\kappa_t^{(2)}$
M7	$\kappa_t^{(1)} + (x - \bar{x})\kappa_t^{(2)} + \left((x - \bar{x})^2 - \hat{\sigma}_x^2\right)\kappa_t^{(3)} + \gamma_{t-x}$

- For consistency, all under a log-Poisson setting:

$$D_{xt} \sim \mathrm{Poisson}(E_{xt}^c \mu_{xt})$$

$$\log \mu_{xt} = \eta_{xt}$$

Model definition: Predefined functions for common models

Model definition: Predefined functions for common models

```r
LC <- lc()
CBD <- cbd(link = "log")
APC <- apc()
M7 <- m7(link = "log")
```

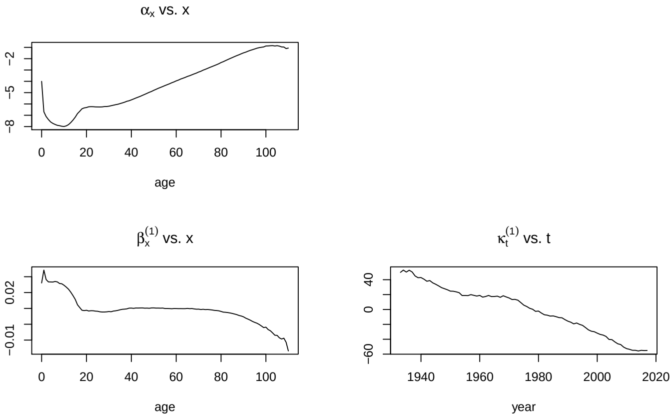
Model definition: Predefined functions for common models

```
LC <- lc()
CBD <- cbd(link = "log")
APC <- apc()
M7 <- m7(link = "log")
```

```
## Poisson model with predictor: log m[x,t] = a[x] + b1[x] k1[t]

## Poisson model with predictor: log m[x,t] = k1[t] + f2[x] k2[t]

## Poisson model with predictor: log m[x,t] = a[x] + k1[t] + g[t-x]

## Poisson model with predictor: log m[x,t] = k1[t] + f2[x] k2[t]
+ f3[x] k3[t] + g[t-x]
```

Model fitting: Lee-Carter Example

```r
USmale <- StMoMoData(USdata, series = "male")
LCfit <- fit(LC, data = USmale)
plot(LCfit)
```

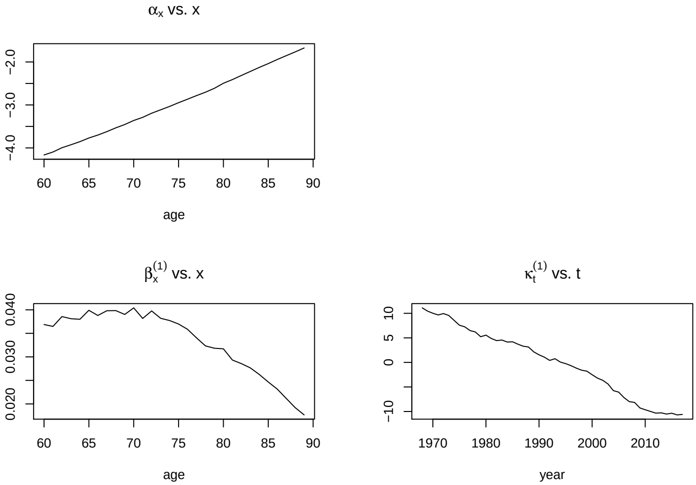


Model fitting: Concentrate on older ages

```r
#Ages for fitting
ages.fit <- 60:89
years.fit <- 1968:2017

#Fit all models
LCfit <- fit(LC, data=USmale, ages.fit=ages.fit,
             years.fit=years.fit)
CBDfit <- fit(CBD, data=USmale, ages.fit=ages.fit,
              years.fit=years.fit)
APCfit <- fit(APC, data=USmale, ages.fit=ages.fit,
              years.fit=years.fit)
M7fit <- fit(M7, data=USmale, ages.fit=ages.fit,
               years.fit=years.fit)
```

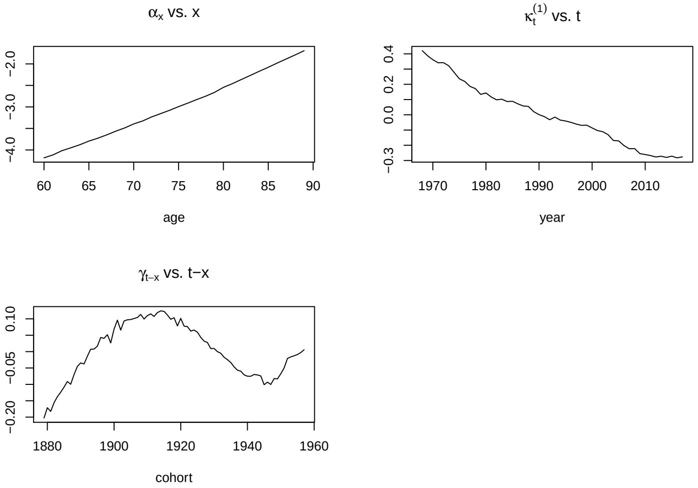
Parameter estimates – LC ($\log \mu_{xt} = \alpha_x + \beta_x^{(1)} \kappa_t^{(1)}$)

```r
plot(LCfit)
```



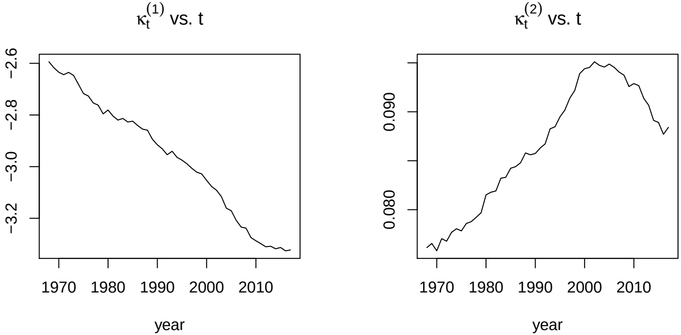
Parameter estimates – APC ($\log \mu_{xt} = \alpha_x + \kappa_t^{(1)} + \gamma_{t-x}$)

```r
plot(APCfit, parametricbx = FALSE)
```



Parameter estimates – CBD ($\log \mu_{xt} = \kappa_t^{(1)} + (x - \bar{x})\kappa_t^{(2)}$)

```r
plot(CBDfit, parametricbx = FALSE)
```



Goodness-of-fit: Deviance Residuals

$$r_{xt} = \operatorname{sign}(d_{xt} - \hat{d}_{xt})\sqrt{\frac{\operatorname{dev}(x,t)}{\hat{\phi}}}$$

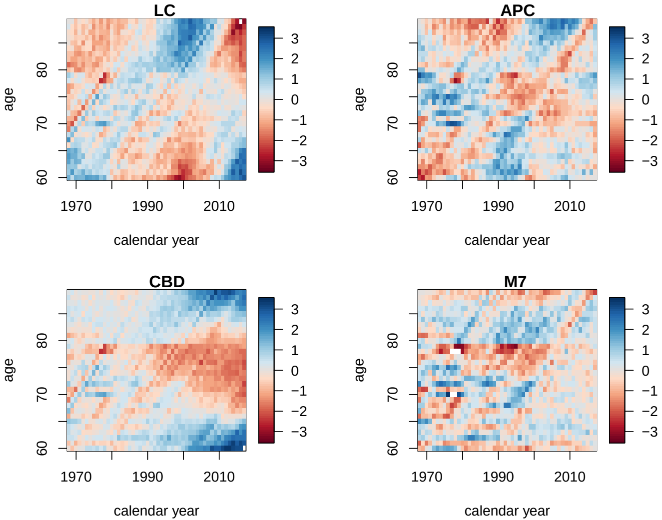
Goodness-of-fit: Deviance Residuals

$$r_{xt} = \mathrm{sign}(d_{xt} - \hat{d}_{xt})\sqrt{\frac{\mathrm{dev}(x,t)}{\hat{\phi}}}$$

```
#Compute residuals
LCres <- residuals(LCfit)
APCres <- residuals(APCfit)
CBDres <- residuals(CBDfit)
M7res <- residuals(M7fit)
```

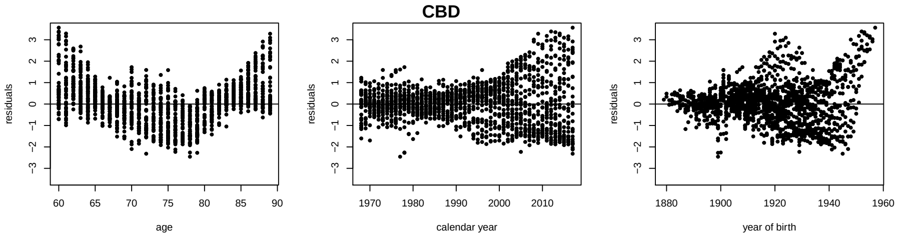
Goodness-of-fit: Residual heatmaps

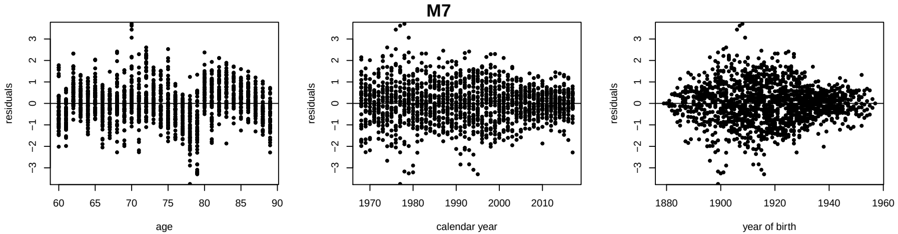
```
plot(LCres, type = "colourmap", reslim = c(-3.5, 3.5))
```



Goodness-of-fit: Residual scatterplots

```r
plot(CBDres, type = "scatter", reslim = c(-3.5, 3.5))
```





Goodness-of-fit vs Parsimony: AIC and BIC

$$AIC = -2\mathcal{L} + 2 \times \nu \qquad BIC = -2\mathcal{L} + \log K \times \nu$$

Goodness-of-fit vs Parsimony: AIC and BIC

$$AIC = -2\mathcal{L} + 2 \times \nu \qquad BIC = -2\mathcal{L} + \log K \times \nu$$

```
#Compute residuals
AIC(CBDfit)
```

```
## [1] 74754.48
```

```
BIC(CBDfit)
```

```
## [1] 75285.8
```

Goodness-of-fit vs Parsimony: AIC and BIC

$$AIC = -2\mathcal{L} + 2 \times \nu \qquad BIC = -2\mathcal{L} + \log K \times \nu$$

```r
#Compute residuals
AIC(CBDfit)
```

```
## [1] 74754.48
```

```r
BIC(CBDfit)
```

```
## [1] 75285.8
```

Criterion	LC	APC	CBD	M7
AIC	56814	32063	74754	**27575**
BIC	57388	32892	75286	**28775**

Forecasting and simulation

- **Period indexes:** Multivariate random walk with drift

$$\boldsymbol{\kappa}_t = \boldsymbol{\delta} + \boldsymbol{\kappa}_{t-1} + \boldsymbol{\xi}^{\kappa}_t, \qquad \boldsymbol{\kappa}_t = \begin{pmatrix} \kappa^{(1)}_t \\ \vdots \\ \kappa^{(N)}_t \end{pmatrix}, \qquad \boldsymbol{\xi}^{\kappa}_t \sim N(\mathbf{0}, \Sigma),$$

- **Cohort effect:** ARIMA(p, q, d) with drift

$$\Delta^d \gamma_c = \delta_0 + \phi_1 \Delta^d \gamma_{c-1} + \cdots + \phi_p \Delta^d \gamma_{c-p} + \epsilon_c + \delta_1 \epsilon_{c-1} + \cdots + \delta_q \epsilon_{c-q}$$

Forecasting

Model	Model for γ_{t-x}
APC	ARIMA$(1, 1, 0)$ with zero mean
M7	ARIMA$(0, 0, 0)$ with zero mean

Forecasting

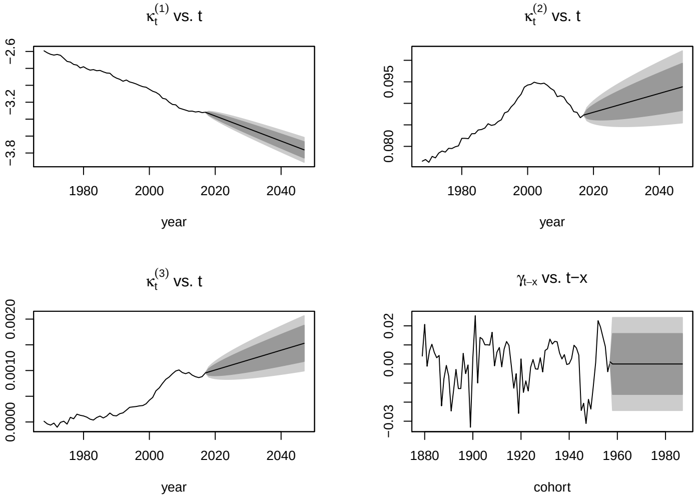
Model	Model for γ_{t-x}
APC	ARIMA$(1, 1, 0)$ with zero mean
M7	ARIMA$(0, 0, 0)$ with zero mean

30-year ahead ($h = 30$) central projections: period indexes, cohort index, and death rates probabilities:

```
LCfor <- forecast(LCfit, h=30)
APCfor <- forecast(APCfit, h=30, gc.order = c(1,1,0),
                   gc.include.constant = FALSE)
CBDfor <- forecast(CBDfit, h=30)
M7for <- forecast(M7fit, h=30, gc.order = c(0,0,0),
                  gc.include.constant = FALSE)
```

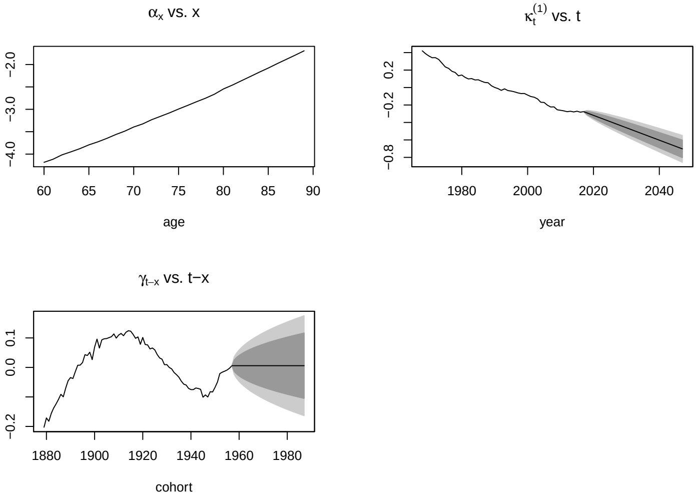
Forecasted period and cohort indexes

```r
plot(M7for, parametricbx = FALSE)
```



Forecasted period and cohort indexes

```r
plot(APCfor, parametricbx = FALSE)
```

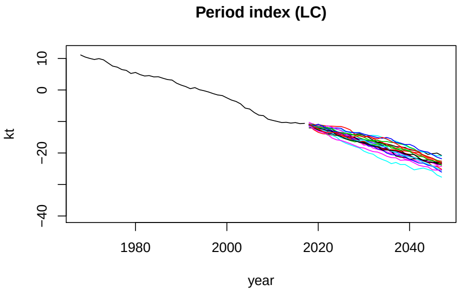


Simulation

```
LCsim <- simulate(LCfit, nsim=1000, h=30)
APCsim <- simulate(APCfit, nsim=1000, h=30,
                   gc.order = c(1,1,0),
                   gc.include.constant = FALSE)
CBDsim <- simulate(CBDfit, nsim=1000, h=30)
M7sim <- simulate(M7fit, nsim=1000, h=30,
                  gc.order = c(0,0,0),
                  gc.include.constant = FALSE)
```

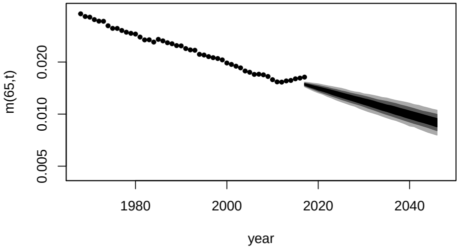
Simulation

```
#Plot period index trajectories for the LC model
plot(LCfit$years, LCfit$kt[1,], type="l", xlim=c(1968,2047),
     ylim=c(-40,12), xlab="year", ylab="kt",
     main="Period index (LC)")
matlines(LCsim$kt.s$years, LCsim$kt.s$sim[1,,1:20], type="l",
         lty=1)
```

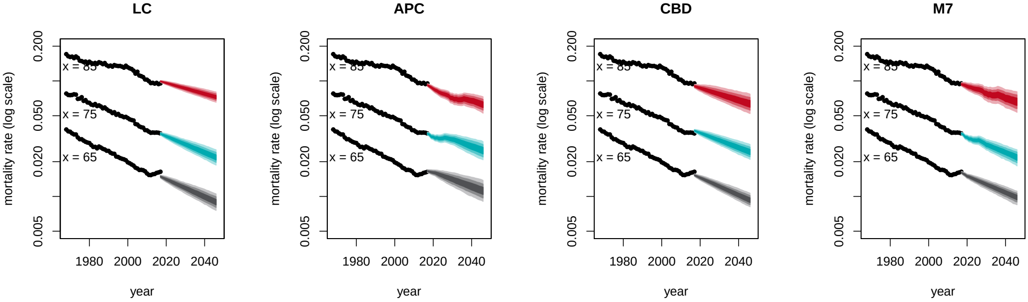


Fancharts

```
library(fanplot)
mxt <- LCfit$Dxt/LCfit$Ext
plot(LCfit$years, mxt["65",], pch=20, xlim=c(1968,2047),
     ylim=c(0.0045,0.04), log="y", xlab="year", ylab="m(65,t)")
fan(t(LCsim$rates["65",,]), start=2017, probs=c(50,75,90,97.5),
  n.fan=4, ln=NULL, fan.col=colorRampPalette(c("black","white")))
```



Fancharts



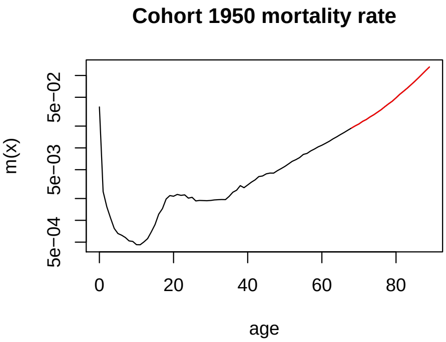
LC
APC
CBD
M7
mortality rate (log scale)
0.005
0.020
0.050
0.200
x = 85
x = 75
x = 65
1980
2000
2020
2040
year

Obtaining projected life tables for a cohort

```r
chosen_cohort <- 1950

#observed rates for ages 0-59
hist_rates <- extractCohort(USmale$Dxt/USmale$Ext,
                            cohort = chosen_cohort)[1:60]

#fitted historical rates for ages 60-67
lc_fit_rates <- extractCohort(fitted(LCfit, type = "rates"),
                              cohort = chosen_cohort)
#forcasted rates for ages 68-89
lc_for_rates <- extractCohort(LCfor$rates,
                             cohort = chosen_cohort)
#all rates
lc_rates_1950 <- c(hist_rates,lc_fit_rates, lc_for_rates)
```

Obtaining projected life tables for a cohort

```
plot(0:89, lc_rates_1950 ,type = "l", log = "y", xlab = "age",
     ylab = "m(x)", main = "Cohort 1950 mortality rate")
lines(68:89, lc_for_rates, col = "red")
```



Computing life expectancies using **lifecontingencies**

```
#Transform mx to qx
lc_qx_1950 <- mx2qx(lc_rates_1950)
#Create lifetable
lc_lifetable_1950 <- probs2lifetable(probs = lc_qx_1950,
                                     type = "qx",
                                     name = "LC-1950")
exn(lc_lifetable_1950,x=65)
```

```
## [1] 17.373
```

Computing life expectancies using **lifecontingencies**

```
#Transform mx to qx
lc_qx_1950 <- mx2qx(lc_rates_1950)
#Create lifetable
lc_lifetable_1950 <- probs2lifetable(probs = lc_qx_1950,
                                     type = "qx",
                                     name = "LC-1950")
exn(lc_lifetable_1950,x=65)
```

```
## [1] 17.373
```

Model	LC	APC	CBD	M7
LE at age 65	17.3729974	17.4886457	17.254208	17.5034425
LE at age 75	10.3314105	10.6308337	10.3993284	10.5521274

Summary

- This has been a whistle-stop introduction to fitting mortality models in R
- Useful example for using **StMoMo** in the package vignette at https://cran.r-project.org/web/packages/StMoMo/vignettes/StMoMoVignette.pdf
  - Also, references within this useful for understanding more about mortality models
- Useful examples of integrating **StMoMo** with **lifecontingencies** at https://rdrr.io/cran/lifecontingencies/f/inst/doc/mortality_projection.pdf

Summary

- **StMoMo provides** easy implementation and comparison of a wide range of models making it useful for:
  - Actuaries analysing longevity risk $\rightsquigarrow$ Model risk
  - Use in the classroom
- Standard packages are starting point – however, to go beyond them you need to understand the principles behind their operation

Work in progress

- Available in the development version and expected to be released in next six months
  - Selection of models using cross-validation
  - Construction of models using regularisation techniques
- Sister packages **iMoMo** for mortality improvement rate modelling
  - Expected to be released in the next year
- Other development plans
  - Multipopulation models
  - Model combination

Cross validation and regularisation

Mortality Models: Key research questions

1. What model features are desired for different applications?

2. Why do we only consider a fixed set of models?

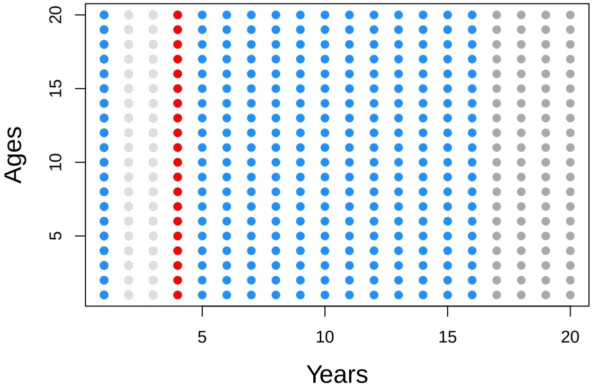
3. How can we be confident we have selected the best model?

Objective

Provide a comprehensive framework to **construct**, **select**, and **evaluate** discrete-time mortality models for forecasting applications, using various **statistical learning** and predictive analytics techniques.

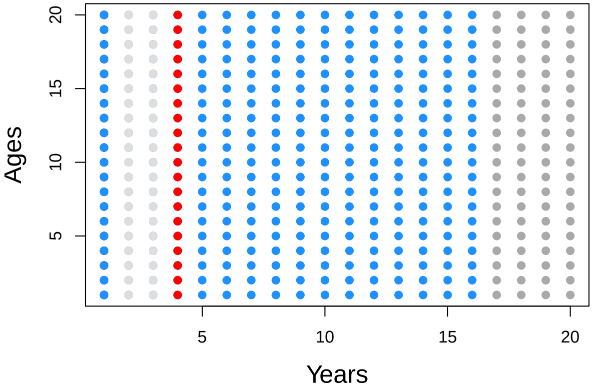
- **Construction** based on regularisation techniques
- **Selection** based on cross-validation techniques

Selection of mortality models using cross validation



- **Training/Test Set**
- **Test Set Width**: Depending on forecasting Horizon
- **Metric**: MSE on the test set

Selection of mortality models using cross validation



- **Training/Test Set**
- **Test Set Width**: Depending on forecasting Horizon
- **Metric**: MSE on the test set

Selection of mortality models using cross validation

```r
# install.packages("devtools")
# devtools::install_github("amvillegas/StMoMo",
#                                  ref = "GroupLasso")

#CV for 1 year ahead
LCcv1 <- cv.StMoMo(LC, h = 1, data=USmale, ages.train=ages.fit,
            years.train=years.fit, type = "lograte")
APCcv1 <- cv.StMoMo(APC, h = 1, data=USmale, ages.train=ages.fit,
            years.train=years.fit, type = "lograte")
CBDcv1 <- cv.StMoMo(CBD, h = 1, data=USmale, ages.train=ages.fit,
            years.train=years.fit, type = "lograte")
M7cv1 <- cv.StMoMo(M7, h = 1, data=USmale, ages.train=ages.fit,
            years.train=years.fit, type = "lograte")
```

Selection of mortality models using cross validation

```
#CV for 10 year ahead
LCcv10 <- cv.StMoMo(LC, h = 10, data=USmale, ages.train=ages.fit,
                    years.train=years.fit, type = "lograte s")
APCcv10 <- cv.StMoMo(APC, h = 10, data=USmale, ages.train=ages.fit,
                    years.train=years.fit, type = "logrates")
CBDcv10 <- cv.StMoMo(CBD, h = 10, data=USmale, ages.train=ages.fit,
                    years.train=years.fit, type = "logrates")
M7cv10 <- cv.StMoMo(M7, h = 10, data=USmale, ages.train=ages.fit,
                    years.train=years.fit, type = "logrates")
```

Selection of mortality models using cross validation

```
LCcv1$cv.mse
```

```
## [1] 0.001367588
```

```
LCcv10$cv.mse
```

```
## [1] 0.004095502
```

Selection of mortality models using cross validation

```
LCcv1$cv.mse
```

```
## [1] 0.001367588
```

```
LCcv10$cv.mse
```

```
## [1] 0.004095502
```

Criterion	LC	APC	CBD	M7
AIC	56814	32063	74754	**27575**
BIC	57388	32892	75286	**28775**
CV 1-year	0.0014	0.0007	0.0017	**0.0005**
CV 10-year	0.0041	**0.0033**	0.0050	0.0046

Construction: "Formalised" Model-Building Framework

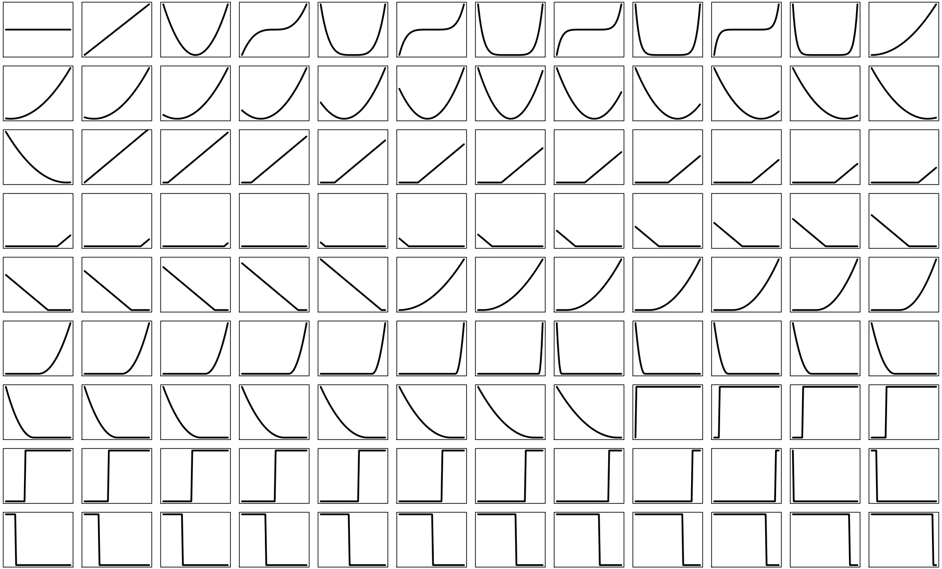
We start with a **huge** model,

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

where the suite of basis functions $(f^{(i)}(x))$ included are:

$$f^{(i)}(x) = \begin{cases} 1, & \text{Unit} \\ (x-\bar{x})^n & \text{Polynomial} \\ (x-n)^+ & \text{Call} \\ (n-x)^+ & \text{Put} \\ 1_{x<n} & \text{Below} \\ 1_{x>n} & \text{Above} \end{cases} \quad . \tag{1}$$

Construction: Our Model



Construction: GLM Representation

$$\eta_{x,t} = \ln(\mu_{x,t}) = \alpha_x + \sum_{i=1}^{B} f^{(i)}(x)\kappa_t^{(i)} + \gamma_c,$$

can be expressed as a GLM (Currie 2014),

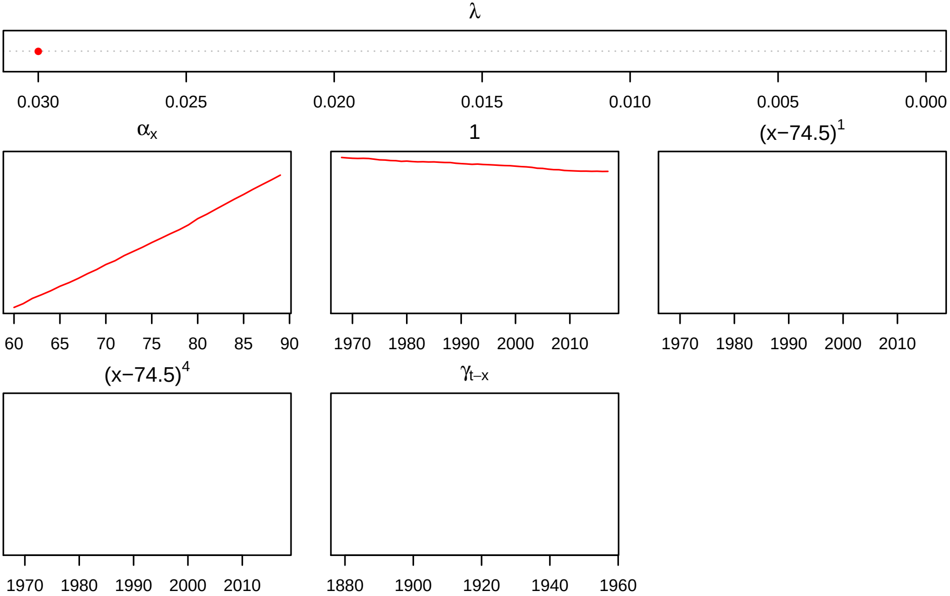
$$\eta = \mathbf{X}\beta = \sum_{j=0}^{B+1} \mathbf{X}_j \boldsymbol{\beta}_j, \qquad \mathbf{X} = [\mathbf{X_0} : \mathbf{X_1} : \mathbf{X_2} : \ldots : \mathbf{X_B} : \mathbf{X_{B+1}}],$$

where,

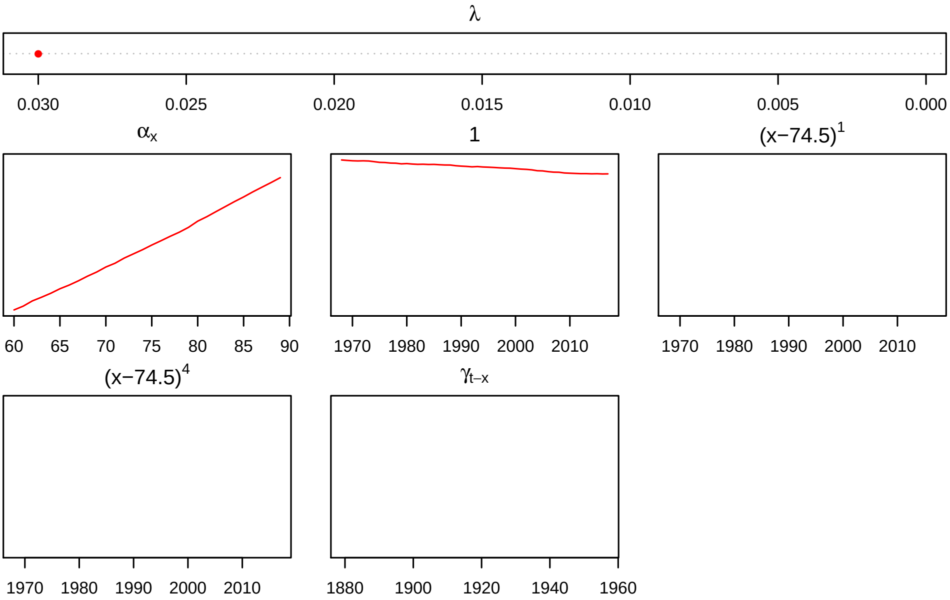
$$\boldsymbol{\beta} = \{\boldsymbol{\beta_i}\}_{i=0}^{B+1}, \quad \boldsymbol{\beta_0} = \{\alpha_x\}_{x=1}^{n_x}, \quad \boldsymbol{\beta_i} = \{\kappa_t^{(i)}\}_{t=1}^{n_t}, \quad \boldsymbol{\beta_{B+1}} = \{\gamma_c\}_{c=1}^{n_c}.$$

- Estimate parameters with group lasso using the R package **grpreg** (Breheny and Huang 2013)

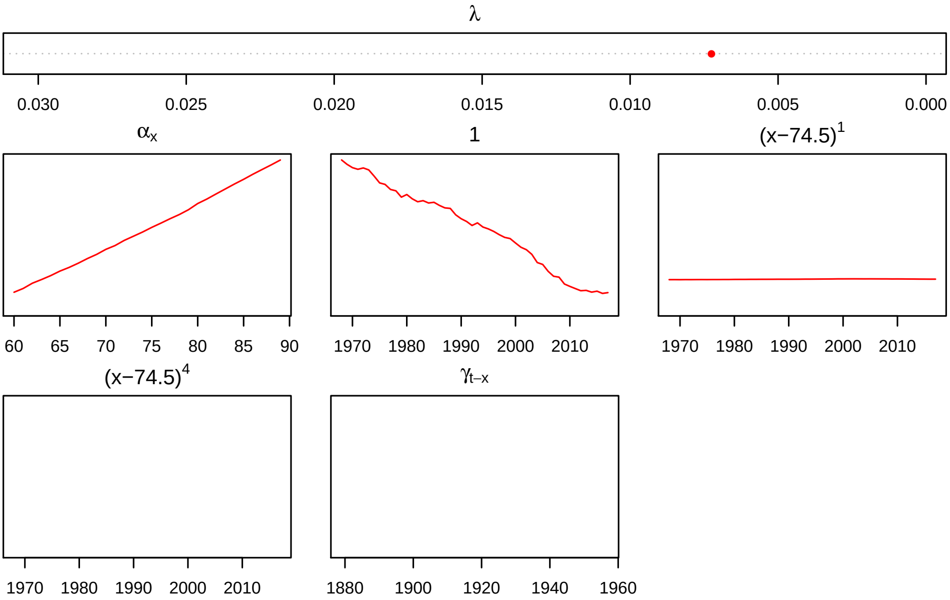
Construction: Estimated Parameters (USA Males)



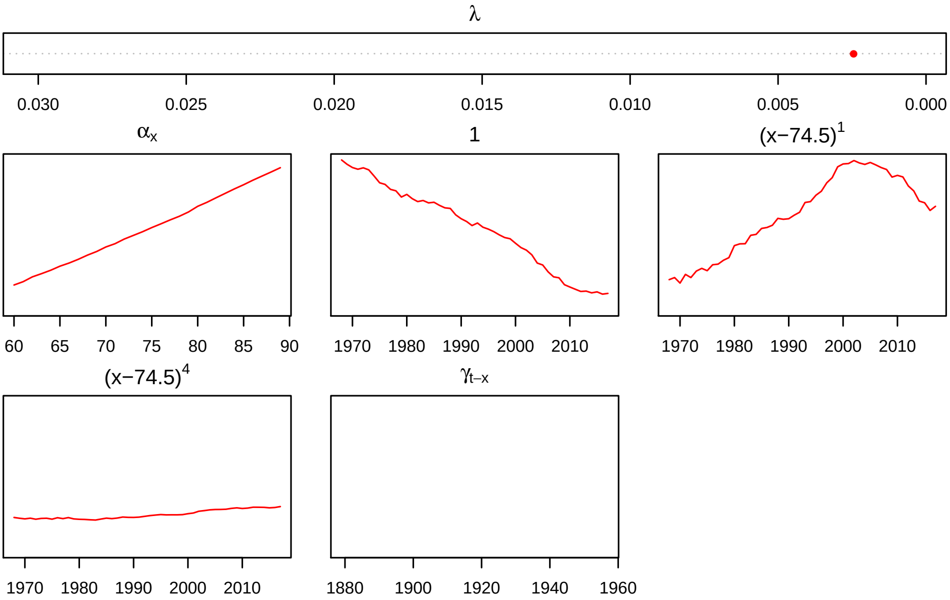
Construction: Estimated Parameters (USA Males)



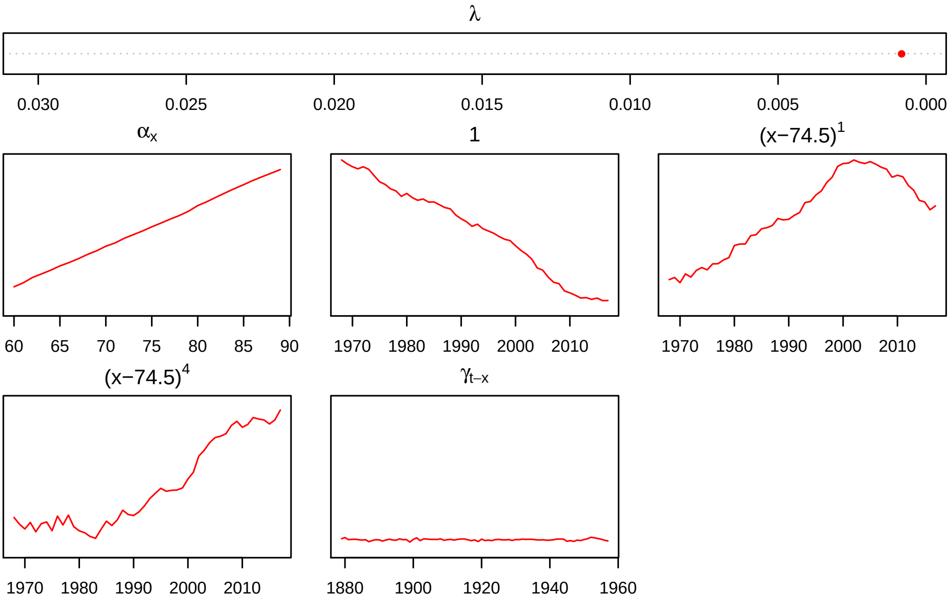
Construction: Estimated Parameters (USA Males)



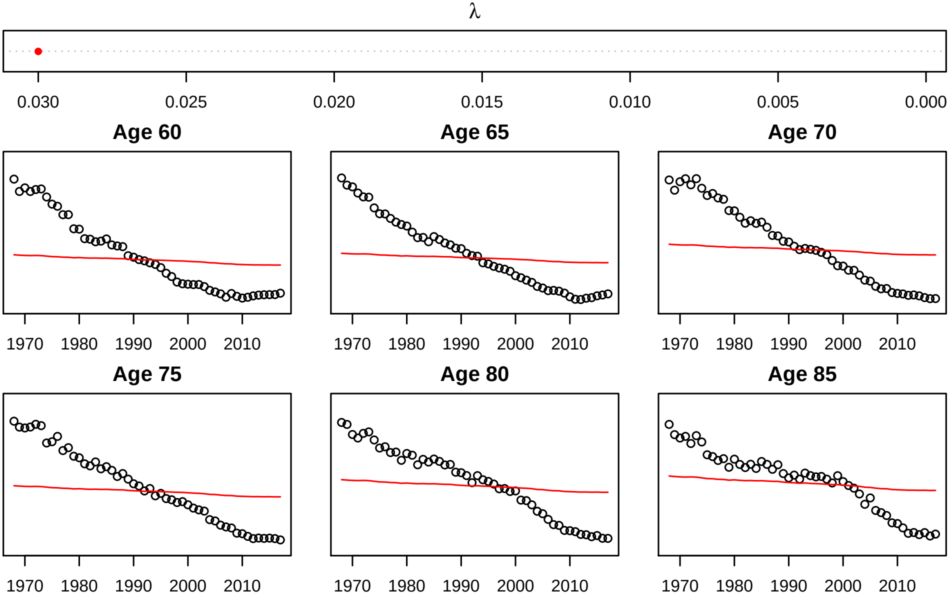
Construction: Estimated Parameters (USA Males)



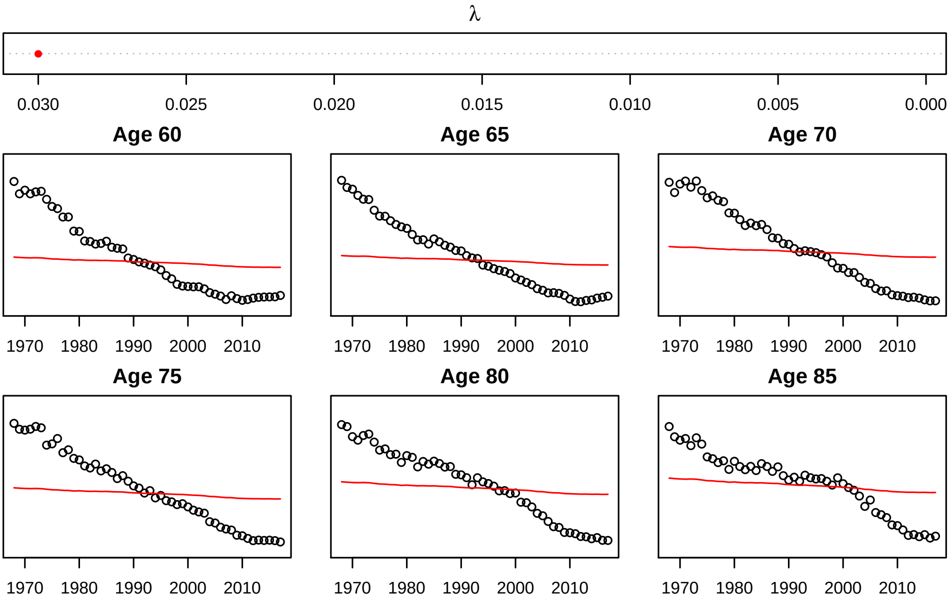
Construction: Estimated Parameters (USA Males)



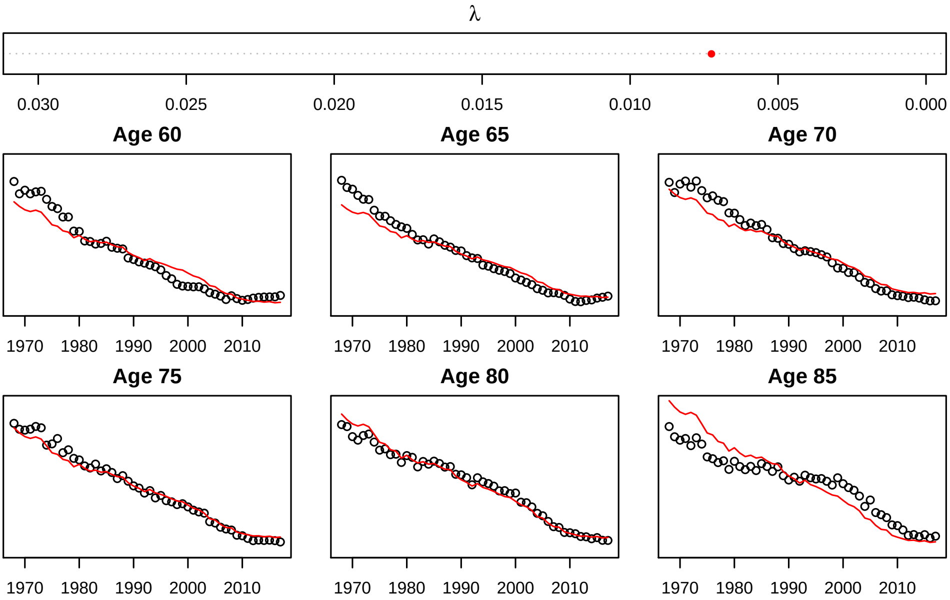
Construction: Fitted Values (USA Males)



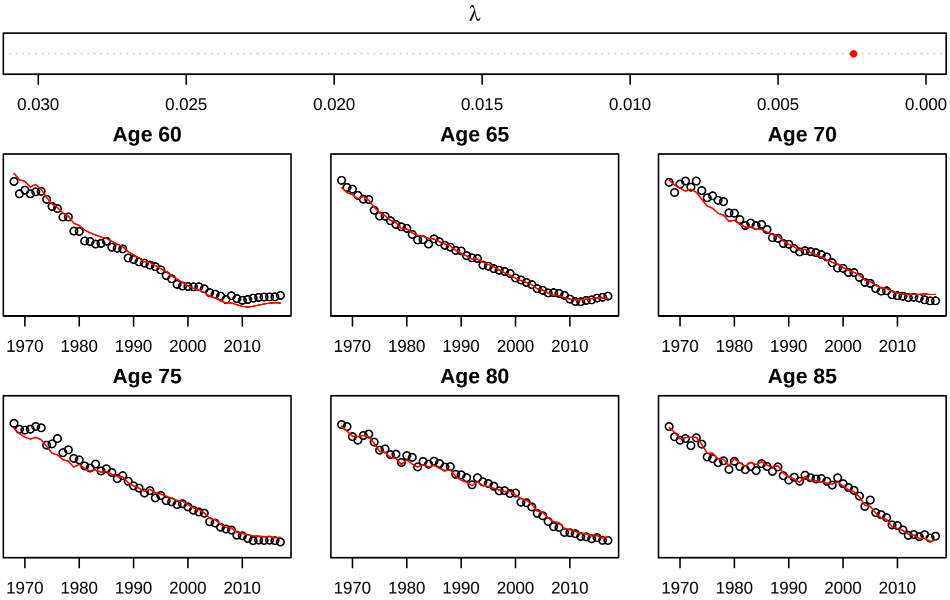
Construction: Fitted Values (USA Males)



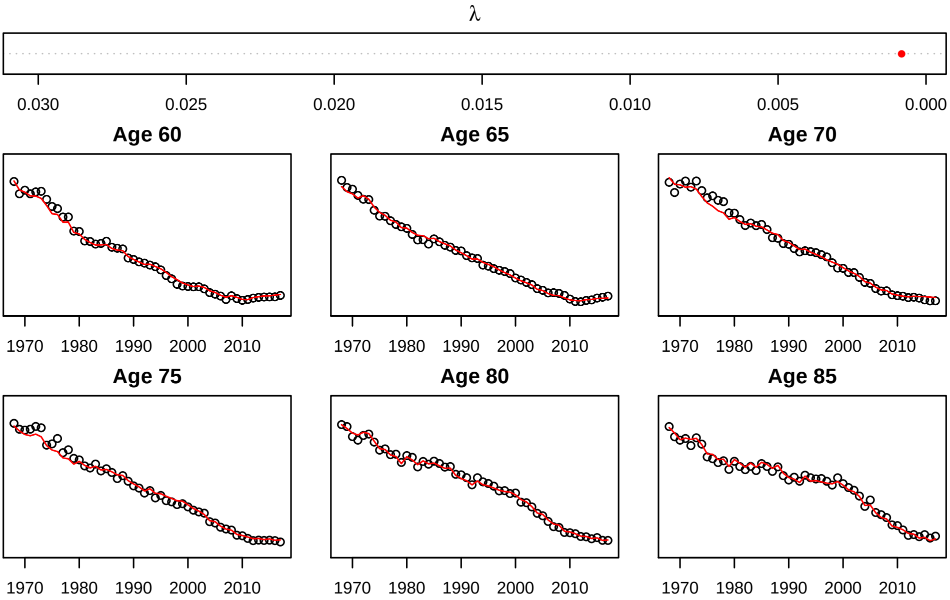
Construction: Fitted Values (USA Males)



Construction: Fitted Values (USA Males)



Construction: Fitted Values (USA Males)



Implementation in StMoMo - Define model

```
#Create big model
strikes <- seq(65,85,5)
bModel <- StMoMo(link = "log-Gaussian", staticAgeFun = TRUE,
                periodAgeFun = c("1", genPoly(1:5), genCall(strikes),
                                 genPut(strikes)),
                cohortAgeFun = "1")
bModel
```

```
## Gaussian model with predictor: log m[x,t] = a[x] + k1[t] +
f2[x] k2[t] + f3[x] k3[t] + f4[x] k4[t] + f5[x] k5[t] + f6[x]
k6[t] + f7[x] k7[t] + f8[x] k8[t] + f9[x] k9[t] + f10[x] k10[t] +
f11[x] k11[t] + f12[x] k12[t] + f13[x] k13[t] + f14[x] k14[t] +
f15[x] k15[t] + f16[x] k16[t] + g[t-x]
```

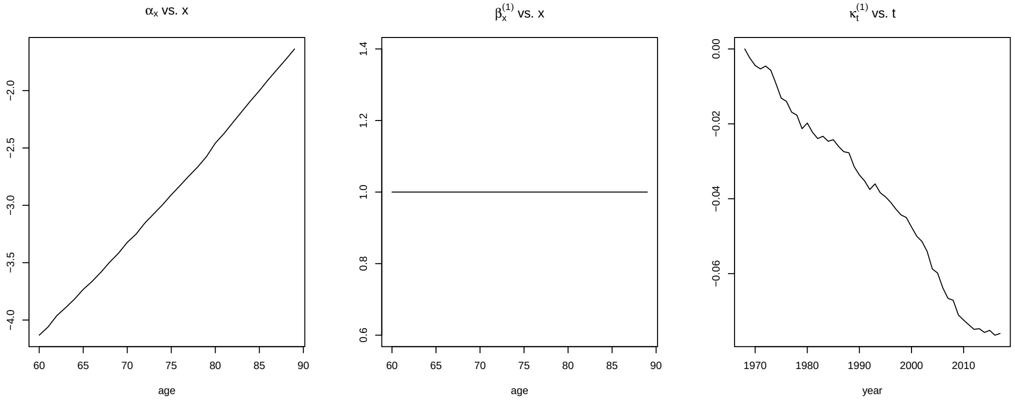
Implementation in StMoMo - Fit model with grouped penalised regularisation

```r
#lambda grid
lambda <- exp(seq(log(0.03), log(0.0005), length.out = 50))
#Penalise all term
bMgrpfit <- grpfit(bModel, lambda = lambda, data = USmale,
                   ages.fit = ages.fit, years.fit = years.fit)
```

```
## StMoMo: Start fitting with grpreg
## StMoMo: Finish fitting with grpreg
```

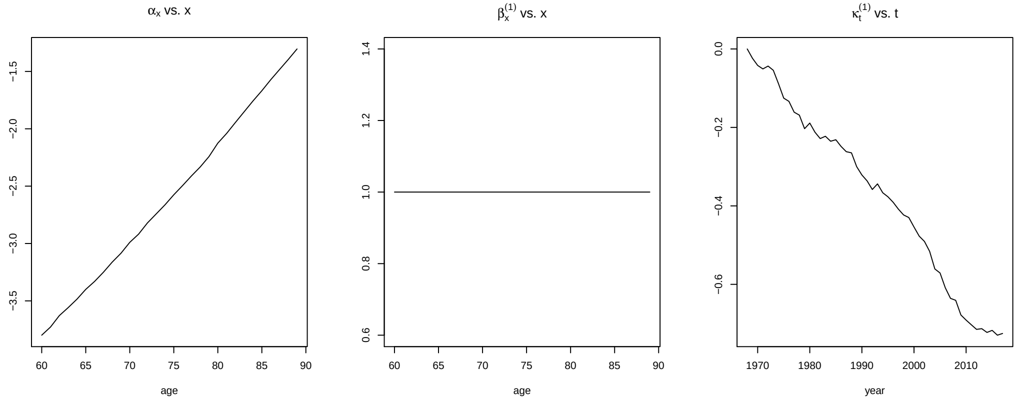
Implementation in StMoMo - Parameter paths

```r
plot(extractStMoMo(bMgrpfit, 1), nCol = 3)
```



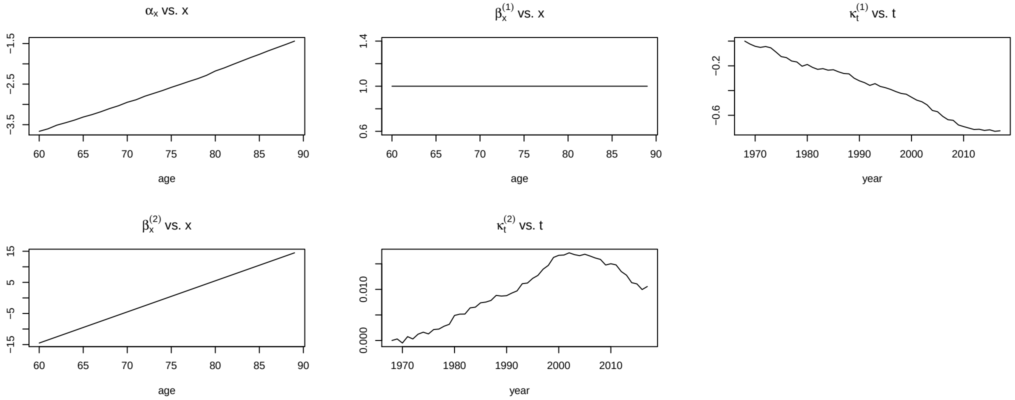
Implementation in StMoMo - Parameter paths

```r
plot(extractStMoMo(bMgrpfit, 17), nCol = 3)
```



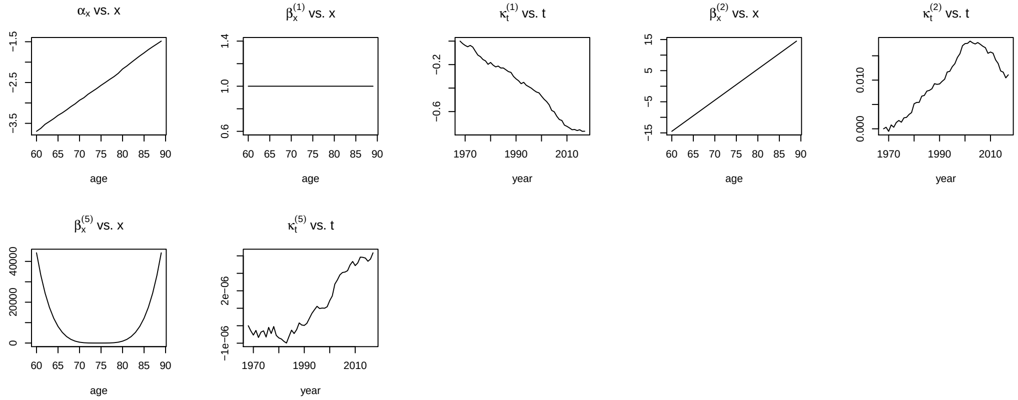
Implementation in StMoMo - Parameter paths

```r
plot(extractStMoMo(bMgrpfit, 30), nCol = 3)
```



Implementation in StMoMo - Parameter paths

```r
plot(extractStMoMo(bMgrpfit, 40), nCol = 5)
```



Improvement rate modelling

The APCi model

- Modelling improvement rates rather than mortality rates is becoming commong
- Consider the APCi model which corresponds broadly to the lates CMI projection approach:

$$-\log \frac{\mu_{xt}}{\mu_{x,t-1}} = \alpha_x + \kappa_t + \gamma_{t-x}$$

Implementation in **iMoMo**

```r
library(iMoMo) #Not yet publicly available
#Define model
APCi <- apci()
APCi
```

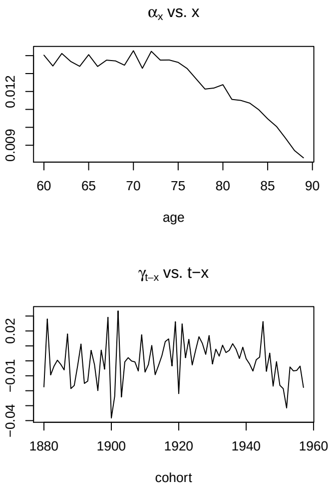
```
## indirect model with predictor: eta[x,t] = a[x] + k1[t] + g[t-x]
```

```r
#fit the model
APCifit <- fit(APCi, data=USmale, ages.fit=ages.fit,
               years.fit=years.fit)
```

```
## StMoMo: Start fitting with gnm
## StMoMo: Finish fitting with gnm
```

Implementation in **iMoMo**

```
plot(APCifit, parametricbx = FALSE)
```



http://cran.r-project.org/web/packages/StMoMo/
https://github.com/amvillegas/StMoMo

Thank you!

a.villegas@unsw.edu.au

andresmauriciovillegas@gmail.com

ARC CENTRE OF EXCELLENCE IN POPULATION AGEING RESEARCH

UNSW | AGSM Business School

References I

Alai, Daniel H., and Michael Sherris. 2014. “Rethinking age-period-cohort mortality trend models.” *Scandinavian Actuarial Journal*, no. 3: 208–27.

Aro, Helena, and Teemu Pennanen. 2011. “A user-friendly approach to stochastic mortality modelling.” *European Actuarial Journal* 1: 151–67.

Booth, Heather, and Leonie Tickle. 2008. “Mortality modelling and forecasting: A review of methods.” *Annals of Actuarial Science* 1 (2). Faculty of Actuaries; Institute of Actuaries: 3–43.

Börger, Matthias, Daniel Fleischer, and Nikita Kuksin. 2013. “Modeling the mortality trend under modern solvency regimes.” *ASTIN Bulletin* 44 (1): 1–38.

Breheny, Patrick, and Jian Huang. 2013. “Group descent algorithms for nonconvex penalized linear and logistic regression models with grouped predictors.” *Statistics and Computing* 25 (2): 173–87.

References II

doi:10.1007/s11222-013-9424-2.

Brouhns, Natacha, Michel Denuit, and Ingrid Van Keilegom. 2005. "Bootstrapping the Poisson log-bilinear model for mortality forecasting." *Scandinavian Actuarial Journal*, no. 3: 212–24.

Butt, Zoltan, Steven Haberman, and Han Lin Shang. 2014. *ilc: Lee-Carter Mortality Models using Iterative Fitting Algorithms*. http://cran.r-project.org/package=ilc.

Cairns, Andrew J.G., David Blake, and Kevin Dowd. 2006. "A two-factor model for stochastic mortality with parameter uncertainty: theory and calibration." *Journal of Risk and Insurance* 73 (4): 687–718.

Cairns, Andrew J.G., David Blake, Kevin Dowd, Guy 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."

References III

North American Actuarial Journal 13 (1): 1–35.

Currie, Iain D. 2014. "On fitting generalized linear and non-linear models of mortality." *Scandinavian Actuarial Journal*.

Hunt, Andrew, and David Blake. 2015. "On the Structure and Classification of Mortality Models." *Pension Institute Working Paper*. http://www.pensions-institute.org/workingpapers/wp1506.pdf.

Hyndman, Rob J. 2014. *demography: Forecasting mortality, fertility, migration and population data*. http://cran.r-project.org/package=demography.

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

Lee, Ronald D., and Lawrence R. Carter. 1992. "Modeling and forecasting U.S. mortality." *Journal of the American Statistical Association* 87 (419):

659–71.

O'Hare, Colin, and Youwei Li. 2012. "Explaining young mortality." *Insurance: Mathematics and Economics* 50 (1): 12–25.

Pitacco, Ermanno. 2004. "Survival models in a dynamic context: a survey." *Insurance: Mathematics and Economics* 35 (April): 279–98. doi:10.1016/j.insmatheco.2004.04.001.

Pitacco, Ermanno, Michel Denuit, Steven Haberman, and Annamaria Olivieri. 2009. *Modelling longevity dynamics for pensions and annuity business*. Oxford: Oxford University Press.

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

Renshaw, A.E., and Steven Haberman. 2003. "Lee-Carter mortality forecasting with age-specific enhancement." *Insurance: Mathematics and*

References V

Economics 33 (2): 255–72.

———. 2006. "A cohort-based extension to the Lee-Carter model for mortality reduction factors." *Insurance: Mathematics and Economics* 38 (3): 556–70.

Villegas, Andrés M., Pietro Millossovich, and Vladimir K Kaishev. 2018. "StMoMo: An R Package for Stochastic Mortality Modelling." *Journal of Statistical Software* 84 (3). doi:10.18637/jss.v084.i03.

Wong-Fupuy, C., and Steven Haberman. 2004. "Projecting mortality trends: recent developments in the United Kingdom and the United States." *North American Actuarial Journal* 8 (2). SOCIETY OF ACTUARIES: 56–83.