Life Insurance
– Tutorial Exercises –

Christoph Hambel

Tilburg University
Tilburg School of Economics and Management
Department of Econometrics and Operations Research

Spring Term 2023

Problem 1

A few years ago, somebody observed that the average age at which people in the Netherlands died each year was around 70 years, and slightly decreasing over the recent years.

1. If life expectancy would be increasing, how would it then be possible that the average age at which people in the Netherlands die (around 70 years) was slightly decreasing over the recent years?

Problem 2

Johan de Witt (1671) was the first to propose to calculate single life annuity values using survival probabilities. He claimed that compared to government bonds (in old Dutch: losrenten), with price 25 guilders (and payoff 1 guilder per year, at the end of each year), the price of a single life annuity (lijfrenten) of 14 guilders was too low, and should be set higher than 16 guilders (payoff 1 guilder, at the end of the year, per year alive). Assume a flat term structure in 1671.

1. What was the interest rate in 1671?
2. Consider a single life annuity in 1671 for somebody whose age is $x = 3$ and priced according to the calculations by Johan de Witt. Assume ${}_\tau p_{t,x}^{(g)} = 1$ for $\tau \leq T$ and ${}_\tau p_{t,x}^{(g)} = 0$ for $\tau > T$ for this person. What is T?

Problem 3

A pool of 1,000 people invests 1,000,000 Euro (each contributing 1,000 Euro). The return of the investment turns out to be 5%. However, only 97% of the people survive.

1. What is the return per surviving individual?
2. The biometric return is the return per surviving individual divided by the return on the investment (using gross returns). How much is the biometric return?

Problem 4

We are now dealing with the COVID-19 pandemic.

1. Would you classify this pandemic as micro- or under macro- longevity risk? Motivate your answer.
2. Would you classify this pandemic as a permanent or as a transitory shock? Motivate your answer.

Problem 5

Assume that somebody of age x dies $T_{x,t}^{(g)}$ years from now.
The density function of $T_{x,t}^{(g)}$ is given by $f_{x,t}^{(g)}$ and the (cumulative) distribution function is given by $F_{x,t}^{(g)}$.
Assume that the maximum age is 120, i.e., $F_{x,t}^{(g)}(120-x)=1$. Let $e_{x,t}^{(g)}$ denote the expected remaining lifetime of somebody of age x.

1. Apply integration by parts to verify that

$$e_{x,t}^{(g)} = \int_0^\infty \tau f_{x,t}^{(g)}(\tau)\mathrm{d}\tau = \int_0^\infty [1-F_{x,t}^{(g)}(\tau)]\mathrm{d}\tau$$

2. How to interpret this equality?

Problem 6

Let ${}_\tau q_{x,t}^{(g)}$ be the probability that somebody of age x dies τ years from now and let ${}_\tau p_{x,t}^{(g)}$ be the probability that somebody of age x survives at least τ years from now. Let $e_{x,t}^{(g)}$ denote the expected remaining lifetime of somebody of age x.

1. Verify

$$e_{x,t}^{(g)} = \sum_{\tau=1}^{\infty} {}_\tau q_{x,t}^{(g)} = \sum_{\tau=1}^{\infty} p_{x,t}^{(g)}$$

Problem 7

Suppose you do not observe the number of deaths in the period Lexis diagrams (left hand side, illustration below) but in the cohort Lexis diagrams (right hand side, illustration below). How would you calculate the number of deaths in the period Lexis diagrams?

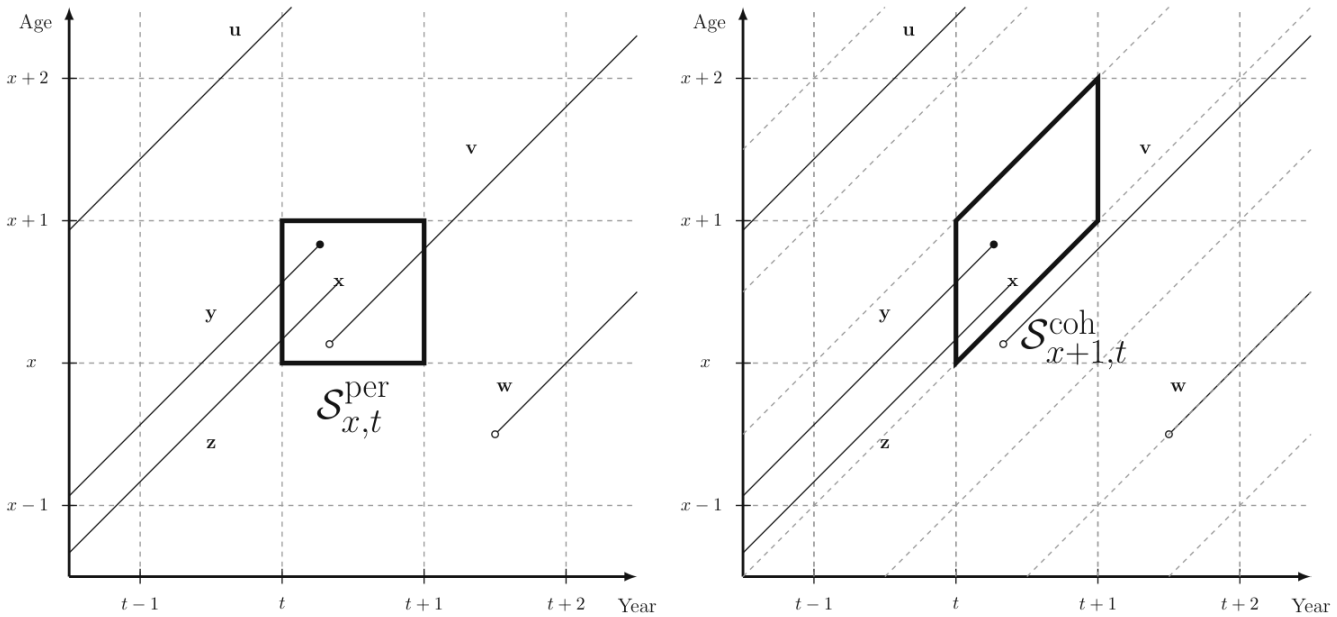


Fig. 1 Lexis diagram with life lines `u, v, w, y, z` and representation of period (left) and cohort (right) surfaces. `x` marks the death of a person, • indicates a person dropping out of the study alive, ∘ specifies a new person entering the study. In our study, only birth and death are observed

Problem 8

An alternative to using the Lee-Carter model (or the model of the Actuarial Society) would be to estimate each year the parameters $a_t^{(g)}, \ldots, h_t^{(g)}$ of the "Heligman-Pollard Law," to make a model for the evolution of these estimated parameters over time, and to use this model to forecast future mortality. The Heligman-Pollard law is given by

$$\frac{q_{x,t}^{(g)}}{p_{x,t}^{(g)}} = (a_t^{(g)})^{(x+b_t^{(g)})^{c_t^{(g)}}} + d_t^{(g)} \exp\left(- e_t^{(g)} (\ln x - \ln f_t^{(g)})^2 \right) + \frac{g_t^{(g)} (h_t^{(g)})^x}{1 + g_t^{(g)} (h_t^{(g)})^x}.$$

What are the advantages/disadvantages of this approach compared to Lee-Carter approach?

Problem 9

Assume the Lee-Carter normalization ($\sum_{x\in\mathcal{X}} \beta_x^{(g)} = 1$ and $\sum_{t\in\mathcal{T}} \kappa_t^{(g)} = 0$)
Motivate why estimation of $\alpha_x^{(g)}, \beta_x^{(g)}, \kappa_t^{(g)}$ by the Singular Value Decomposition (SVD) is the same as minimizing the sum of squared errors

$$\min_{\alpha_x^{(g)},\beta_x^{(g)},\kappa_t^{(g)}} \sum_{x,t} \left[\ln(m_{x,t}^{(g)}) - \alpha_x^{(g)} - \beta_x^{(g)} \kappa_t^{(g)} \right]^2$$

w.r.t $\alpha_x^{(g)}, \beta_x^{(g)}, \kappa_t^{(g)}$.

Problem 10

If one applies the Lee-Carter model to different groups g then the resulting trends in $\kappa_t^{(g)}$ might diverge (following from modeling $\kappa_t^{(g)}$ as a random walk with drift).

1. Discuss why this might happen.
2. What will happen with the corresponding one-year death probabilities $q_{x,t}^{(g)}$ in the long run?

Problem 11

1. Start from the traditional Lee-Carter normalization, i.e., $\sum_{x\in\mathcal{X}} \beta_x^{(g)} = 1$ and $\sum_{t\in\mathcal{T}} \kappa_t^{(g)} = 0$. How to transform the values of $\alpha_x^{(g)}, \beta_x^{(g)}, \kappa_t^{(g)}$ to arrive at the normalization of Liu et al. (2019a b)?

2. Conversely, start from the normalization of Liu et al. (2019a b), i.e., $\sum_{x\in\mathcal{X}} \beta_x^{(g)} = 1$ and $\sum_{x\in\mathcal{X}} \alpha_x^{(g)} = 0$. How to transform the values of $\alpha_x^{(g)}, \beta_x^{(g)}, \kappa_t^{(g)}$ to arrive at the Lee-Carter normalization?

Problem 12

Girosi and King (2006, 2008, ...) proposed a reformulation of the Lee-Carter model, by combining

$$\ell_t^{(g)} = \begin{pmatrix} \ln m_{0,t}^{(g)} \\ \vdots \\ \ln m_{X,t}^{(g)} \end{pmatrix}, \qquad \Delta\kappa_t^{(g)} = c^{(g)} + \delta_t^{(g)}$$

Let $\alpha^{(g)} = \begin{pmatrix} \alpha_0^{(g)} \\ \vdots \\ \alpha_X^{(g)} \end{pmatrix}$, $\beta^{(g)} = \begin{pmatrix} \beta_0^{(g)} \\ \vdots \\ \beta_X^{(g)} \end{pmatrix}$, $\varepsilon_t^{(g)} = \begin{pmatrix} \varepsilon_{0,t}^{(g)} \\ \vdots \\ \varepsilon_{X,t}^{(g)} \end{pmatrix}$. Then

$$\begin{aligned} \ell_t^{(g)} &= \alpha^{(g)} + \beta^{(g)}\kappa_t^{(g)} + \varepsilon_t^{(g)} \\ &= \alpha^{(g)} + \beta^{(g)}[\kappa_{t-1}^{(g)} + c^{(g)} + \delta_t^{(g)}] + \varepsilon_t^{(g)}. \end{aligned}$$

Problem 12 cont'd

1. Rewrite the last equation into the following form

$$\Delta \ell_t^{(g)} = \theta^{(g)} + \zeta_t^{(g)}$$

where $\theta^{(g)}$ is a vector of parameters and $\zeta_t^{(g)}$ a vector of error terms. What is What is the link between $\theta^{(g)}$ and the original Lee- Carter parameters and what is the link between $\zeta_t^{(g)}$ and the original Lee-Carter error terms?

2. Suppose you would assume that $\zeta_t^{(g)}$ are i.i.d. and normally distributed with zero mean vector. How would you estimate $\theta^{(g)}$? How would you quantify parameter risk, i.e., the risk due to sampling error?

3. Do you think the assumption $\zeta_t^{(g)}$ are i.i.d. is reasonable (given the underlying Lee-Carter model)?

Problem 13

Start from

$$Z_t^{(g)} = c^{(g)} + Z_{t-1}^{(g)} + \left(\delta_t^{(g)} + e_t^{(g)} - e_{t-1}^{(g)}\right)$$
$$\ln(m_{x,t}^{(g)}) = \alpha_x^{(g)} + \beta_x^{(g)} Z_t^{(g)} + \left(\varepsilon_{x,t}^{(g)} - \beta_x^{(g)} e_t^{(g)}\right)$$

Discuss how you would estimate these equations in a single step, applying standard linear regression techniques (possibly using instrumental variables).

Problem 14

1. How would you estimate the following extended Lee-Carter models:
   - Extra factor:

$$\ln(m_{x,t}^{(g)}) = \alpha_x^{(g)} + \beta_{1,x}^{(g)}\kappa_{1,t}^{(g)} + \beta_{2,x}^{(g)}\kappa_{2,t}^{(g)} + \varepsilon_{x,t}^{(g)}$$

   - Cohort effect:

$$\ln(m_{x,t}^{(g)}) = \alpha_x^{(g)} + \beta_x^{(g)}\kappa_t^{(g)} + \gamma_x^{(g)}\iota_{t-x}^{(g)} + \varepsilon_{x,t}^{(g)}$$

2. Do we need (additional) normalization assumptions? Which one(s)?
3. How could one investigate whether a cohort effect might be relevant using the output of the standard Lee-Carter model (without cohort effect)?

Problem 15

Consider the CBD-model

$$\ln\left(\frac{q_{x,t}^{(g)}}{p_{x,t}^{(g)}}\right) = \kappa_{1,t}^{(g)} + \kappa_{2,t}^{(g)} x + \varepsilon_{x,t}$$

1. How would you estimate $\kappa_{1,t}^{(g)}$ and $\kappa_{2,t}^{(g)}$ (after imposing appropriate distributional assumptions)?
2. How would you proceed to forecast future mortality using the CBD-model?
3. How to obtain $q_{x,t}^{(g)}$ from $y = \ln\left(\frac{q_{x,t}^{(g)}}{p_{x,t}^{(g)}}\right)$?

Problem 16

Somebody proposes the following extension of the Lee-Carter model

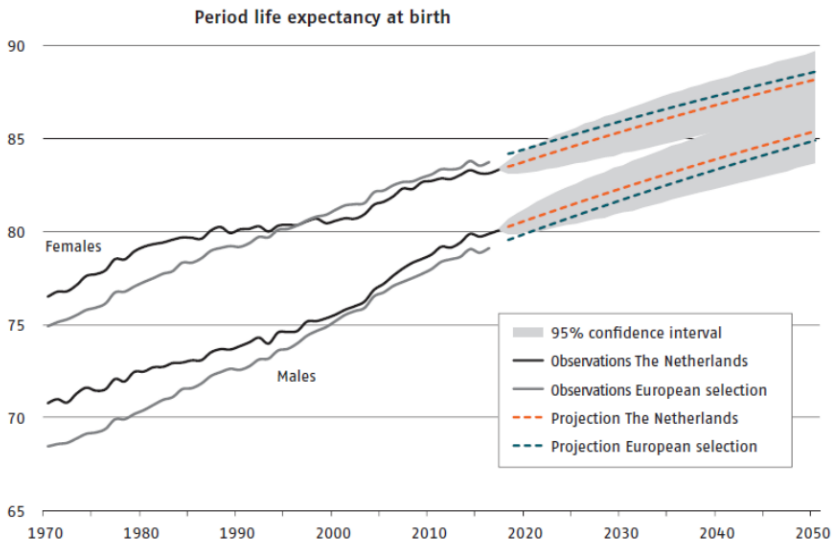
$$\ln(m_{x,t}^{(g)}) = \alpha_x^{(g)} + \beta_{1,x}^{(g)}\kappa_{1,t}^{(g)} + \beta_{2,x}^{(g)}\kappa_{2,t}^{(g)} + \varepsilon_{x,t}^{(g)}$$

with normalizations $\sum_{x\in\mathcal{X}} \beta_{1,x}^{(g)} = 1$, $\sum_{x\in\mathcal{X}} \beta_{2,x}^{(g)} = 1$ and $\sum_{x\in\mathcal{X}} \alpha_x^{(g)} = 0$.

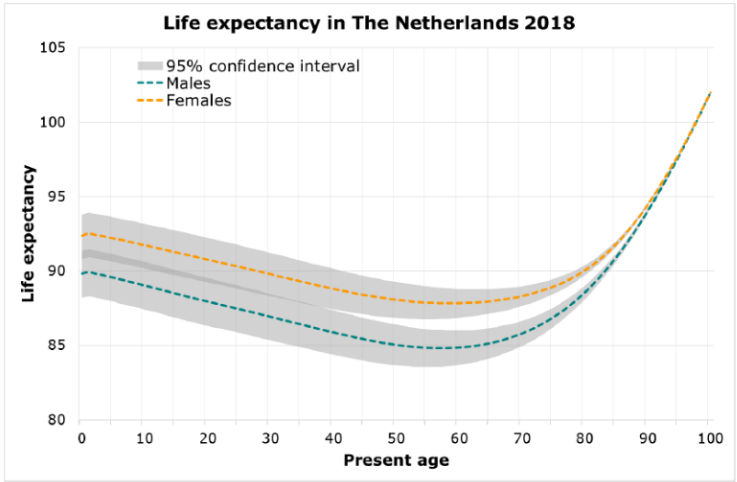
1. Show that this extension with the given normalizations is not identified.
2. Can you propose an extra normalization that identifies this extension of the Lee-Carter model? You do not have to prove that your choice guarantees identification, but you have to motivate your choice!

Problem 17

The graphs below illustrate macro longevity according to AG2018 (cf. the similar graphs based on Lee-Carter). How would you construct such graphs in case of AG2022?



Graph 9.1 Confidence interval around the best estimate of the period life expectancy for Dutch males and females



Graph 9.2: *Confidence interval around the best estimate cohort life expectancy of Dutch males and females in 2018*

Problem 18

Instead of a three-layer Li-Lee model, one could also have used for the AG2022 model a standard two-layer Li-Lee model, i.e., just using the AG2020 model specification, but re-estimating its parameters using also the (available) data of 2020 (EU and NL) and 2021 (NL only).

1. Can you describe the possible outcome of such a version of the AG2022 model?
2. Use your answer to a) to motivate the choice for the current version of AG2022.

Problem 19

1. How would you aggregate the week effects $\mathfrak{K}_{w,t}^{(g)}$ of the COVID-19 layer to the year effects $\mathfrak{X}_t^{(g)}$? What about the week age effects $\mathfrak{B}_t^{(g)}$ compared to the corresponding annualized age effects $\widetilde{\mathfrak{B}}_t^{(g)}$?
2. Using the new closing method, the confidence interval does not become smaller. However, would you not expect it to become wider? Can you motivate the resulting shape of the confidence interval?

Problem 20

In the AG2020-report, you can find the graph below, showing period life expectancies in-sample and projected into the future. In case of females, the difference between the European and Dutch seems to increase. In case of males the difference seems to decrease. How would you expect the projections will evolve further into the future? Motivate your answer!