IDEA-LR: Frequentist Estimation

Jackie Wong

2024-03-01

In this document, I will briefly illustrate the results of fitting various extensions of Lee-Carter mortality models using a **classical/frequentist** approach. This is so that we have a clearer idea of what to expect from each variation of the models being considered to help us better understand the model, mechanics, parameter interpretations, etc. Might be useful for Bayesian implementations later.

The data

```
setwd("G:/IDEA-LR/") #set working directory
load("data_summarised.rda") #load the data

head(data_summarised)
```

```
##   Product Age Year Exposure  Claim   ExpClaim          Qx        ExpQx
## 1     ACI  18 2016 186.4372 0.0000 0.09689145 0.000000000 0.0005197003
## 2     ACI  18 2017 224.4137 1.0331 0.11851849 0.004603551 0.0005281250
## 3     ACI  18 2018 232.0658 0.0000 0.12093372 0.000000000 0.0005211183
## 4     ACI  18 2019 162.4630 0.0000 0.08201585 0.000000000 0.0005048278
## 5     ACI  18 2020 158.8978 0.0000 0.08213127 0.000000000 0.0005168811
## 6     ACI  19 2016 691.6257 1.0094 0.36785802 0.001459460 0.0005318744
##         StdQx
## 1 0.000000000
## 2 0.004518766
## 3 0.000000000
## 4 0.000000000
## 5 0.000000000
## 6 0.001451588
```

```
dim(data_summarised)
```

```
## [1] 1278    9
```

```
names(data_summarised)
```

```
## [1] "Product"  "Age"      "Year"     "Exposure" "Claim"    "ExpClaim" "Qx"
## [8] "ExpQx"    "StdQx"
```

```r
table(data_summarised$Product) # 4 products
```

```
##
##       ACI Annuities        DB       SCI
##       318       285       365       310
```

```r
#stratify the data by product type
data_summarised_ACI<-data_summarised[data_summarised$Product=="ACI",]
data_summarised_DB<-data_summarised[data_summarised$Product=="DB",]
data_summarised_SCI<-data_summarised[data_summarised$Product=="SCI",]
data_summarised_Annuities<-data_summarised[data_summarised$Product=="Annuities",]

table(data_summarised_ACI$Age)
```

```
##
## 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43
##  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5
## 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69
##  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5
## 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84
##  5  5  5  5  5  5  5  3  3  3  3  4  4  2  1
```

```r
table(data_summarised_DB$Age)
```

```
##
## 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43
##  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5
## 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69
##  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5
## 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90
##  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5
```

```r
table(data_summarised_SCI$Age)
```

```
##
## 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43
##  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5
## 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69
##  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5
## 70 71 72 73 74 75 76 77 78 79 80 83 84
##  5  5  5  5  5  5  5  4  4  4  1  1  1
```

```r
table(data_summarised_Annuities$Age) #annuities data does not have much data by age
```

```
##
##  20  21  22  23  24  25  26  27  28  29  30  31  32  33  34  35  36  37  38  39
##   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1
##  40  41  42  43  44  45  46  47  48  49  50  51  52  53  54  55  56  57  58  59
##   1   1   1   1   1   1   1   1   1   1   5   5   5   5   5   5   5   5   5   5
##  60  61  62  63  64  65  66  67  68  69  70  71  72  73  74  75  76  77  78  79
```

```
##   5   5   5   5   5   5   5   5   5   5   5   5   5   5   5   5   5   5   5   5
##  80  81  82  83  84  85  86  87  88  89  90  91  92  93  94  95  96  97  98  99
##   5   5   5   5   5   5   5   5   5   5   5   5   5   5   5   5   5   5   5   5
## 100
##   5
```

We see that for some ages, particularly at both ends, are quite sparse (makes sense from actuarial perspective). Moreover, for annuities, we observe more data at higher ages, again makes actuarial sense.

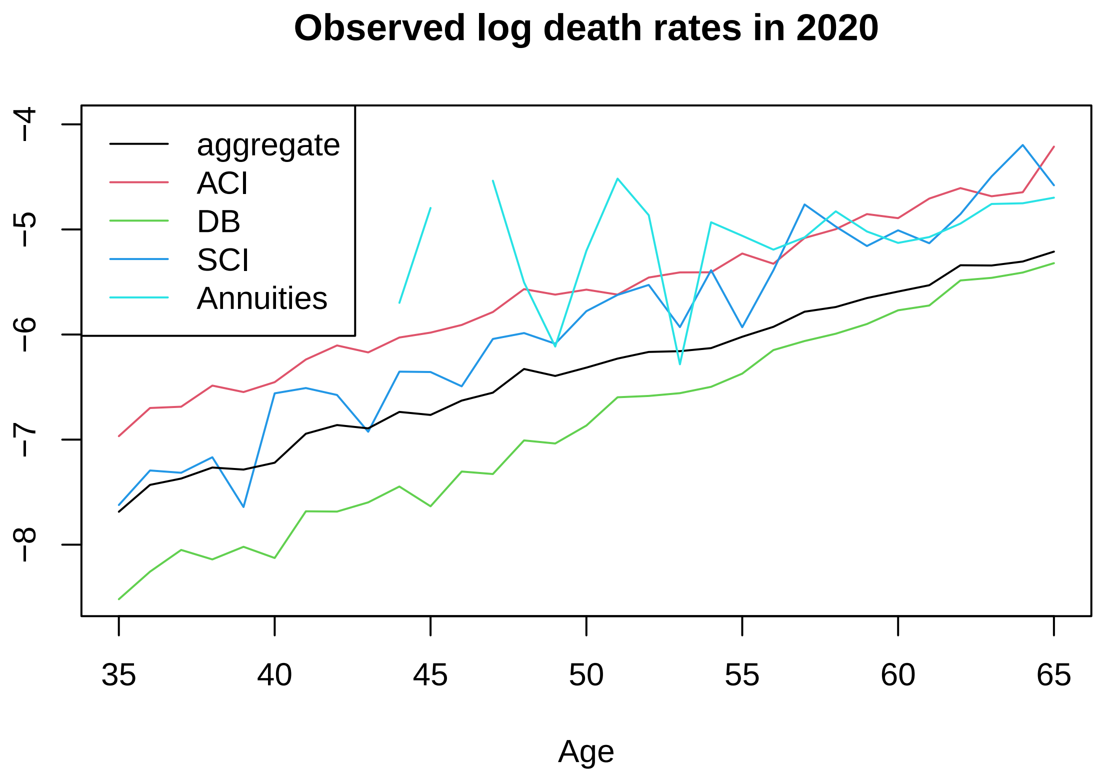
In order to avoid having to deal with ages with missing yearly deaths, we will extract data corresponding to ages 35-65 where we have sufficient information. We may revisit to refine the results by including the full dataset later once we have developed the general methodology.

```
#subset ages 35-65 (to avoid 0 number of claims)

data_subset_ACI<-data_summarised_ACI[(34<data_summarised_ACI$Age & data_summarised_ACI$Age<=65)
                                     ,c("Age","Year","Claim","Exposure")]
data_subset_DB<-data_summarised_DB[(34<data_summarised_DB$Age & data_summarised_DB$Age<=65)
                                   ,c("Age","Year","Claim","Exposure")]
data_subset_SCI<-data_summarised_SCI[(34<data_summarised_SCI$Age & data_summarised_SCI$Age<=65)
                                     ,c("Age","Year","Claim","Exposure")]
data_subset_Annuities<-data_summarised_Annuities[(34<data_summarised_Annuities$Age & data_summarised_Ann
                                                 ,c("Age","Year","Claim","Exposure")]

#round off the number of claims/deaths to be consistent with the model specification due to Poisson/Bin

data_subset_ACI$Qx<-round(data_subset_ACI$Claim)/data_subset_ACI$Exposure
data_subset_DB$Qx<-round(data_subset_DB$Claim)/data_subset_DB$Exposure
data_subset_SCI$Qx<-round(data_subset_SCI$Claim)/data_subset_SCI$Exposure
data_subset_Annuities$Qx<-round(data_subset_Annuities$Claim)/data_subset_Annuities$Exposure
```

For illustrative purposes, we could also aggregate the death data (over `ACI`, `DB`, `SCI`).

```
##aggregated data

data_subset_aggregate<-data_subset_ACI
data_subset_aggregate[,3:4]<-cbind((data_subset_ACI[,3]+data_subset_DB[,3]+data_subset_SCI[,3])
                                   ,(data_subset_ACI[,4]+data_subset_DB[,4]+data_subset_SCI[,4]))

data_subset_aggregate$Qx<-round(data_subset_aggregate$Claim)/data_subset_aggregate$Exposure
```

```
############
##plots by product (trends observed in the data)
###########

plot(unique(data_subset_ACI$Age),log(data_subset_ACI$Qx[data_subset_ACI$Year==2020]),xlab="Age"
     ,ylab="",type="l",col=2,main="Observed log death rates in 2020",ylim=c(-8.5,-4))
lines(unique(data_subset_DB$Age),log(data_subset_DB$Qx[data_subset_DB$Year==2020]),col=3)
lines(unique(data_subset_SCI$Age),log(data_subset_SCI$Qx[data_subset_SCI$Year==2020]),col=4)
lines(unique(data_subset_aggregate$Age),log(data_subset_aggregate$Qx[data_subset_aggregate$Year==2020])
lines(unique(data_subset_Annuities$Age),log(data_subset_Annuities$Qx[data_subset_Annuities$Year==2020])
#some issues with annuities data,death rates are unexpectedly high and contain many zeros
legend("topleft",c("aggregate","ACI","DB","SCI","Annuities"),lty=1,col=1:5)
```




```r
#I have chosen to ignore annuities from now
```

Fitting the Lee-Carter separately for each product

The most naive approach would be to fit the Lee-Carter (LC) model separately on each product, corresponding to the model below.

$$
\begin{aligned}
d_{x,t,p} &\sim \text{Pois}(E_{x,t,p} m_{x,t,p}) \\
\log(m_{x,t,p}) &= a_{x,p} + b_{x,p} k_{t,p}
\end{aligned}
$$

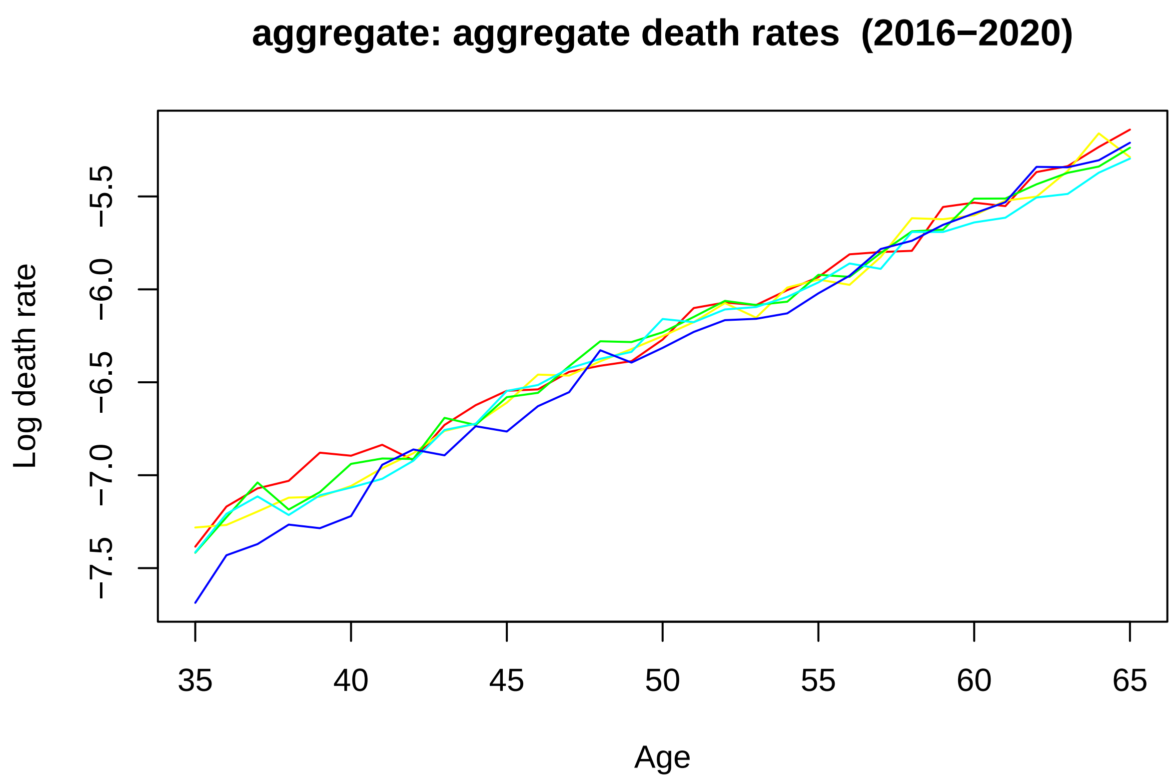
Note we only consider the Poisson specification for number of death but it is straightforward to switch them to modelling using the binomial model, i.e.

$$
\begin{aligned}
d_{x,t,p} &\sim \text{Binomial}(E_{x,t,p}, q_{x,t,p}) \\
\text{logit}(q_{x,t,p}) &= a_{x,p} + b_{x,p} k_{t,p}
\end{aligned}
$$

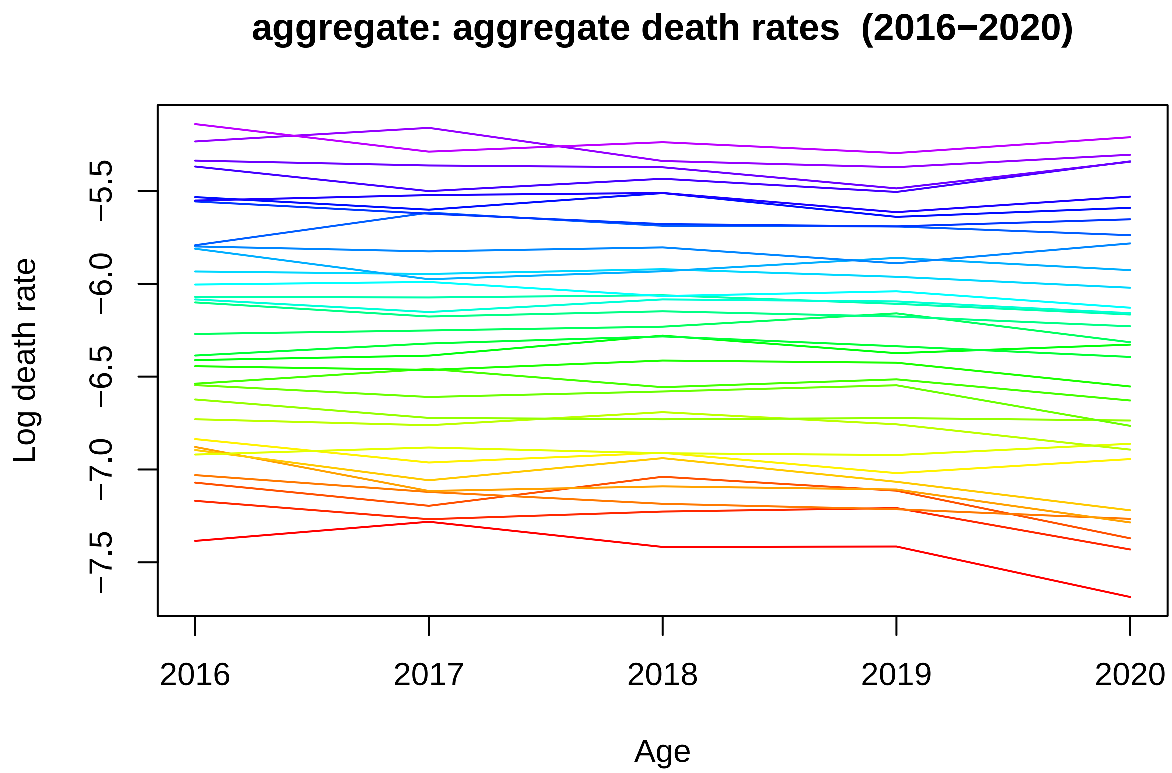
There are two readily available `R` packages to fit the LC model, `demography` (documentation here) which was developed earlier on and `StMoMo` (documentation here) which is more recent and contain more functionalities and models in the package. We will show both to ensure consistencies. Note of course they are not the same, since `demography` package execute the estimation of the original LC approach, which comes with normal model error, SVD etc. But the `StMoMo` corresponds to the Poisson LC model (Brouhns et al., 2002).

For aggregate data

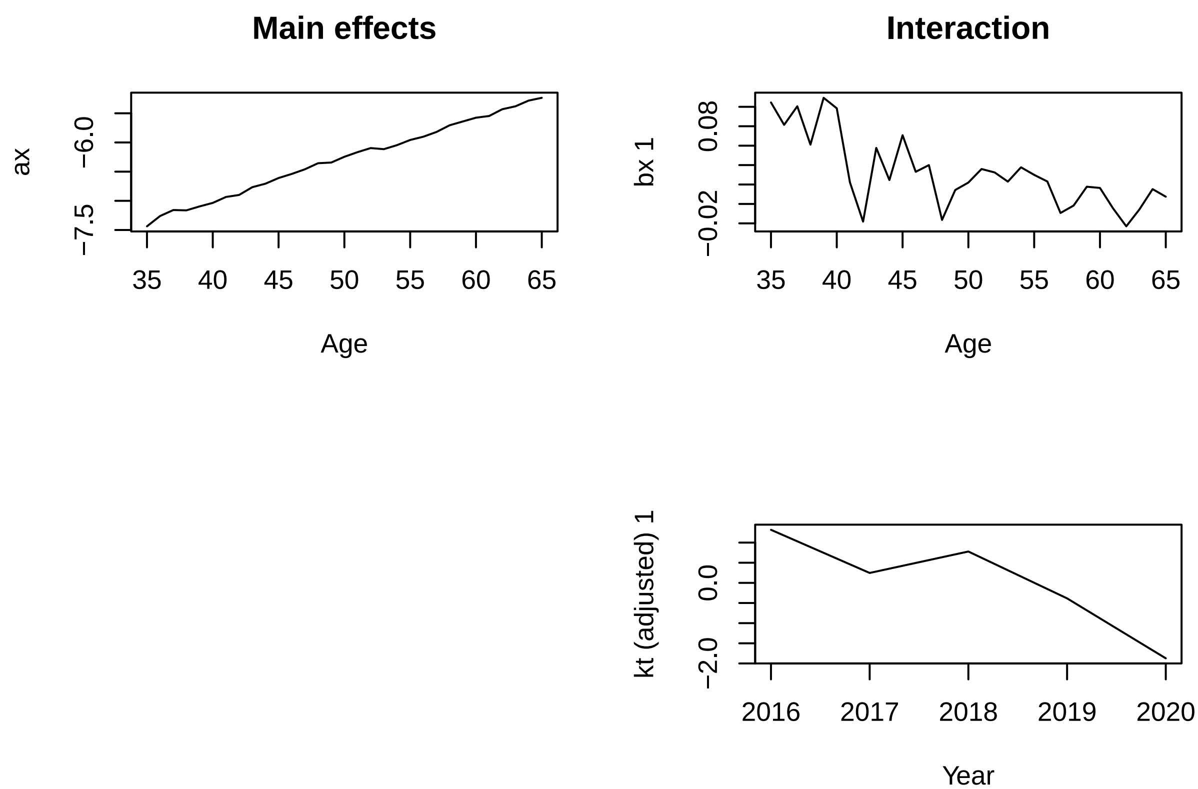
```
#prepare the death rates and exposures in matrix format
rates_subset_aggregate_mat<-matrix(data_subset_aggregate$Qx,nrow=length(unique(data_subset_aggregate$Age
rownames(rates_subset_aggregate_mat)<-unique(data_subset_aggregate$Age)
colnames(rates_subset_aggregate_mat)<-unique(data_subset_aggregate$Year)

expo_subset_aggregate_mat<-matrix(data_subset_aggregate$Exposure,nrow=length(unique(data_subset_aggrega
rownames(expo_subset_aggregate_mat)<-unique(data_subset_aggregate$Age)
colnames(expo_subset_aggregate_mat)<-unique(data_subset_aggregate$Year)

#LC by demography

#install.packages("demography")
library(demography)

rates_subset_aggregate_demog<-demogdata(rates_subset_aggregate_mat,pop=expo_subset_aggregate_mat,ages=un
plot(rates_subset_aggregate_demog,plot.type = "functions") #rainbow scale: red for earliest years, then
```




```
plot(rates_subset_aggregate_demog,plot.type = "time")
```



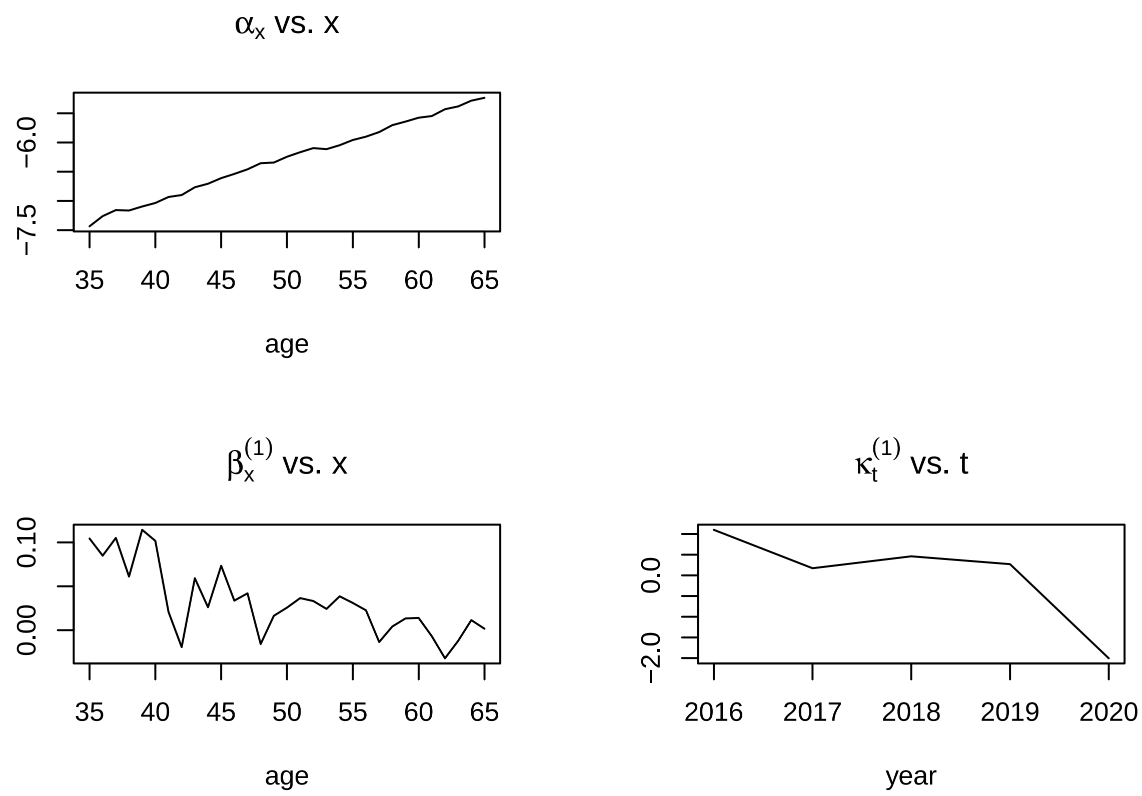

```
lca_subset_aggregate_demog<-lca(rates_subset_aggregate_demog)
plot(lca_subset_aggregate_demog)
```




```r
#LC by StMoMo

#install.package("StMoMo")
library(StMoMo)

rates_subset_aggregate_stmomo<-StMoMoData(rates_subset_aggregate_demog)

LC_log<-lc()

lca_log_subset_aggregate_stmomo<-fit(LC_log,data=rates_subset_aggregate_stmomo)
```

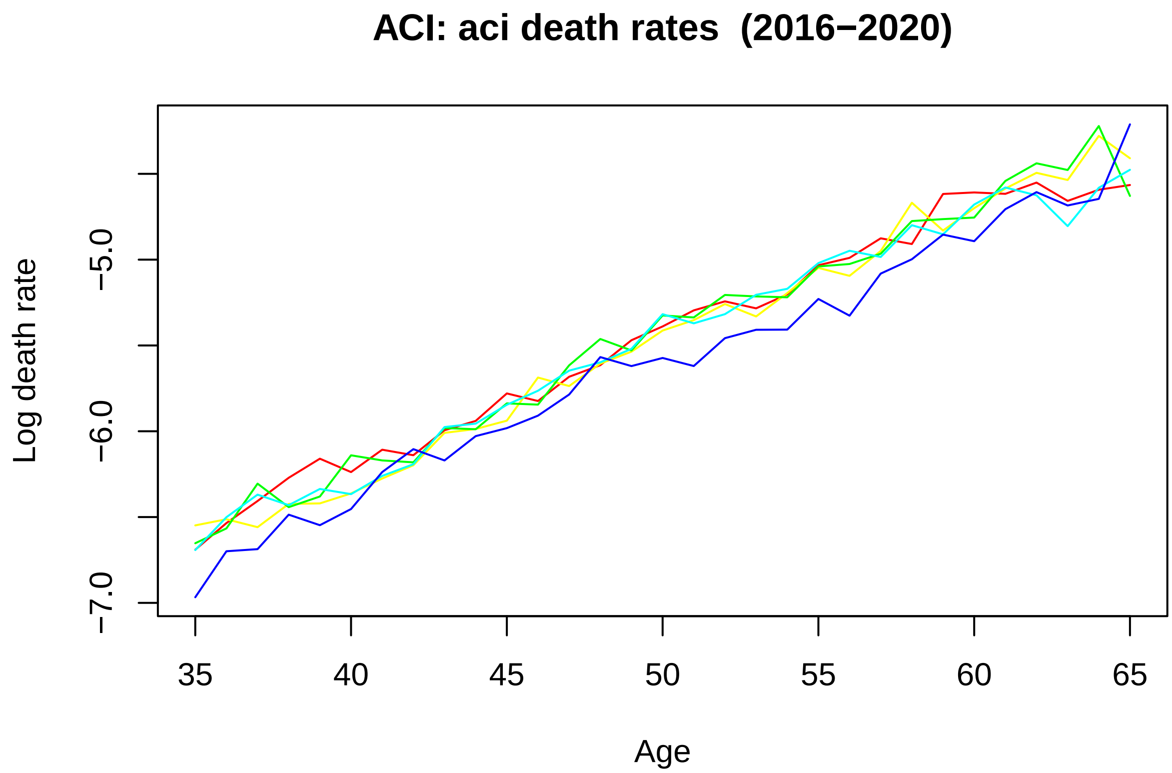
```
## StMoMo: Start fitting with gnm
## Initialising
## Running start-up iterations..
## Running main iterations.........................
## Done
## StMoMo: Finish fitting with gnm
```

```r
plot(lca_log_subset_aggregate_stmomo)
```

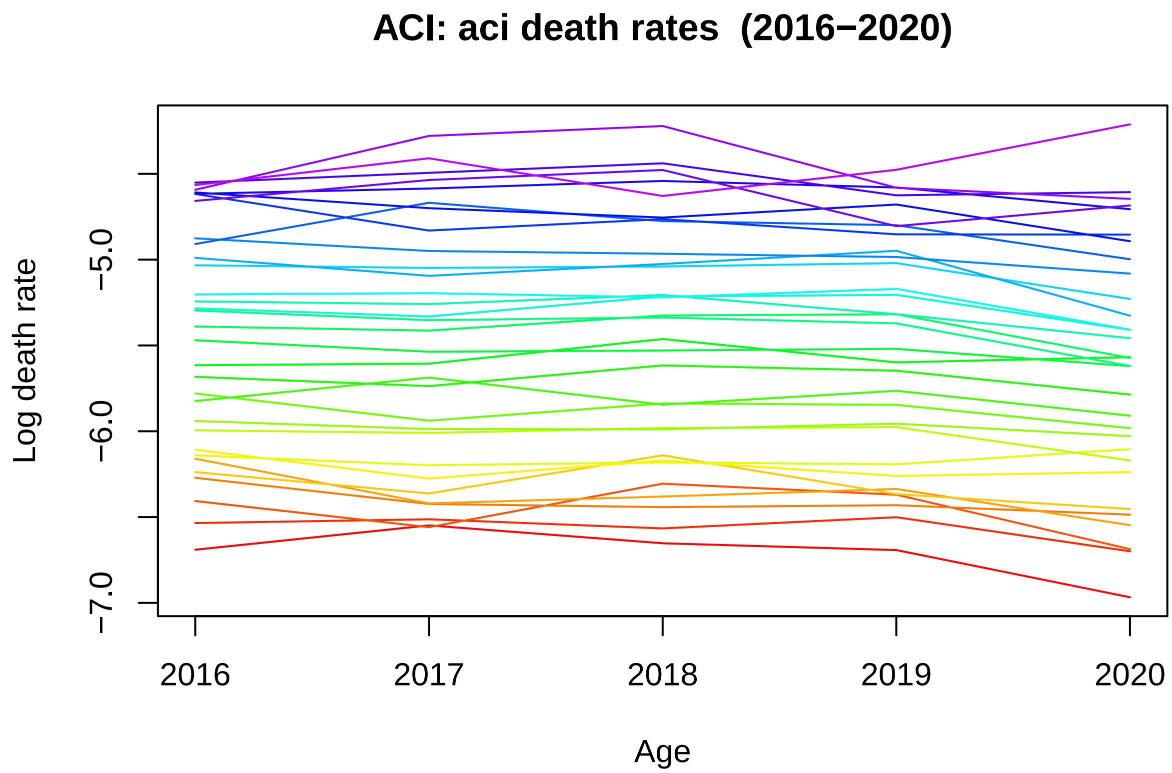


For other products

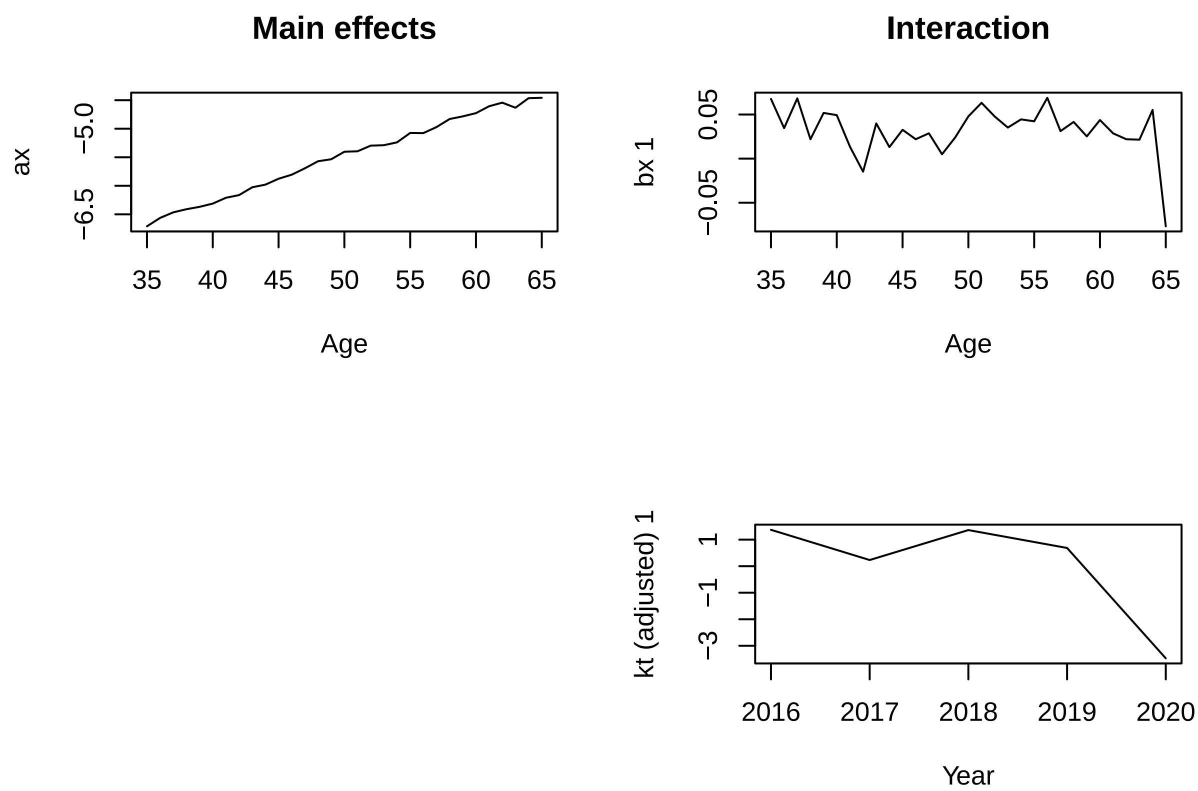
```
###########
## only ACI
###########

rates_subset_ACI_mat<-matrix(data_subset_ACI$Qx,nrow=length(unique(data_subset_ACI$Age)),ncol=length(uni
rownames(rates_subset_ACI_mat)<-unique(data_subset_ACI$Age)
colnames(rates_subset_ACI_mat)<-unique(data_subset_ACI$Year)

expo_subset_ACI_mat<-matrix(data_subset_ACI$Exposure,nrow=length(unique(data_subset_ACI$Age)),ncol=leng
rownames(expo_subset_ACI_mat)<-unique(data_subset_ACI$Age)
colnames(expo_subset_ACI_mat)<-unique(data_subset_ACI$Year)

#LC by demography
rates_subset_ACI_demog<-demogdata(rates_subset_ACI_mat,pop=expo_subset_ACI_mat,ages=unique(data_subset_
plot(rates_subset_ACI_demog)
```




```
plot(rates_subset_ACI_demog,plot.type = "time")
```



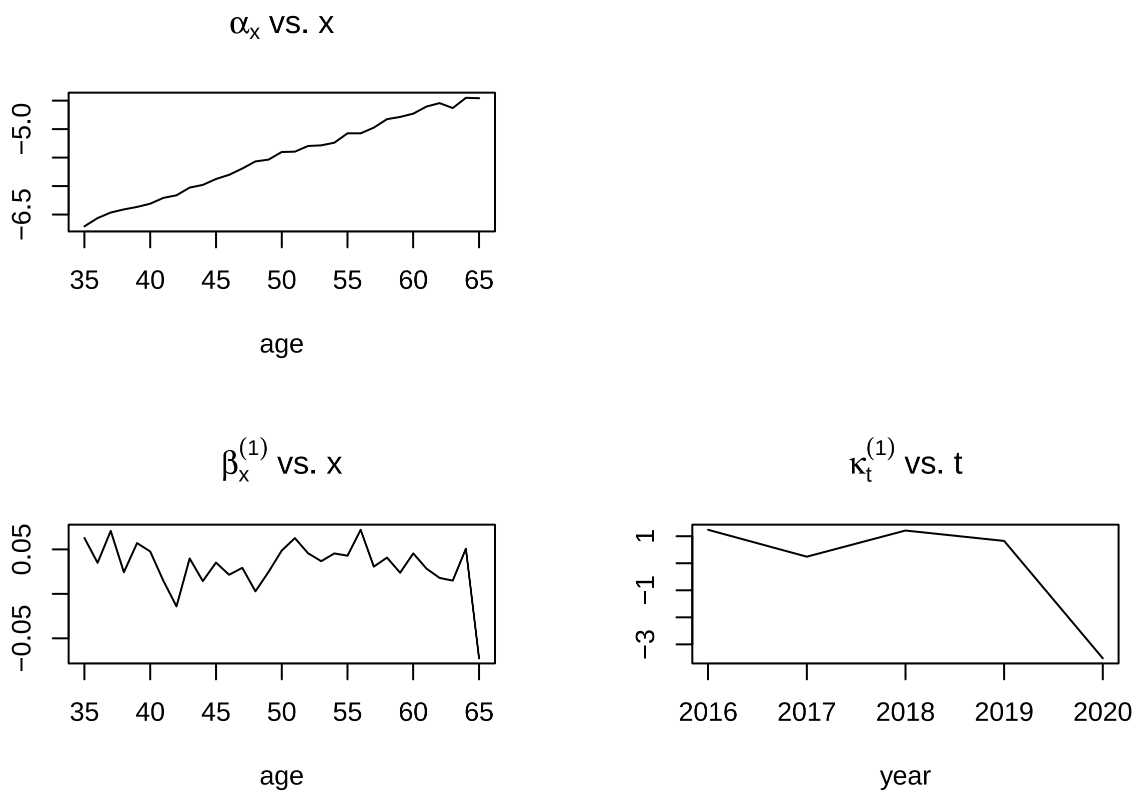

```
lca_subset_ACI_demog<-lca(rates_subset_ACI_demog)
plot(lca_subset_ACI_demog)
```



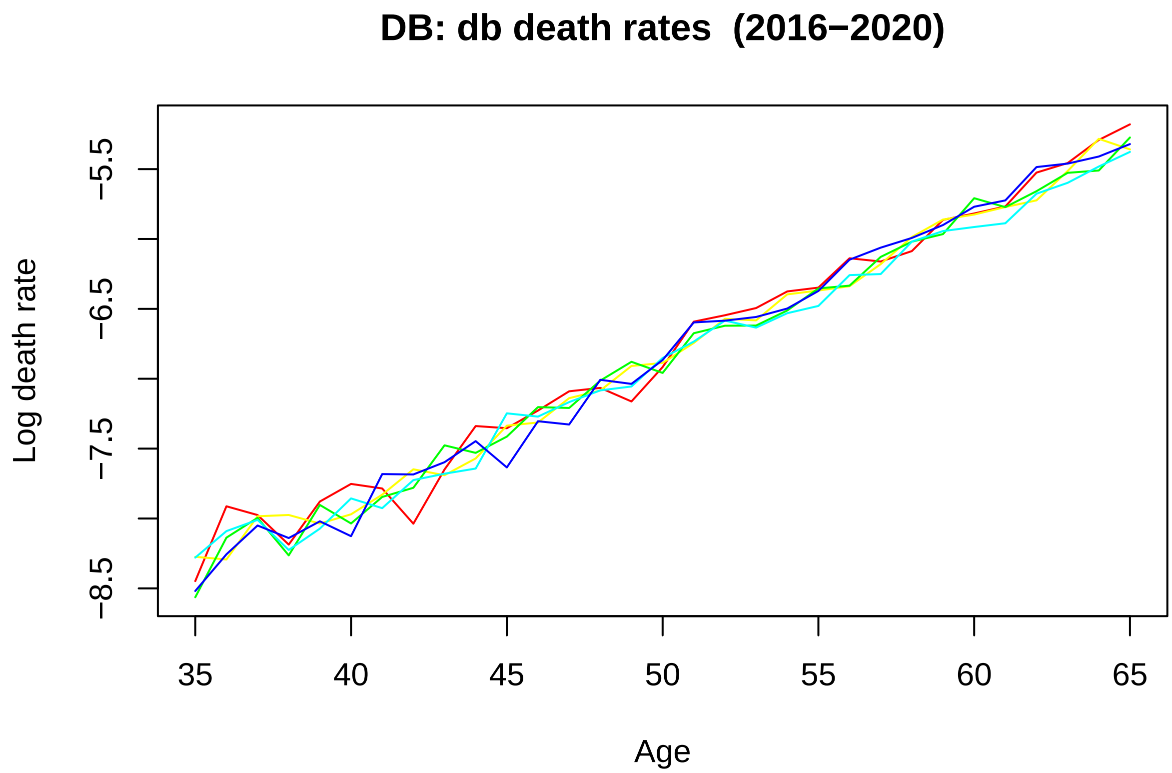

```r
#LC by StMoMo
rates_subset_ACI_stmomo<-StMoMoData(rates_subset_ACI_demog)

lca_log_subset_ACI_stmomo<-fit(LC_log,data=rates_subset_ACI_stmomo)
```

```
## StMoMo: Start fitting with gnm
## Initialising
## Running start-up iterations..
## Running main iterations................
## Done
## StMoMo: Finish fitting with gnm
```

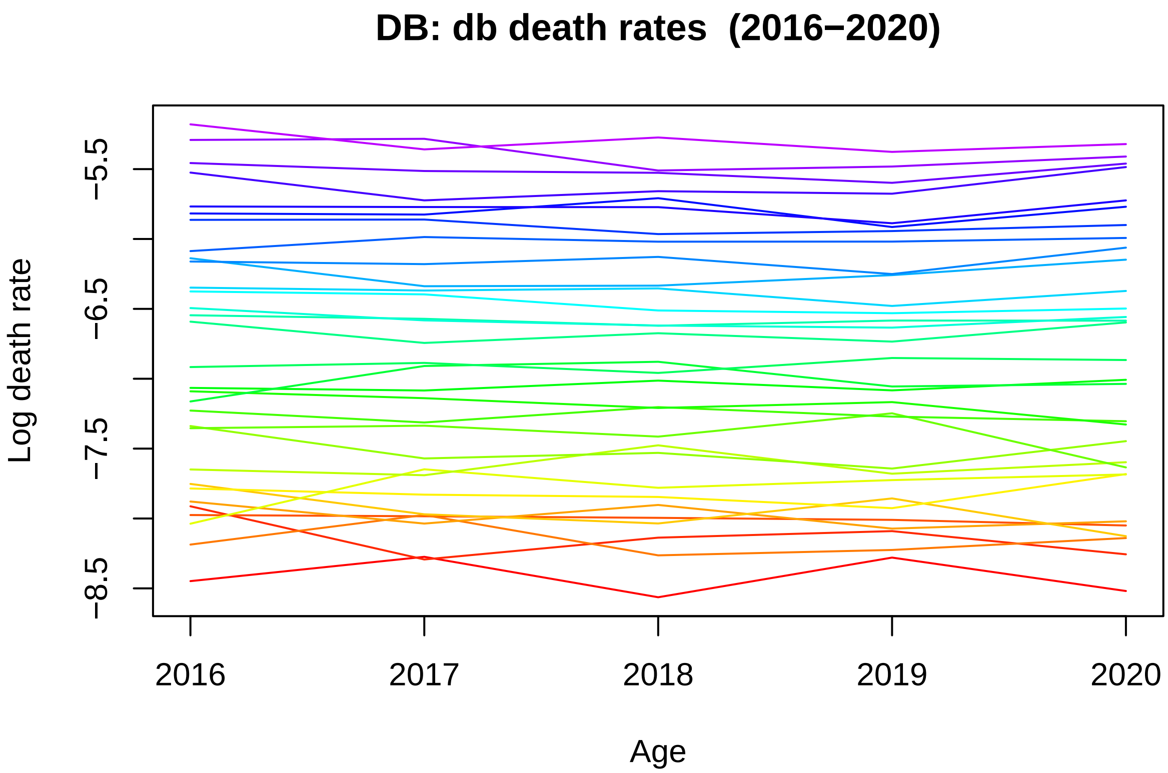
```r
plot(lca_log_subset_ACI_stmomo)
```



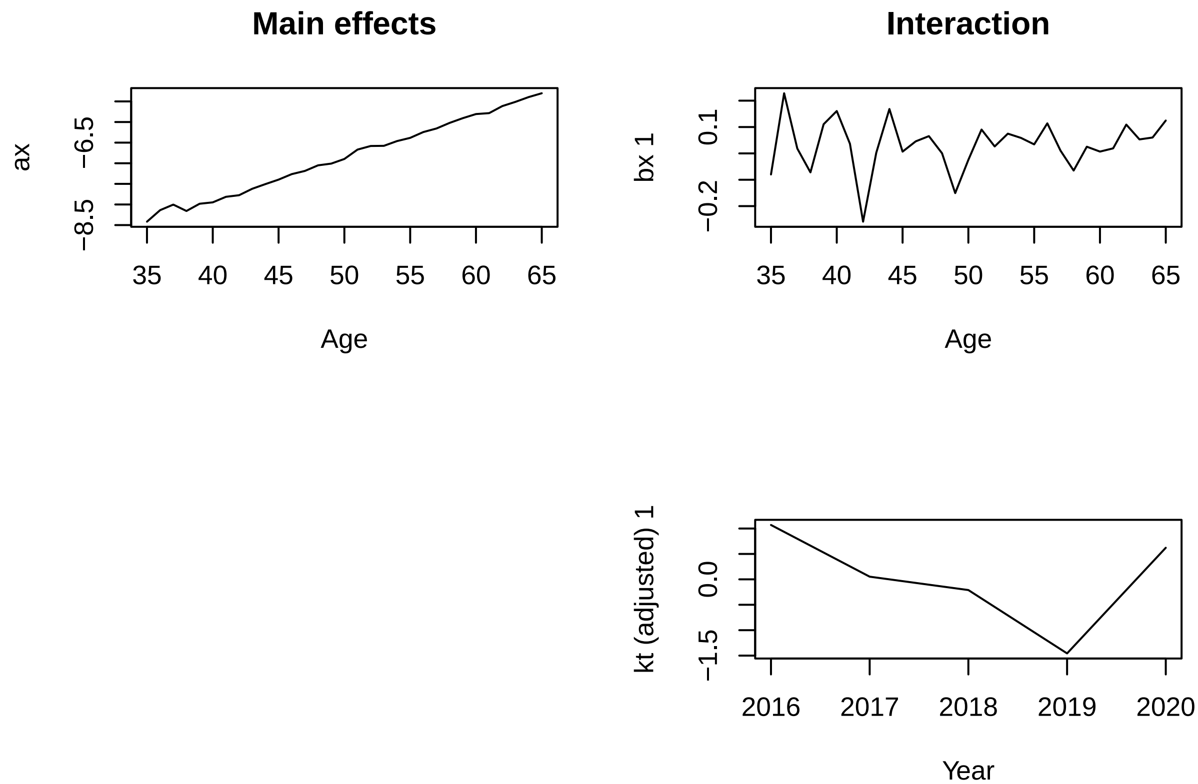

```
##########
## only DB
##########

rates_subset_DB_mat<-matrix(data_subset_DB$Qx,nrow=length(unique(data_subset_DB$Age)),ncol=length(unique
rownames(rates_subset_DB_mat)<-unique(data_subset_DB$Age)
colnames(rates_subset_DB_mat)<-unique(data_subset_DB$Year)

expo_subset_DB_mat<-matrix(data_subset_DB$Exposure,nrow=length(unique(data_subset_DB$Age)),ncol=length(u
rownames(expo_subset_DB_mat)<-unique(data_subset_DB$Age)
colnames(expo_subset_DB_mat)<-unique(data_subset_DB$Year)

#LC by demography
rates_subset_DB_demog<-demogdata(rates_subset_DB_mat,pop=expo_subset_DB_mat,ages=unique(data_subset_DB$A
plot(rates_subset_DB_demog)
```




```
plot(rates_subset_DB_demog,plot.type = "time")
```



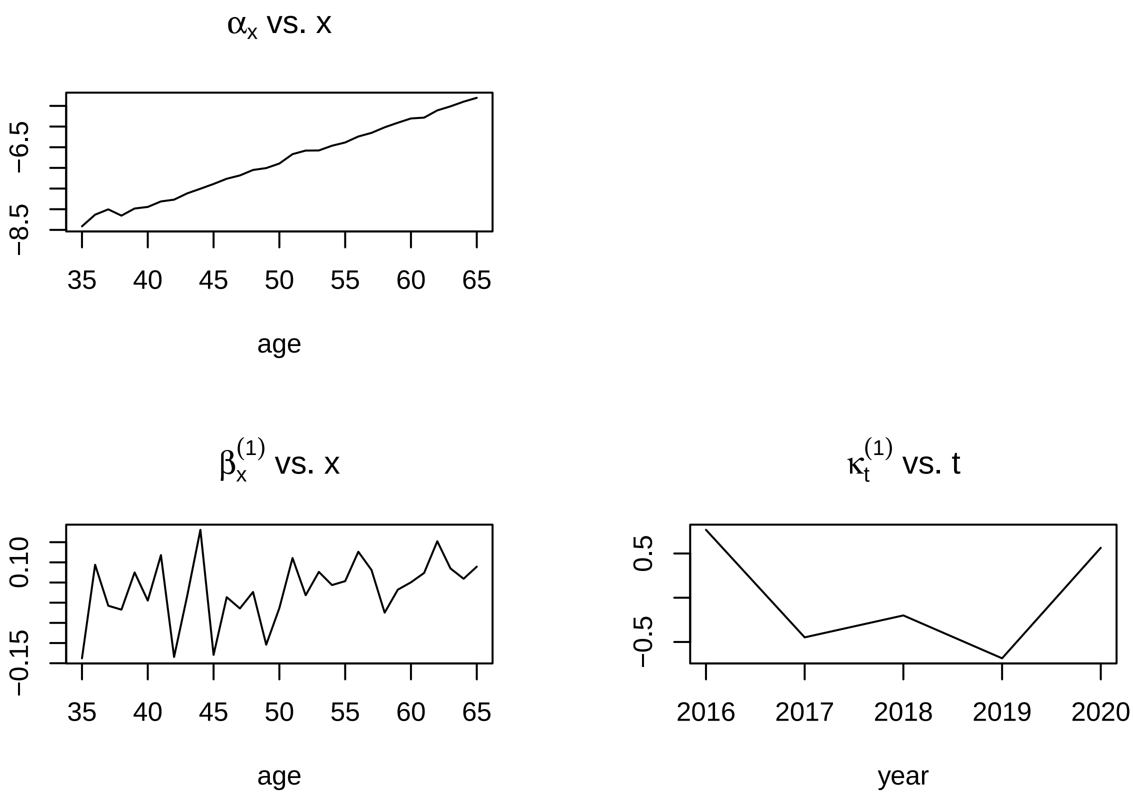

```
#LC by StMoMo
lca_subset_DB_demog<-lca(rates_subset_DB_demog)
plot(lca_subset_DB_demog)
```



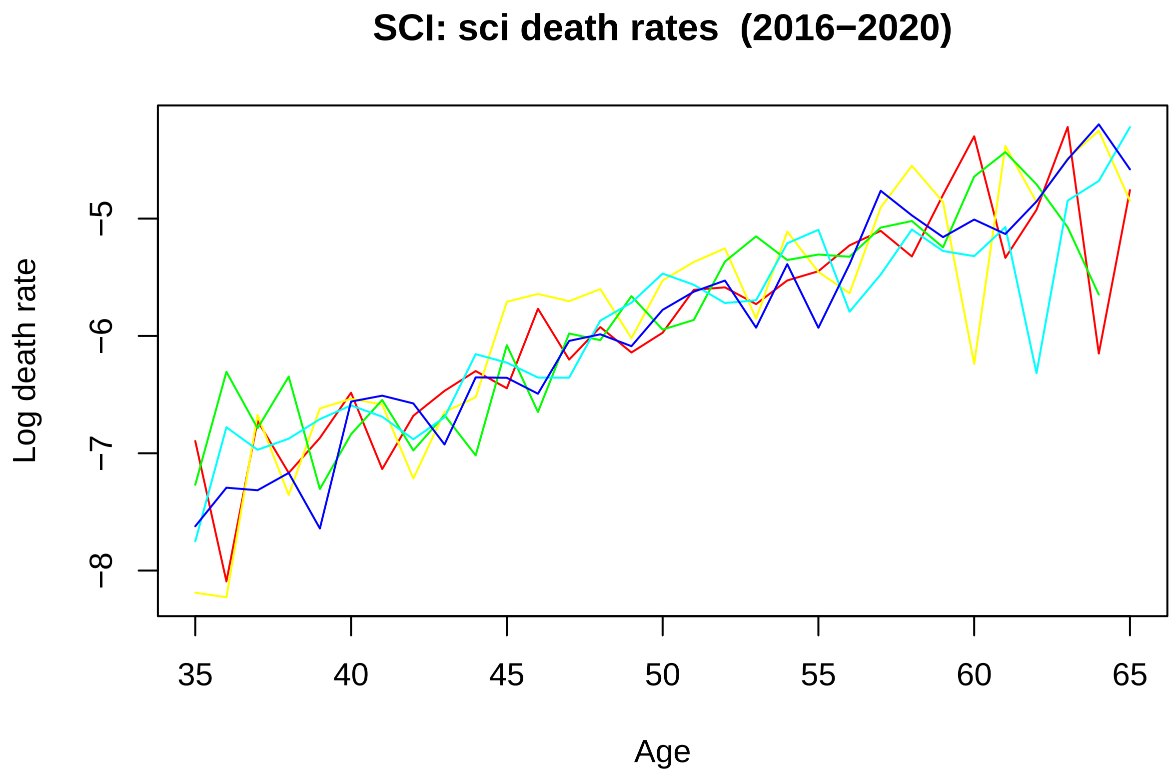

```
rates_subset_DB_stmomo<-StMoMoData(rates_subset_DB_demog)

lca_log_subset_DB_stmomo<-fit(LC_log,data=rates_subset_DB_stmomo)
```

```
## StMoMo: Start fitting with gnm
## Initialising
## Running start-up iterations..
## Running main iterations.........................................................
## .........
## Done
## StMoMo: Finish fitting with gnm
```

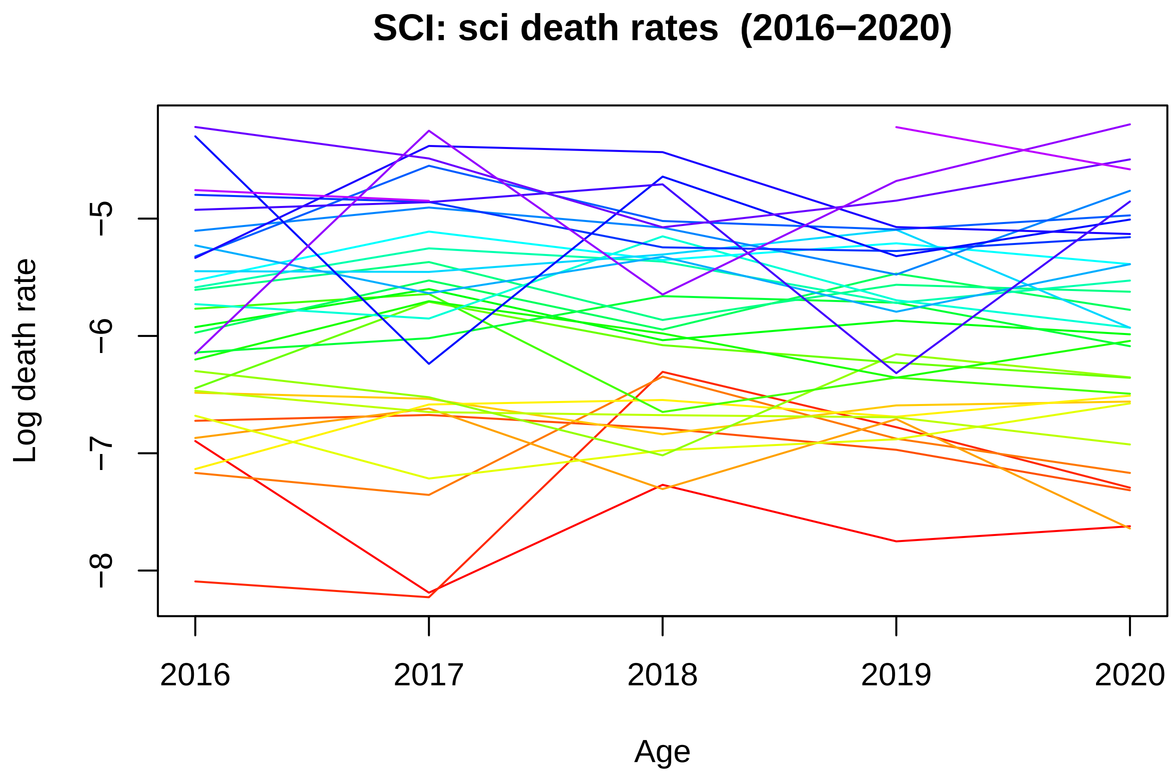
```
plot(lca_log_subset_DB_stmomo)
```




```
############
## only SCI
############

rates_subset_SCI_mat<-matrix(data_subset_SCI$Qx,nrow=length(unique(data_subset_SCI$Age)),ncol=length(uni
rownames(rates_subset_SCI_mat)<-unique(data_subset_SCI$Age)
colnames(rates_subset_SCI_mat)<-unique(data_subset_SCI$Year)

expo_subset_SCI_mat<-matrix(data_subset_SCI$Exposure,nrow=length(unique(data_subset_SCI$Age)),ncol=lengt
rownames(expo_subset_SCI_mat)<-unique(data_subset_SCI$Age)
colnames(expo_subset_SCI_mat)<-unique(data_subset_SCI$Year)

#LC by demography
rates_subset_SCI_demog<-demogdata(rates_subset_SCI_mat,pop=expo_subset_SCI_mat,ages=unique(data_subset_S
plot(rates_subset_SCI_demog)
```



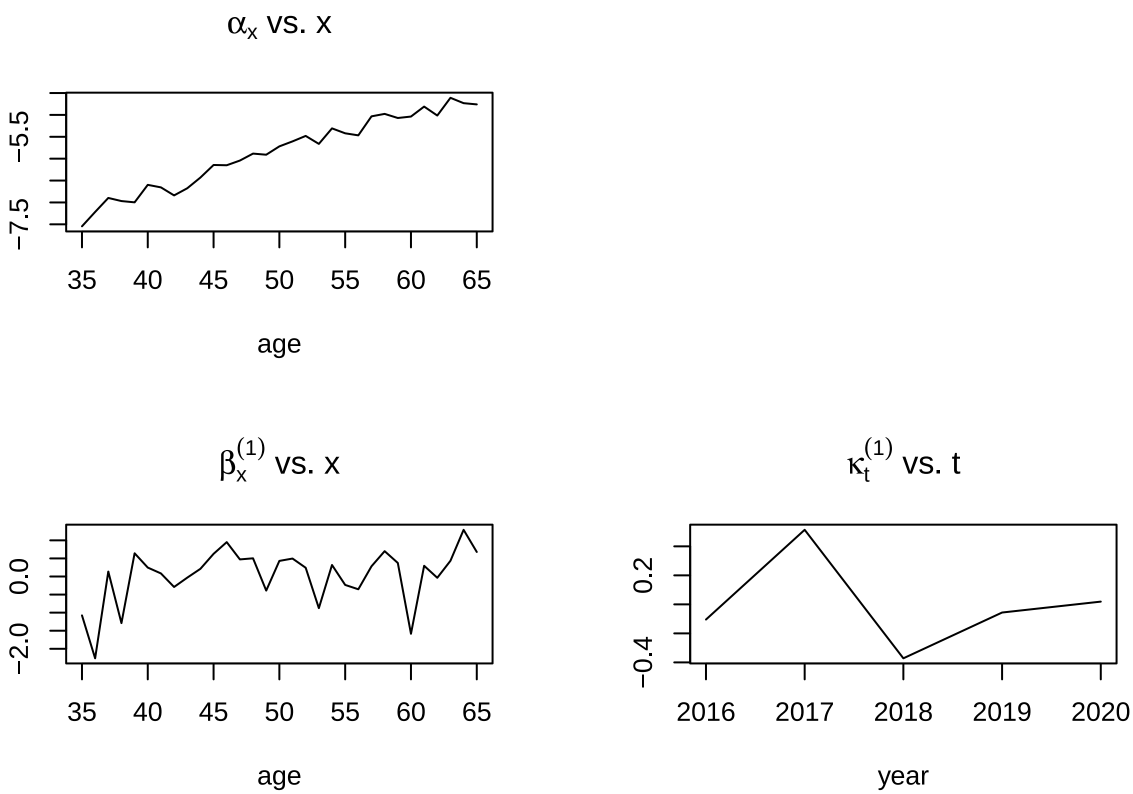

```
plot(rates_subset_SCI_demog,plot.type = "time")
```



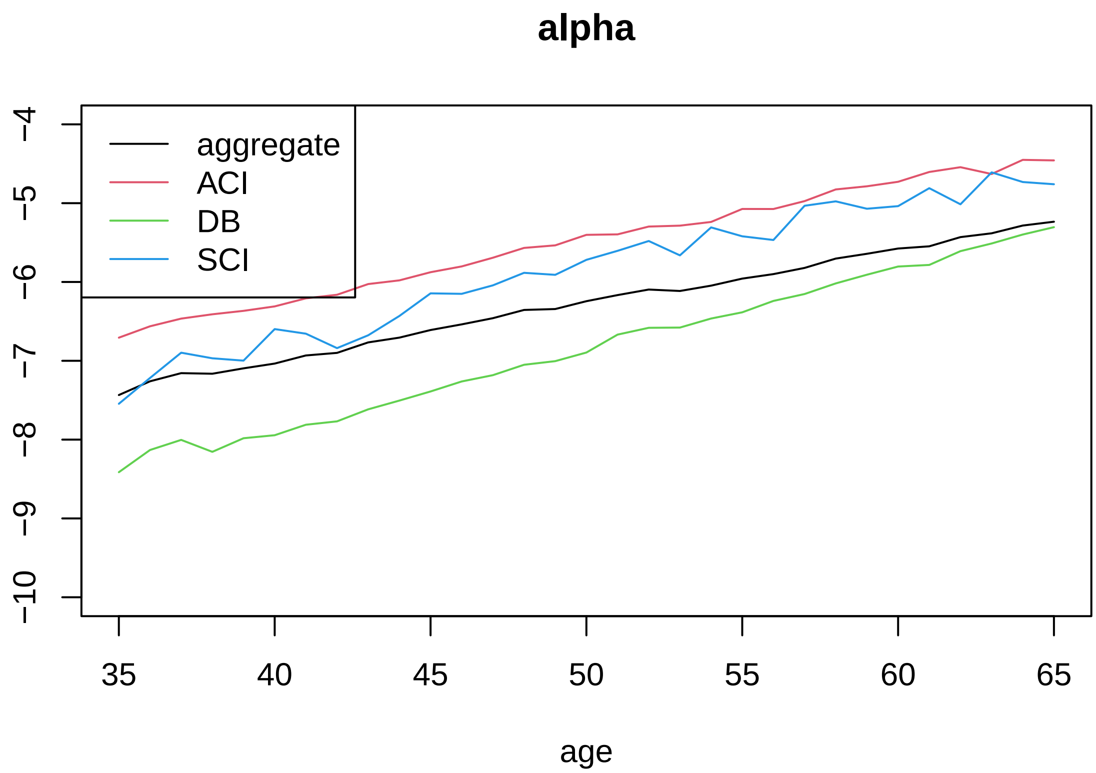

```
#the following failed to run
#lca_subset_SCI_demog<-lca(rates_subset_SCI_demog)
#plot(lca_subset_SCI_demog)

#LC by StMoMo
rates_subset_SCI_stmomo<-StMoMoData(rates_subset_SCI_demog)

lca_log_subset_SCI_stmomo<-fit(LC_log,data=rates_subset_SCI_stmomo)
```

```
## StMoMo: Start fitting with gnm
## Initialising
## Running start-up iterations..
## Running main iterations.....................................................
## .............
## Done
## StMoMo: Finish fitting with gnm
```

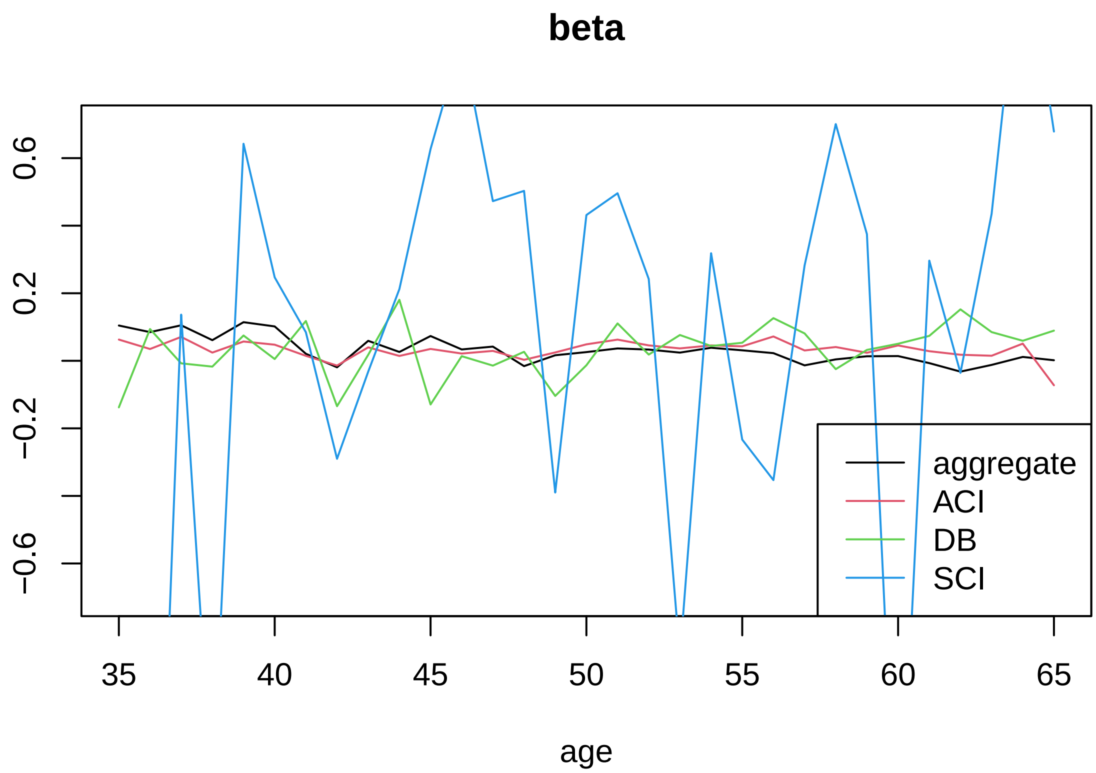
```
plot(lca_log_subset_SCI_stmomo)
```



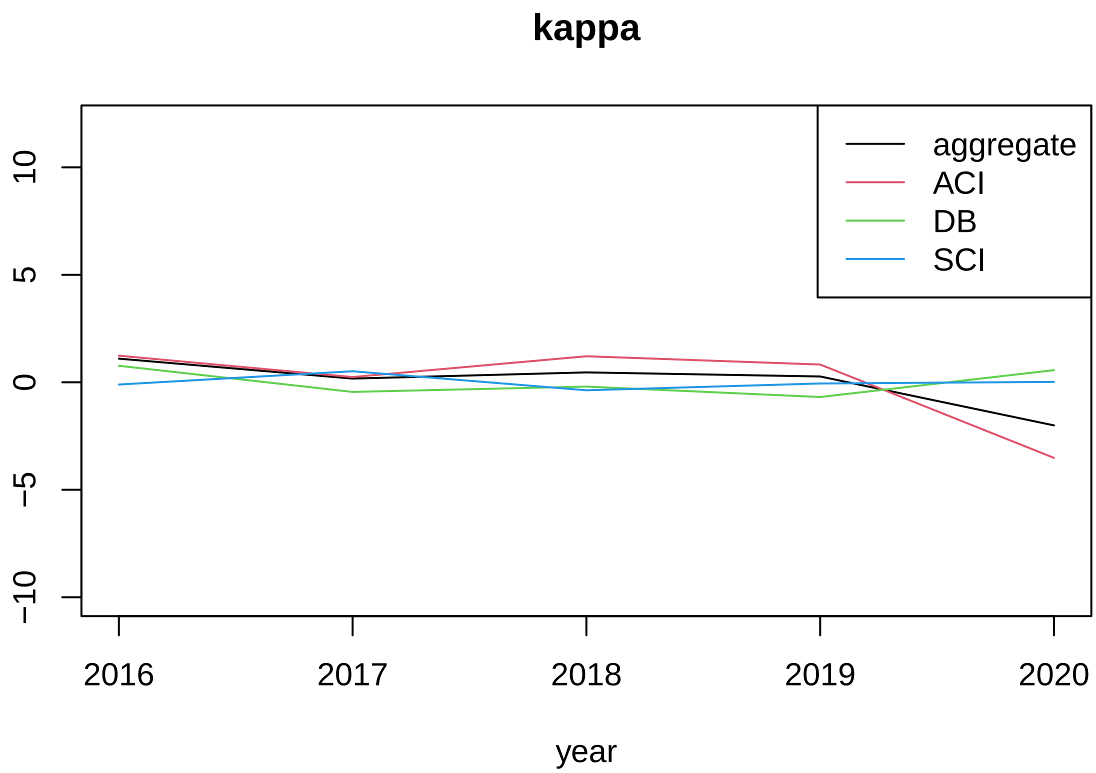

```
############
##plots of parameters in a single plot (individually fitting LC model)
###########

par(mfrow=c(1,1))
plot(unique(data_subset_aggregate$Age),lca_log_subset_aggregate_stmomo$ax,type="l",xlab="age",main="alph
lines(unique(data_subset_aggregate$Age),lca_log_subset_ACI_stmomo$ax,col=2)
lines(unique(data_subset_aggregate$Age),lca_log_subset_DB_stmomo$ax,col=3)
lines(unique(data_subset_aggregate$Age),lca_log_subset_SCI_stmomo$ax,col=4)
legend("topleft",c("aggregate","ACI","DB","SCI"),lty=1,col=1:4)
```




```
plot(unique(data_subset_aggregate$Age),lca_log_subset_aggregate_stmomo$bx,type="l",xlab="age",main="bet
lines(unique(data_subset_aggregate$Age),lca_log_subset_ACI_stmomo$bx,col=2)
lines(unique(data_subset_aggregate$Age),lca_log_subset_DB_stmomo$bx,col=3)
lines(unique(data_subset_aggregate$Age),lca_log_subset_SCI_stmomo$bx,col=4)
legend("bottomright",c("aggregate","ACI","DB","SCI"),lty=1,col=1:4)
```




```
plot(unique(data_subset_aggregate$Year),lca_log_subset_aggregate_stmomo$kt,ylim=c(-10,12),type="l",xlab=
lines(unique(data_subset_aggregate$Year),lca_log_subset_ACI_stmomo$kt,col=2)
lines(unique(data_subset_aggregate$Year),lca_log_subset_DB_stmomo$kt,col=3)
lines(unique(data_subset_aggregate$Year),lca_log_subset_SCI_stmomo$kt,col=4)
legend("topright",c("aggregate","ACI","DB","SCI"),lty=1,col=1:4)
```



For a_x, the patterns look very similar across the products, with only slight variations. For b_x, only SCI demonstrated substantially different pattern. So this may indicate that the model with product-specific betas ($b_{x,p}$), and shared a_x, k_t to be superior (Model 2A2)? For k_t, all looks the same.

Various extensions of the LC model

After fitting the LC model on each of the products, we inferred that there is evidence that the fitted parameters demonstrate common trends (e.g. a_x and k_t), but not all of them share the same features (e.g. b_x). So here, we consider multiple ways to extending the original LC model specification that may be suitable. All the functions to compile the frequentist fitting of these models are included in the file `idea-lr-frequentist.R` and loaded as below. They will not be printed here so click into the file to check out the functions if interested.

```
source("idea-lr-frequentist.R")
```

Before applying the functions, we will prepare the following quantities.

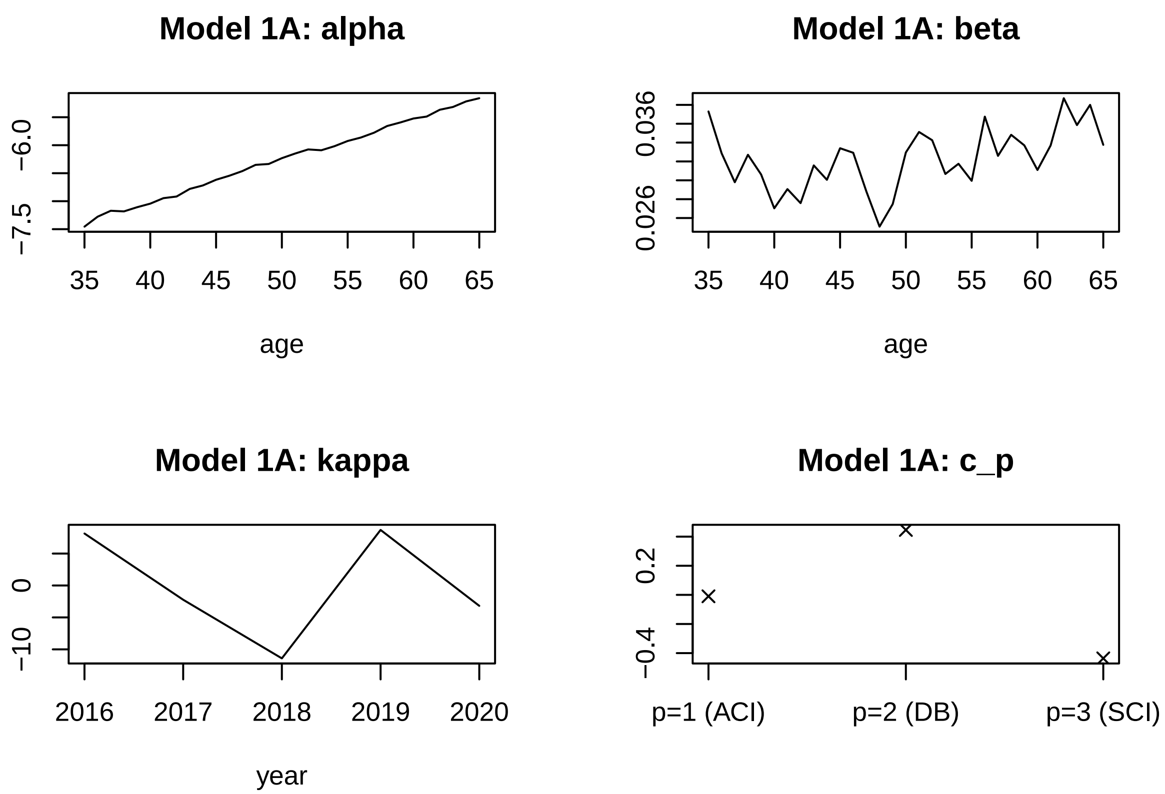
```
#number of deaths matrix rounded to nearest integers
deaths_subset_aggregate_mat<-round(rates_subset_aggregate_mat*expo_subset_aggregate_mat)
deaths_subset_ACI_mat<-round(rates_subset_ACI_mat*expo_subset_ACI_mat)
deaths_subset_DB_mat<-round(rates_subset_DB_mat*expo_subset_DB_mat)
deaths_subset_SCI_mat<-round(rates_subset_SCI_mat*expo_subset_SCI_mat)

A<-nrow(deaths_subset_ACI_mat);T<-ncol(deaths_subset_ACI_mat)
```

Model 1A

$$\log(m_{x,t,p}) = a_x + c_p + b_x k_t$$

Constraints: $\sum_{x=1}^{A} b_x = 1$, $\sum_{t=1}^{T} k_t{=}0$, $\sum_{p=1}^{3} c_p = 0$.

```
deaths_combine=round(rbind(deaths_subset_ACI_mat,deaths_subset_DB_mat,deaths_subset_SCI_mat))
expo_combine=rbind(expo_subset_ACI_mat,expo_subset_DB_mat,expo_subset_SCI_mat)

lca_subset_iterative_poisson_glm_M1A<-iteration.methodB.LC.poisson_M1A(deaths_combine=deaths_combine,
                                                                    exposure_combine=expo_combine,m=

par(mfrow=c(2,2))
plot(unique(data_subset_aggregate$Age),lca_subset_iterative_poisson_glm_M1A$alpha,type="l"
     ,xlab="age",main="Model 1A: alpha",ylab="")
plot(unique(data_subset_aggregate$Age),lca_subset_iterative_poisson_glm_M1A$beta,type="l"
     ,xlab="age",main="Model 1A: beta",ylab="")
plot(unique(data_subset_aggregate$Year),lca_subset_iterative_poisson_glm_M1A$kappa,type="l"
     ,xlab="year",main="Model 1A: kappa",ylab="")
plot(1:3,c(lca_subset_iterative_poisson_glm_M1A$intercept,-sum(lca_subset_iterative_poisson_glm_M1A$inte
     pch=4,xlab="",ylab="",xaxt="n",main="Model 1A: c_p")
axis(1,at=1:3,label=c("p=1 (ACI)","p=2 (DB)","p=3 (SCI)"))
```




```
lca_subset_iterative_poisson_glm_M1A$deviance[length(lca_subset_iterative_poisson_glm_M1A$deviance)]
```

```
## [1] 26705
```

```
lca_subset_iterative_poisson_glm_M1A$model_bic[length(lca_subset_iterative_poisson_glm_M1A$model_bic)]
```

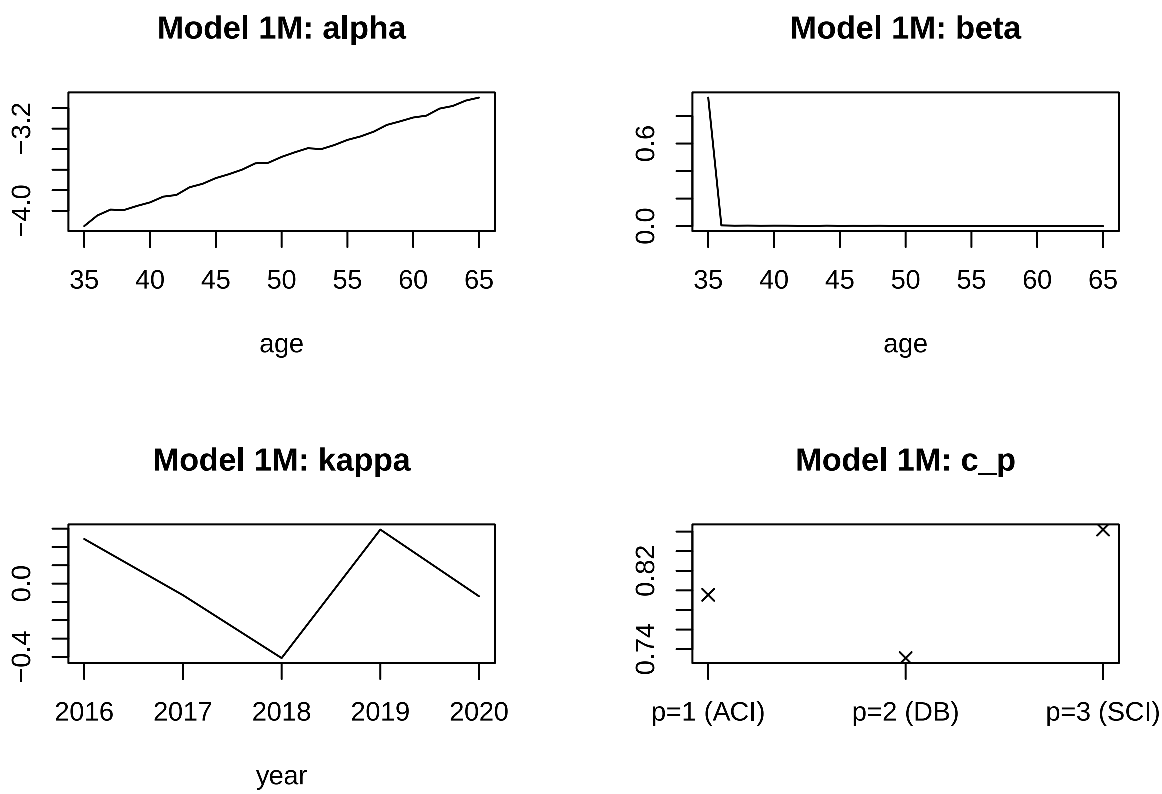
```
## [1] 29978.92
```

Model 1M

$$\log(m_{x,t,p}) = a_x(1+c_p) + b_x k_t$$
$$Constraints : \sum_{x=1}^{A} b_x = 1, \sum_{t=1}^{T} k_t = 0.$$

We do not need the constraint $\sum_{p=1}^{3} c_p = 0$ because of the way the model is parameterised, so c_p is identifiable relative to additive constants. I am not sure about the results though, they don't look correct. The model might not have converged due to computational issues (e.g. due to insufficient information from the data to estimate c_p).

```
lca_subset_iterative_poisson_glm_M1M<-iteration.methodB.LC.poisson_M1M(deaths_combine=deaths_combine,
                                                                       exposure_combine=expo_combine,m=

par(mfrow=c(2,2))
plot(unique(data_subset_aggregate$Age),lca_subset_iterative_poisson_glm_M1M$alpha,type="l",
     xlab="age",main="Model 1M: alpha",ylab="")
plot(unique(data_subset_aggregate$Age),lca_subset_iterative_poisson_glm_M1M$beta,type="l",
     xlab="age",main="Model 1M: beta",ylab="")
plot(unique(data_subset_aggregate$Year),lca_subset_iterative_poisson_glm_M1M$kappa,type="l",
     xlab="year",main="Model 1M: kappa",ylab="")
plot(1:3,lca_subset_iterative_poisson_glm_M1M$c.new,pch=4,xlab="",ylab="",xaxt="n",main="Model 1M: c_p")
axis(1,at=1:3,label=c("p=1 (ACI)","p=2 (DB)","p=3 (SCI)"))
```




```
lca_subset_iterative_poisson_glm_M1M$deviance[length(lca_subset_iterative_poisson_glm_M1M$deviance)]
```

```
## [1] 28860.33
```

```
lca_subset_iterative_poisson_glm_M1M$model_bic[length(lca_subset_iterative_poisson_glm_M1M$model_bic)]
```

```
## [1] 31974.56
```

Model 1M2

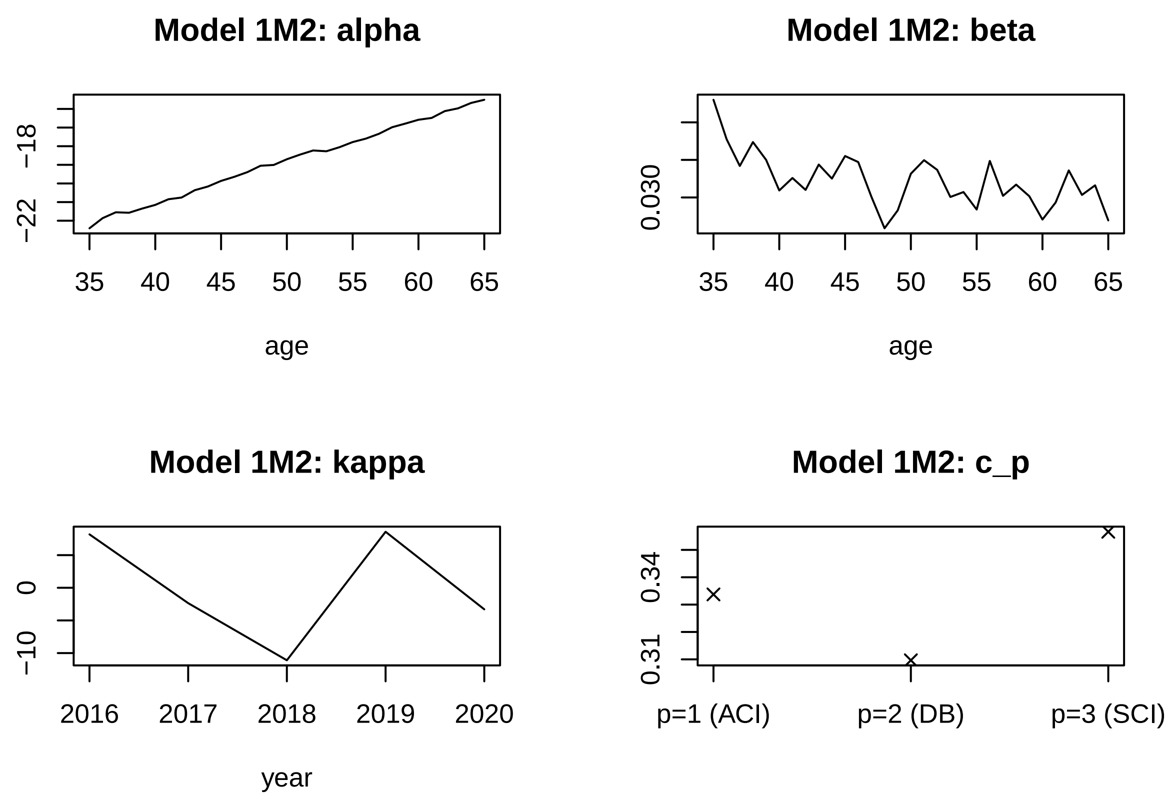
$$\log(m_{x,t,p}) = a_x c_p + b_x k_t$$

Model 1M2 is a similar variant as Model 1M, but should be easier to estimate comparatively.

$$Constraints: \sum_{x=1}^{A} b_x = 1, \sum_{t=1}^{T} k_t = 0, \sum_{p=1}^{3} c_p = 1.$$

```
lca_subset_iterative_poisson_glm_M1M2<-iteration.methodB.LC.poisson_M1M2(deaths_combine=deaths_combine,
                                                                          exposure_combine=expo_combine,

par(mfrow=c(2,2))
plot(unique(data_subset_aggregate$Age),lca_subset_iterative_poisson_glm_M1M2$alpha,type="l"
     ,xlab="age",main="Model 1M2: alpha",ylab="")
```

```r
plot(unique(data_subset_aggregate$Age),lca_subset_iterative_poisson_glm_M1M2$beta,type="l"
     ,xlab="age",main="Model 1M2: beta",ylab="")
plot(unique(data_subset_aggregate$Year),lca_subset_iterative_poisson_glm_M1M2$kappa,type="l"
     ,xlab="year",main="Model 1M2: kappa",ylab="")
plot(1:3,lca_subset_iterative_poisson_glm_M1M2$c.new,pch=4,xlab="",ylab="",xaxt="n",main="Model 1M2: c_p
axis(1,at=1:3,label=c("p=1 (ACI)","p=2 (DB)","p=3 (SCI)"))
```




```r
lca_subset_iterative_poisson_glm_M1M2$deviance[length(lca_subset_iterative_poisson_glm_M1M2$deviance)]
```

```
## [1] 26664.88
```

```r
lca_subset_iterative_poisson_glm_M1M2$model_bic[length(lca_subset_iterative_poisson_glm_M1M2$model_bic)]
```

```
## [1] 29748.39
```

As observed above, the estimated parameters behave in a more stable manner.

Model 1U

$$\log(m_{x,t,p}) = a_{x,p} + b_x k_t$$

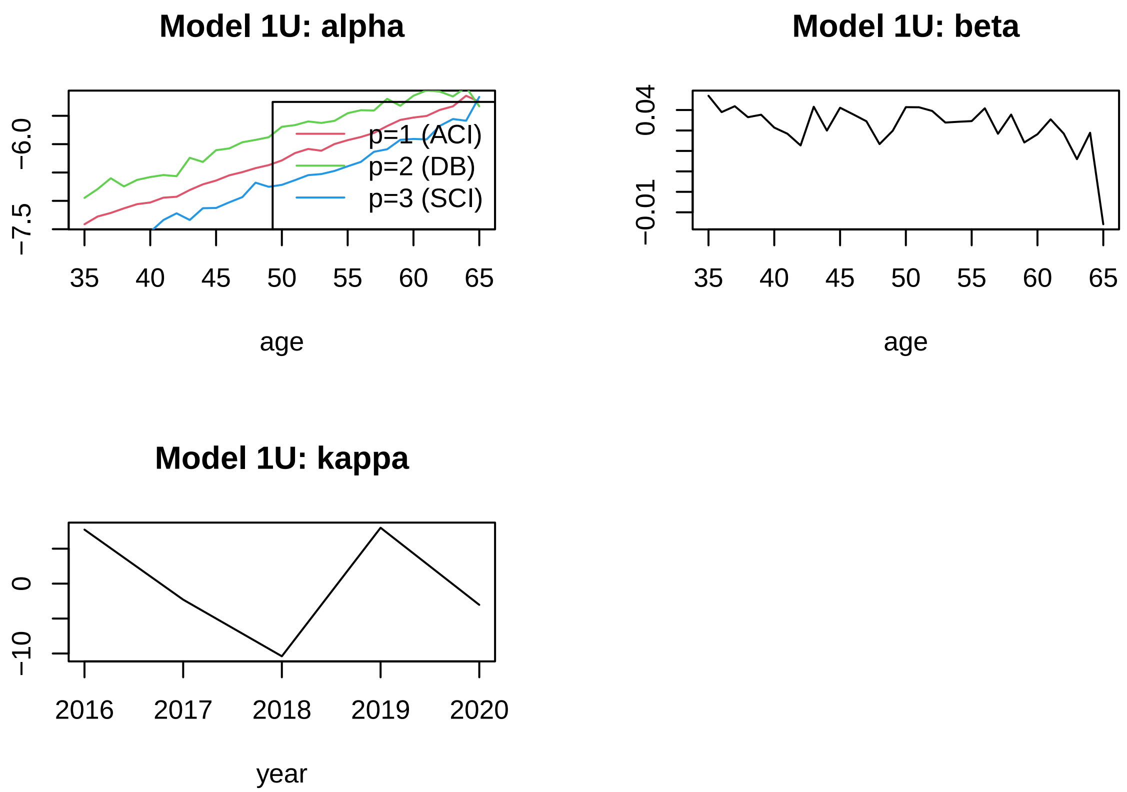
$$Constraints : \sum_{x=1}^{A} b_x = 1, \sum_{t=1}^{T} k_t = 0.$$

```
lca_subset_iterative_poisson_glm_M1U<-iteration.methodB.LC.poisson_M1U(deaths_combine=deaths_combine,
                                                                      exposure_combine=expo_combine,m=

par(mfrow=c(2,2))
plot(unique(data_subset_aggregate$Age),lca_subset_iterative_poisson_glm_M1U$alpha.prod1,col=2,type="l",
lines(unique(data_subset_aggregate$Age),lca_subset_iterative_poisson_glm_M1U$alpha.prod2,col=3)
lines(unique(data_subset_aggregate$Age),lca_subset_iterative_poisson_glm_M1U$alpha.prod3,col=4)
legend("bottomright",c("p=1 (ACI)","p=2 (DB)","p=3 (SCI)"),lty=1,col=2:4)
plot(unique(data_subset_aggregate$Age),lca_subset_iterative_poisson_glm_M1U$beta,type="l",xlab="age",ma
plot(unique(data_subset_aggregate$Year),lca_subset_iterative_poisson_glm_M1U$kappa,type="l",xlab="year"
lca_subset_iterative_poisson_glm_M1U$deviance[length(lca_subset_iterative_poisson_glm_M1U$deviance)]
```

```
## [1] 26513.65
```

```
lca_subset_iterative_poisson_glm_M1U$model_bic[length(lca_subset_iterative_poisson_glm_M1U$deviance)]
```

```
## [1] 30156.09
```

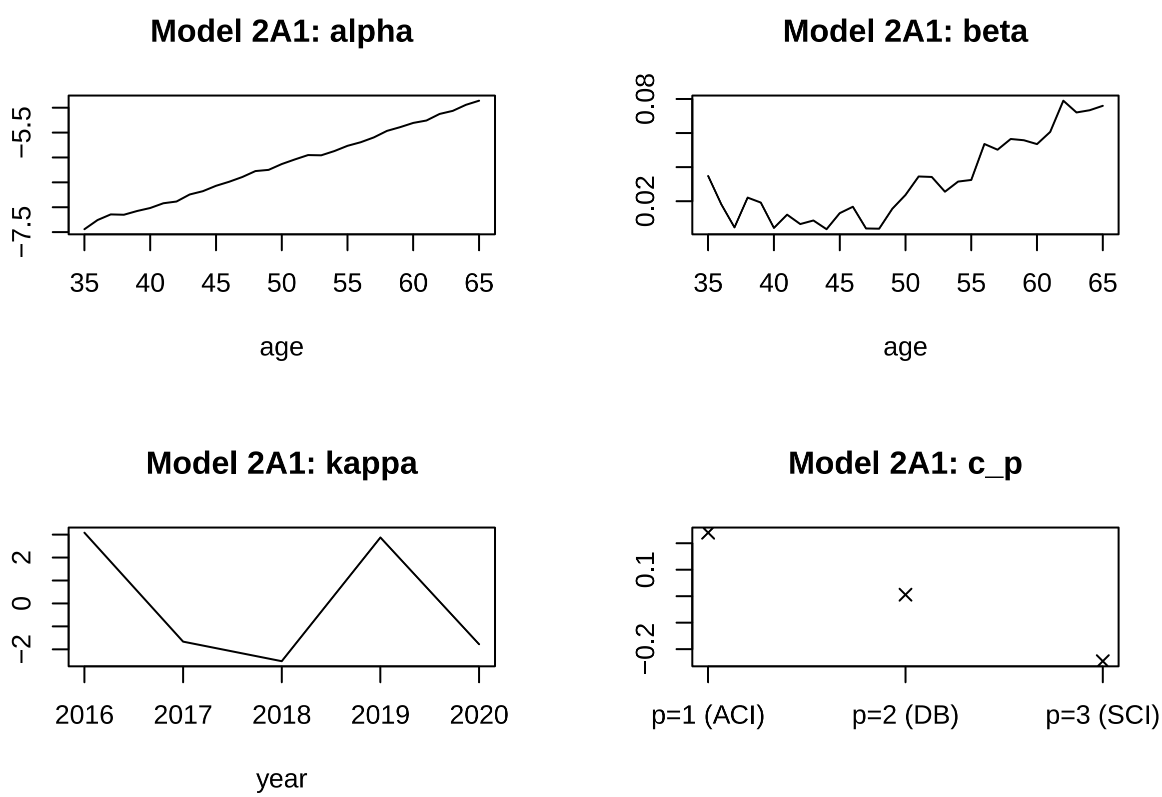


Model 2A1

$$\log(m_{x,t,p}) = a_x + (b_x + c_p)k_t$$

$$Constraints: \sum_{x=1}^{A} b_x = 1, \sum_{t=1}^{T} k_t = 0, \sum_{p=1}^{3} c_p = 0.$$

```
lca_subset_iterative_poisson_glm_M2A1<-iteration.methodB.LC.poisson_M2A1(deaths_combine=deaths_combine,
                                                                         exposure_combine=expo_combine,

par(mfrow=c(2,2))
plot(unique(data_subset_aggregate$Age),lca_subset_iterative_poisson_glm_M2A1$alpha,type="l",
     xlab="age",main="Model 2A1: alpha",ylab="")
plot(unique(data_subset_aggregate$Age),lca_subset_iterative_poisson_glm_M2A1$beta,type="l",
     xlab="age",main="Model 2A1: beta",ylab="")
plot(unique(data_subset_aggregate$Year),lca_subset_iterative_poisson_glm_M2A1$kappa,type="l",
     xlab="year",main="Model 2A1: kappa",ylab="")

plot(1:3,c(lca_subset_iterative_poisson_glm_M2A1$c.new,-sum(lca_subset_iterative_poisson_glm_M2A1$c.new
     xlab="",ylab="",xaxt="n",main="Model 2A1: c_p")
axis(1,at=1:3,label=c("p=1 (ACI)","p=2 (DB)","p=3 (SCI)"))
```




```
lca_subset_iterative_poisson_glm_M2A1$deviance[length(lca_subset_iterative_poisson_glm_M2A1$deviance)]
```

```
## [1] 12160.73
```

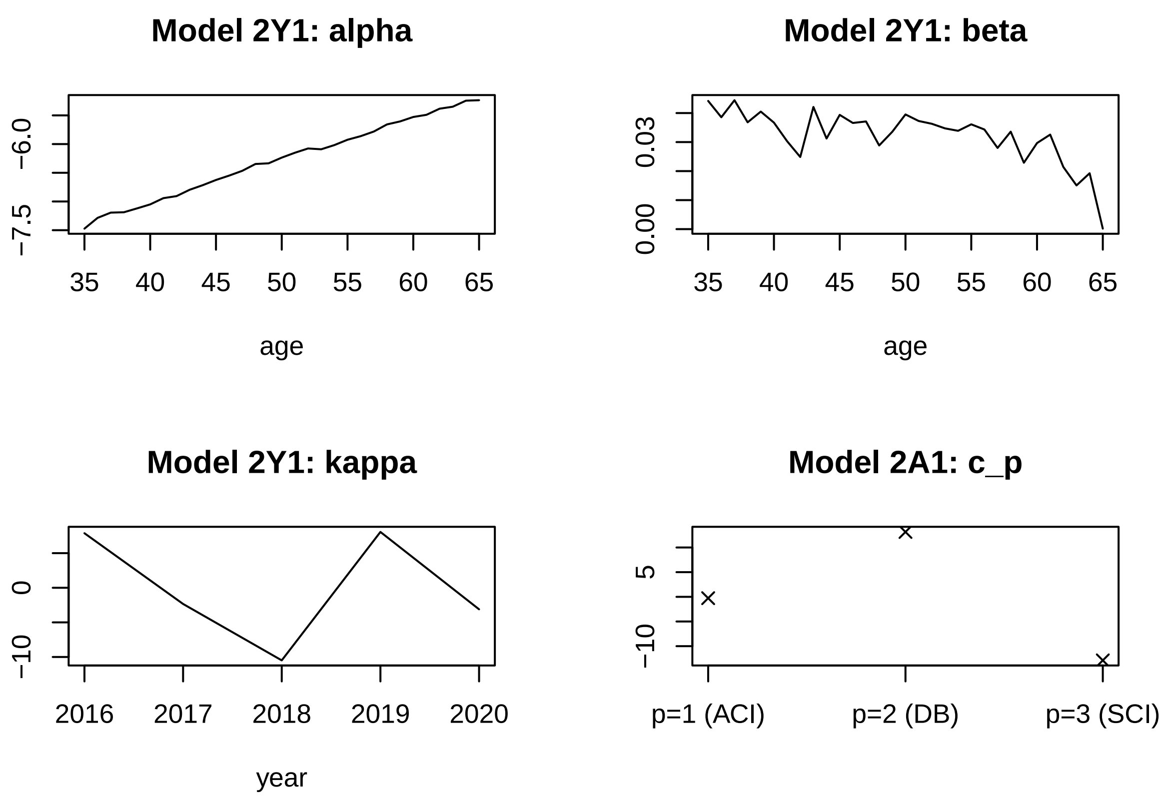
```
lca_subset_iterative_poisson_glm_M2A1$model_bic[(lca_subset_iterative_poisson_glm_M2A1$model_bic)]
```

```
##   [1] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
##  [26] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
##  [51] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
##  [76] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
```

Model 2Y1

$$\log(m_{x,t,p}) = a_x + b_x(k_t + c_p)$$
$$Constraints : \sum_{x=1}^{A} b_x = 1, \sum_{t=1}^{T} k_t = 0, \sum_{p=1}^{3} c_p = 0.$$

```
lca_subset_iterative_poisson_glm_M2Y1<-iteration.methodB.LC.poisson_M2Y1(deaths_combine=deaths_combine,
                                                                          exposure_combine=expo_combine,

par(mfrow=c(2,2))
plot(unique(data_subset_aggregate$Age),lca_subset_iterative_poisson_glm_M2Y1$alpha,type="l",
     xlab="age",main="Model 2Y1: alpha",ylab="")
plot(unique(data_subset_aggregate$Age),lca_subset_iterative_poisson_glm_M2Y1$beta,type="l",
     xlab="age",main="Model 2Y1: beta",ylab="")
plot(unique(data_subset_aggregate$Year),lca_subset_iterative_poisson_glm_M2Y1$kappa,type="l",
     xlab="year",main="Model 2Y1: kappa",ylab="")

plot(1:3,c(lca_subset_iterative_poisson_glm_M2Y1$c.new,-sum(lca_subset_iterative_poisson_glm_M2Y1$c.new
     pch=4,xlab="",ylab="",xaxt="n",main="Model 2A1: c_p")
axis(1,at=1:3,label=c("p=1 (ACI)","p=2 (DB)","p=3 (SCI)"))
```




```
lca_subset_iterative_poisson_glm_M2Y1$deviance[length(lca_subset_iterative_poisson_glm_M2Y1$deviance)]
```

```
## [1] 26595.23
```

```
lca_subset_iterative_poisson_glm_M2Y1$model_bic[length(lca_subset_iterative_poisson_glm_M2Y1$model_bic)]
```

```
## [1] 29856.87
```

Model 2A2

$$\log(m_{x,t,p}) = a_x + b_{x,p}k_t$$

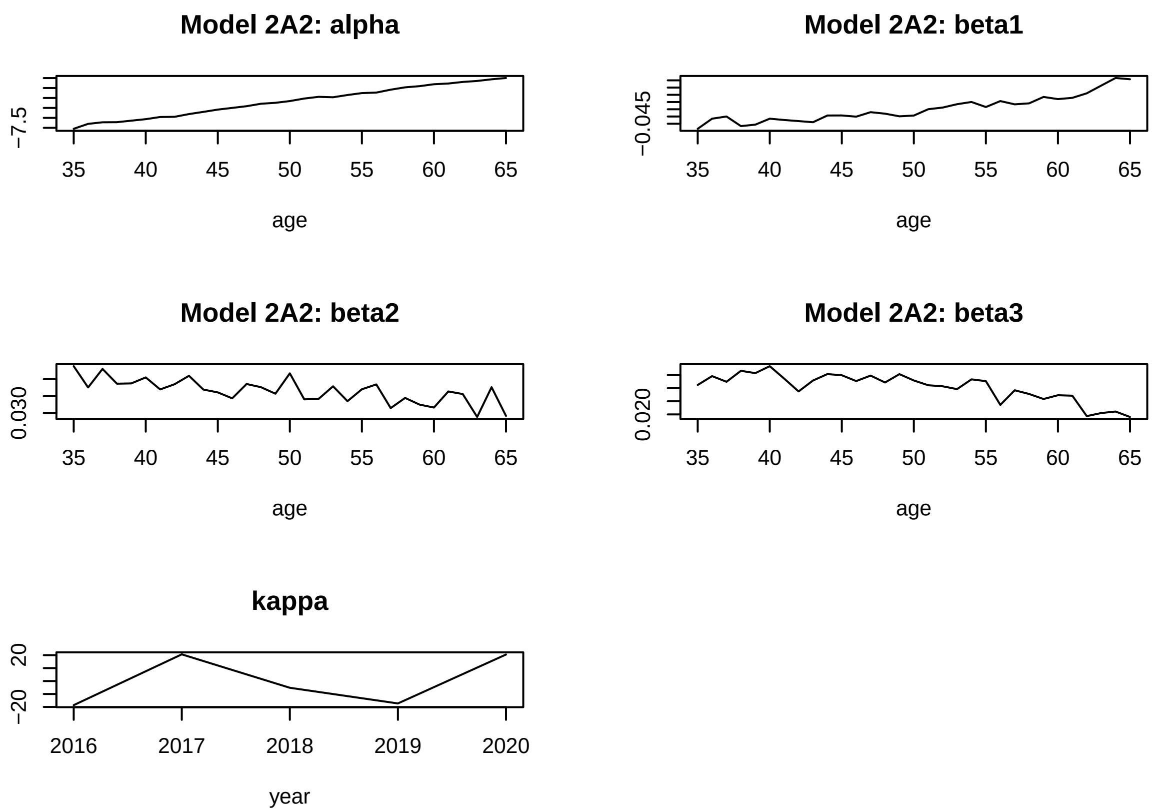
$$Constraints : \sum_{x,p} b_{x,p} = 1, \sum_{t=1}^{T} k_t = 0.$$

```
lca_subset_iterative_poisson_glm_M2A2<-iteration.methodB.LC.poisson_M2A2(deaths_combine=deaths_combine,
                                                                          exposure_combine=expo_combine,

par(mfrow=c(3,2))
plot(unique(data_subset_aggregate$Age),lca_subset_iterative_poisson_glm_M2A2$alpha,type="l",
     xlab="age",main="Model 2A2: alpha",ylab="")
plot(unique(data_subset_aggregate$Age),lca_subset_iterative_poisson_glm_M2A2$beta1,type="l",
     xlab="age",main="Model 2A2: beta1",ylab="")
plot(unique(data_subset_aggregate$Age),lca_subset_iterative_poisson_glm_M2A2$beta2,type="l",
     xlab="age",main="Model 2A2: beta2",ylab="")
plot(unique(data_subset_aggregate$Age),lca_subset_iterative_poisson_glm_M2A2$beta3,type="l",
     xlab="age",main="Model 2A2: beta3",ylab="")
plot(unique(data_subset_aggregate$Year),lca_subset_iterative_poisson_glm_M2A2$kappa,type="l",
     xlab="year",main="kappa",ylab="")
lca_subset_iterative_poisson_glm_M2A2$deviance[length(lca_subset_iterative_poisson_glm_M2A2$deviance)]
```

```
## [1] 6628.092
```

```
lca_subset_iterative_poisson_glm_M2A2$model_bic[length(lca_subset_iterative_poisson_glm_M2A2$model_bic)]
```

```
## [1] 10270.53
```



Model 2Y2

$$\log(m_{x,t,p}) = a_x + b_x k_{t,p}$$

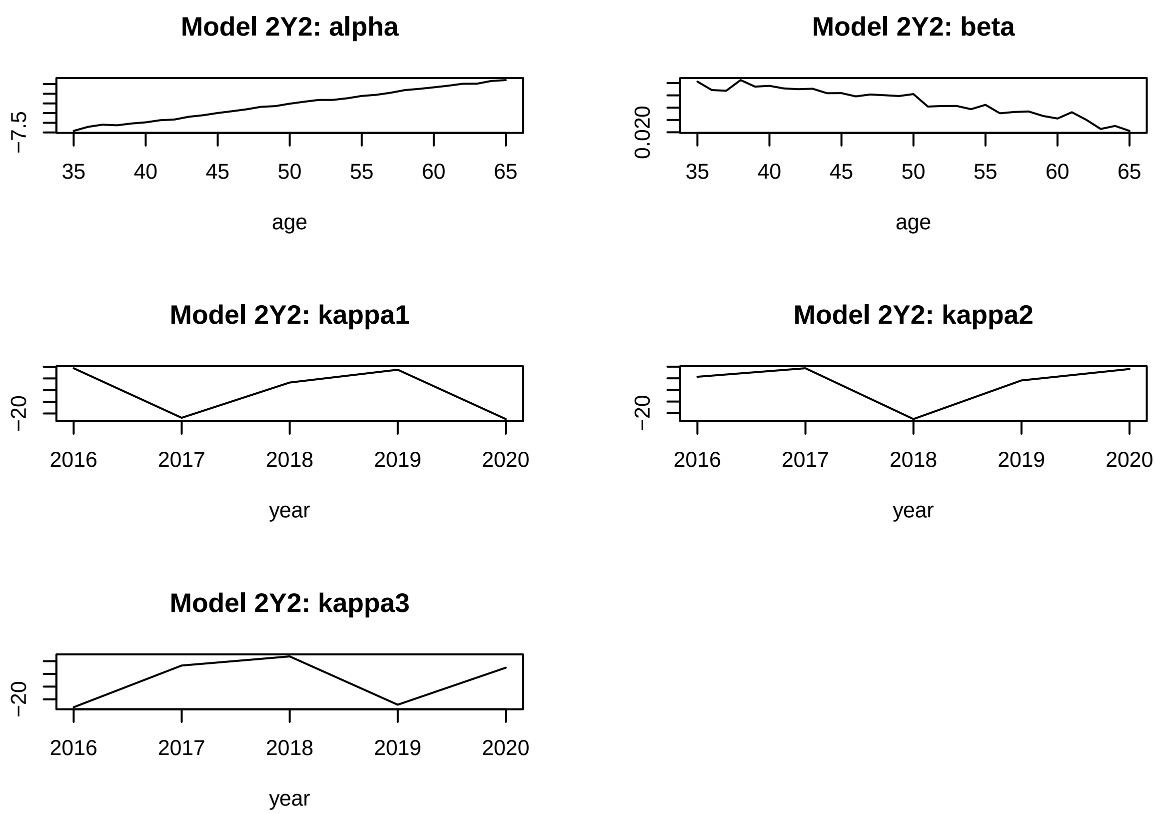
$$Constraints : \sum_{x=1}^{A} b_x = 1, \sum_{t,p} k_{t,p} = 0.$$

```
lca_subset_iterative_poisson_glm_M2Y2<-iteration.methodB.LC.poisson_M2Y2(deaths_combine=deaths_combine,
                                                                         exposure_combine=expo_combine,

par(mfrow=c(3,2))
plot(unique(data_subset_aggregate$Age),lca_subset_iterative_poisson_glm_M2Y2$alpha,type="l",
     xlab="age",main="Model 2Y2: alpha",ylab="")
plot(unique(data_subset_aggregate$Age),lca_subset_iterative_poisson_glm_M2Y2$beta,type="l",
     xlab="age",main="Model 2Y2: beta",ylab="")
plot(unique(data_subset_aggregate$Year),lca_subset_iterative_poisson_glm_M2Y2$kappa1,type="l",
     xlab="year",main="Model 2Y2: kappa1",ylab="")
plot(unique(data_subset_aggregate$Year),lca_subset_iterative_poisson_glm_M2Y2$kappa2,type="l",
     xlab="year",main="Model 2Y2: kappa2",ylab="")
plot(unique(data_subset_aggregate$Year),lca_subset_iterative_poisson_glm_M2Y2$kappa3,type="l",
     xlab="year",main="Model 2Y2: kappa3",ylab="")

lca_subset_iterative_poisson_glm_M2Y2$deviance[length(lca_subset_iterative_poisson_glm_M2Y2$deviance)]
```

```
## [1] 512.8938
```

```
lca_subset_iterative_poisson_glm_M2Y2$model_bic[length(lca_subset_iterative_poisson_glm_M2Y2$model_bice
```

```
## numeric(0)
```



Li and Lee (2005) Model

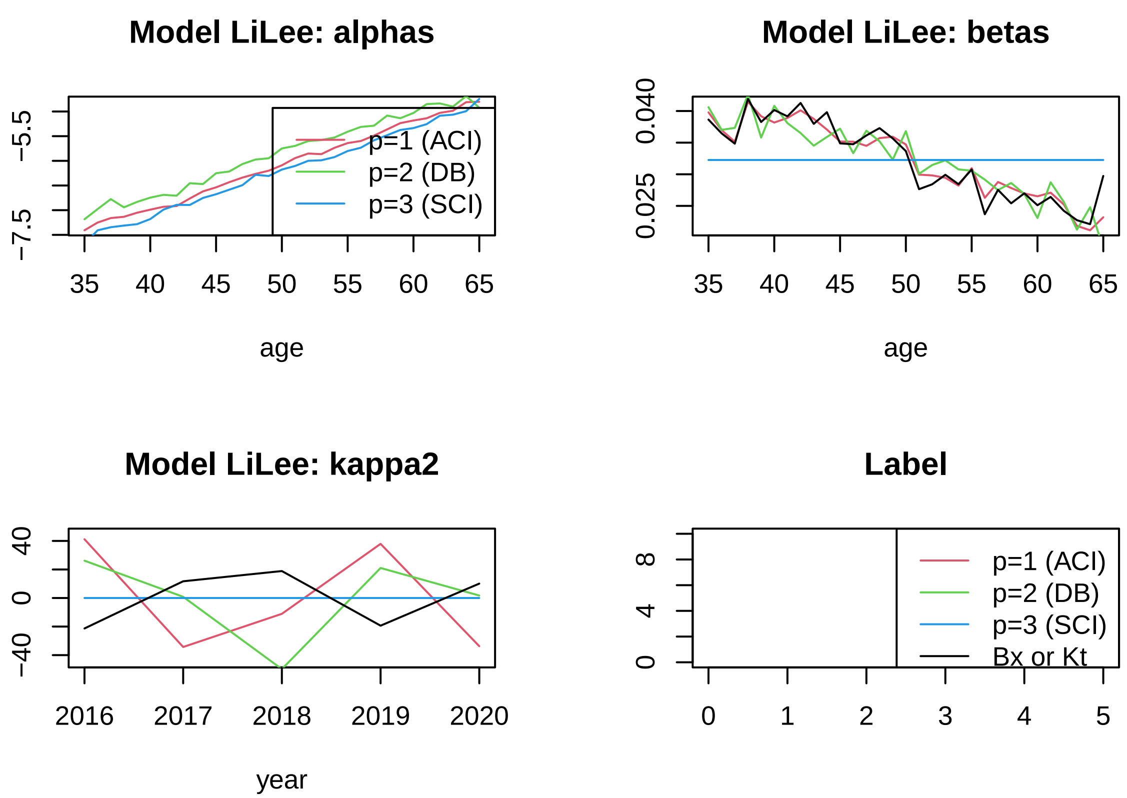
$$\log(m_{x,t,p}) = a_{x,p} + b_{x,p}k_{t,p} + B_xK_t$$

$$Constraints: \sum_{x=1}^{A} b_{x,1} = 1, \sum_{x=1}^{A} b_{x,2} = 1, \sum_{t=1}^{T} k_{t,1} = 0, \sum_{t=1}^{T} k_{t,2} = 0, \sum_{x=1}^{A} B_x = 1, \sum_{t=1}^{T} K_t = 0.$$

We additionally need to impose $b_{x,3} = 0$ and $k_{t,3} = 0$ for all x and t, corresponding to setting the additional product-specific bi-linear term for product 3 to be zero. In other words, we use product 3 as the reference group (i.e. $\log(m_{x,t,3}) = a_{x,3} + B_xK_t$).

```
lca_subset_iterative_poisson_glm_lilee<-iteration.methodB.LC.poisson_lilee(deaths_combine=deaths_combine
                                                                              exposure_combine=expo_combine

par(mfrow=c(2,2))
plot(unique(data_subset_aggregate$Age),lca_subset_iterative_poisson_glm_lilee$alpha1,type="l",col=2,
     xlab="age",main="Model LiLee: alphas",ylab="")
lines(unique(data_subset_aggregate$Age),lca_subset_iterative_poisson_glm_lilee$alpha2,type="l",col=3,
      xlab="age",main="Model LiLee: alpha2",ylab="")
lines(unique(data_subset_aggregate$Age),lca_subset_iterative_poisson_glm_lilee$alpha3,type="l",col=4,
```

```
      xlab="age",main="Model LiLee: alpha3",ylab="")
legend("bottomright",c("p=1 (ACI)","p=2 (DB)","p=3 (SCI)"),lty=1,col=2:4)

plot(unique(data_subset_aggregate$Age),lca_subset_iterative_poisson_glm_lilee$beta1,col=2,type="l",
     xlab="age",main="Model LiLee: betas",ylab="")
lines(unique(data_subset_aggregate$Age),lca_subset_iterative_poisson_glm_lilee$beta2,col=3,type="l",
      xlab="age",main="Model LiLee: beta2",ylab="")
lines(unique(data_subset_aggregate$Age),lca_subset_iterative_poisson_glm_lilee$beta3,col=4,type="l",
      xlab="age",main="Model LiLee: beta3",ylab="")
lines(unique(data_subset_aggregate$Age),lca_subset_iterative_poisson_glm_lilee$Beta,type="l",xlab="age"

plot(unique(data_subset_aggregate$Year),lca_subset_iterative_poisson_glm_lilee$kappa1,col=2,type="l",
     ylim=c(-45,45),xlab="year",main="Model LiLee: kappa2",ylab="")
lines(unique(data_subset_aggregate$Year),lca_subset_iterative_poisson_glm_lilee$kappa2,col=3,type="l",
      xlab="year",main="Model LiLee: kappa2",ylab="")
lines(unique(data_subset_aggregate$Year),lca_subset_iterative_poisson_glm_lilee$kappa3,col=4,type="l",
      xlab="year",main="Model LiLee: kappa3",ylab="")
lines(unique(data_subset_aggregate$Year),lca_subset_iterative_poisson_glm_lilee$Kappa,type="l",
      xlab="year",main="Model LiLee: Kappa",ylab="")
plot(NULL,xlim=c(0,5),ylim=c(0,10),xlab="",ylab="",main="Label")
legend("topright",c("p=1 (ACI)","p=2 (DB)","p=3 (SCI)","Bx or Kt"),lty=1,col=c(2:4,1))
```




```
lca_subset_iterative_poisson_glm_lilee$deviance[length(lca_subset_iterative_poisson_glm_lilee$deviance)]
```

```
## [1] 368.1237
```

```
lca_subset_iterative_poisson_glm_lilee$model_bic[length(lca_subset_iterative_poisson_glm_lilee$model_bic
```

```
## [1] 4379.088
```

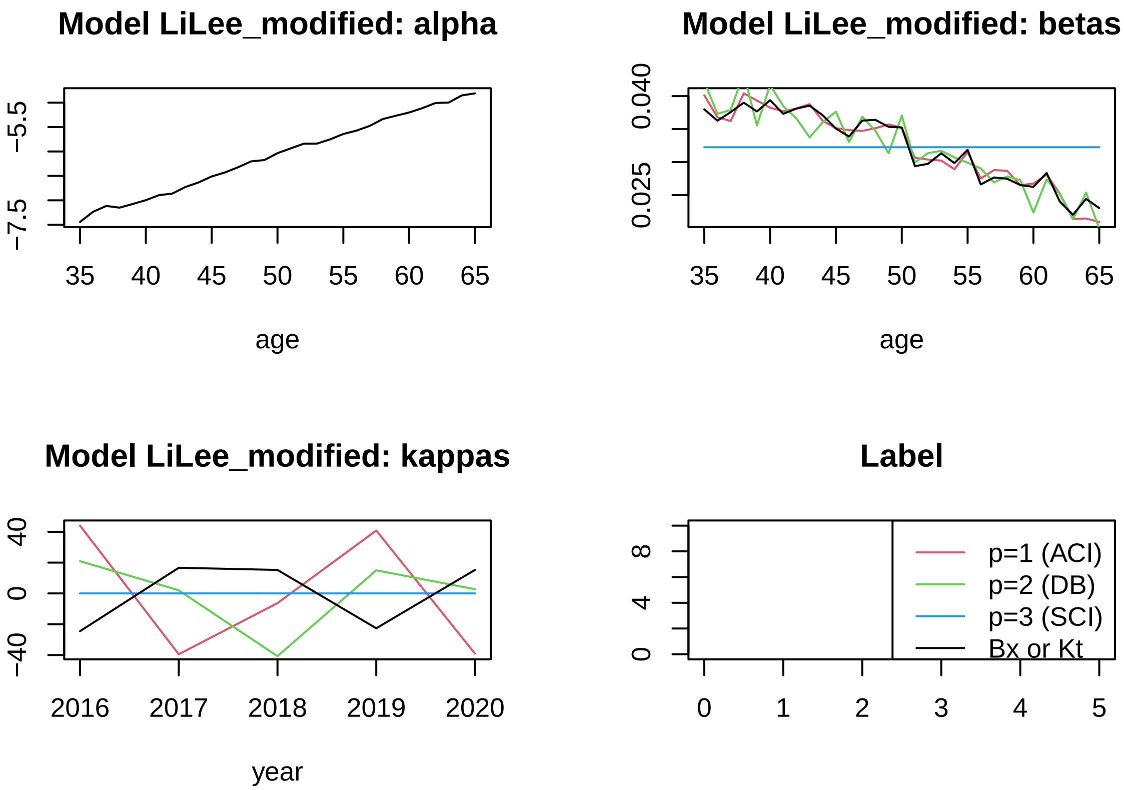
Modified Li and Lee (2005) Model

This model is a simple special of the original Li and Lee (2005) model, but the age-specific term a_x (representing average log-mortality rates) is shared between the products.

$$\log(m_{x,t,p}) = a_x + b_{x,p}k_{t,p} + B_x K_t$$
$$Constraints : \sum_{x=1}^{A} b_{x,1} = 1, \sum_{x=1}^{A} b_{x,2} = 1, \sum_{t=1}^{T} k_{t,1} = 0, \sum_{t=1}^{T} k_{t,2} = 0, \sum_{x=1}^{A} B_x = 1, \sum_{t=1}^{T} K_t = 0.$$

Similarly, we impose $b_{x,3} = 0$ and $k_{t,3} = 0$ for all x and t as explained above.

```
lca_subset_iterative_poisson_glm_lilee_modified<-iteration.methodB.LC.poisson_lilee_modified(deaths_comb
                                                                                             exposure_c

par(mfrow=c(2,2))
plot(unique(data_subset_aggregate$Age),lca_subset_iterative_poisson_glm_lilee_modified$alpha,type="l",
     xlab="age",main="Model LiLee_modified: alpha",ylab="")

plot(unique(data_subset_aggregate$Age),lca_subset_iterative_poisson_glm_lilee_modified$beta1,col=2,
     type="l",xlab="age",main="Model LiLee_modified: betas",ylab="")
lines(unique(data_subset_aggregate$Age),lca_subset_iterative_poisson_glm_lilee_modified$beta2,col=3,
      type="l",xlab="age",main="Model LiLee_modified: beta2",ylab="")
lines(unique(data_subset_aggregate$Age),lca_subset_iterative_poisson_glm_lilee_modified$beta3,col=4,
      type="l",xlab="age",main="Model LiLee_modified: beta3",ylab="")
lines(unique(data_subset_aggregate$Age),lca_subset_iterative_poisson_glm_lilee_modified$Beta,type="l",
      xlab="age",main="Model LiLee_modified: Beta",ylab="")

plot(unique(data_subset_aggregate$Year),lca_subset_iterative_poisson_glm_lilee_modified$kappa1,col=2,
     type="l",xlab="year",main="Model LiLee_modified: kappas",ylab="")
lines(unique(data_subset_aggregate$Year),lca_subset_iterative_poisson_glm_lilee_modified$kappa2,col=3,
      type="l",xlab="year",main="Model LiLee_modified: kappa2",ylab="")
lines(unique(data_subset_aggregate$Year),lca_subset_iterative_poisson_glm_lilee_modified$kappa3,col=4,
      type="l",xlab="year",main="Model LiLee_modified: kappa3",ylab="")
lines(unique(data_subset_aggregate$Year),lca_subset_iterative_poisson_glm_lilee_modified$Kappa,type="l"
      xlab="year",main="Model LiLee_modified: Kappa",ylab="")
plot(NULL,xlim=c(0,5),ylim=c(0,10),xlab="",ylab="",main="Label")
legend("topright",c("p=1 (ACI)","p=2 (DB)","p=3 (SCI)","Bx or Kt"),lty=1,col=c(2:4,1))
```




```
lca_subset_iterative_poisson_glm_lilee_modified$deviance[length(lca_subset_iterative_poisson_glm_lilee_r
```

```
## [1] 816.028
```

```
lca_subset_iterative_poisson_glm_lilee_modified$model_bic[length(lca_subset_iterative_poisson_glm_lilee_
```

```
## [1] 4446.186
```

References

Brouhns, N., Denuit, M., & Vermunt, J. K. (2002). A Poisson log-bilinear regression approach to the construction of projected lifetables. Insurance: Mathematics and Economics, 31(3), 373-393.