TOOLS 4F

Czech Mortality Predictions

focused on pension insurance

Martin Matějka, Pavel Zimmermann

MFF UK, Praha 2014

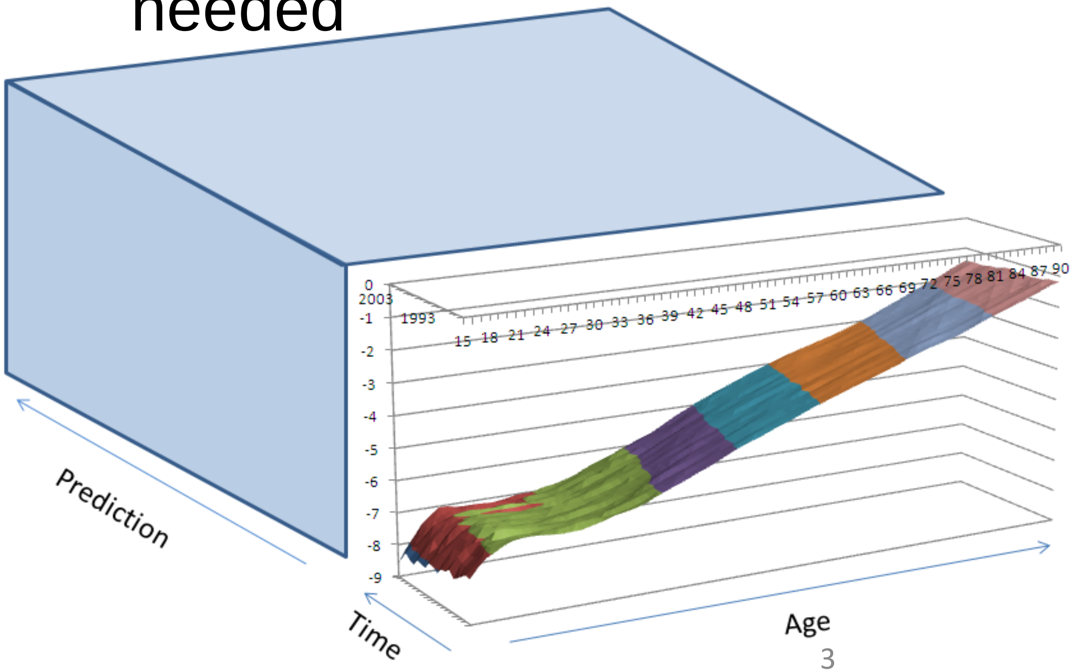
WE UNDERSTAND YOUR JOB

Agenda

- Historical Context
- Convergence
- Mortality Models
  - Adult ages
  - High ages
- Application to the Czech data
- Further extensions

Life table prediction

- Life tables are changing over time
- The life table prediction should be
  - based on the historical data
  - (expert) assumptions about future behavior are needed



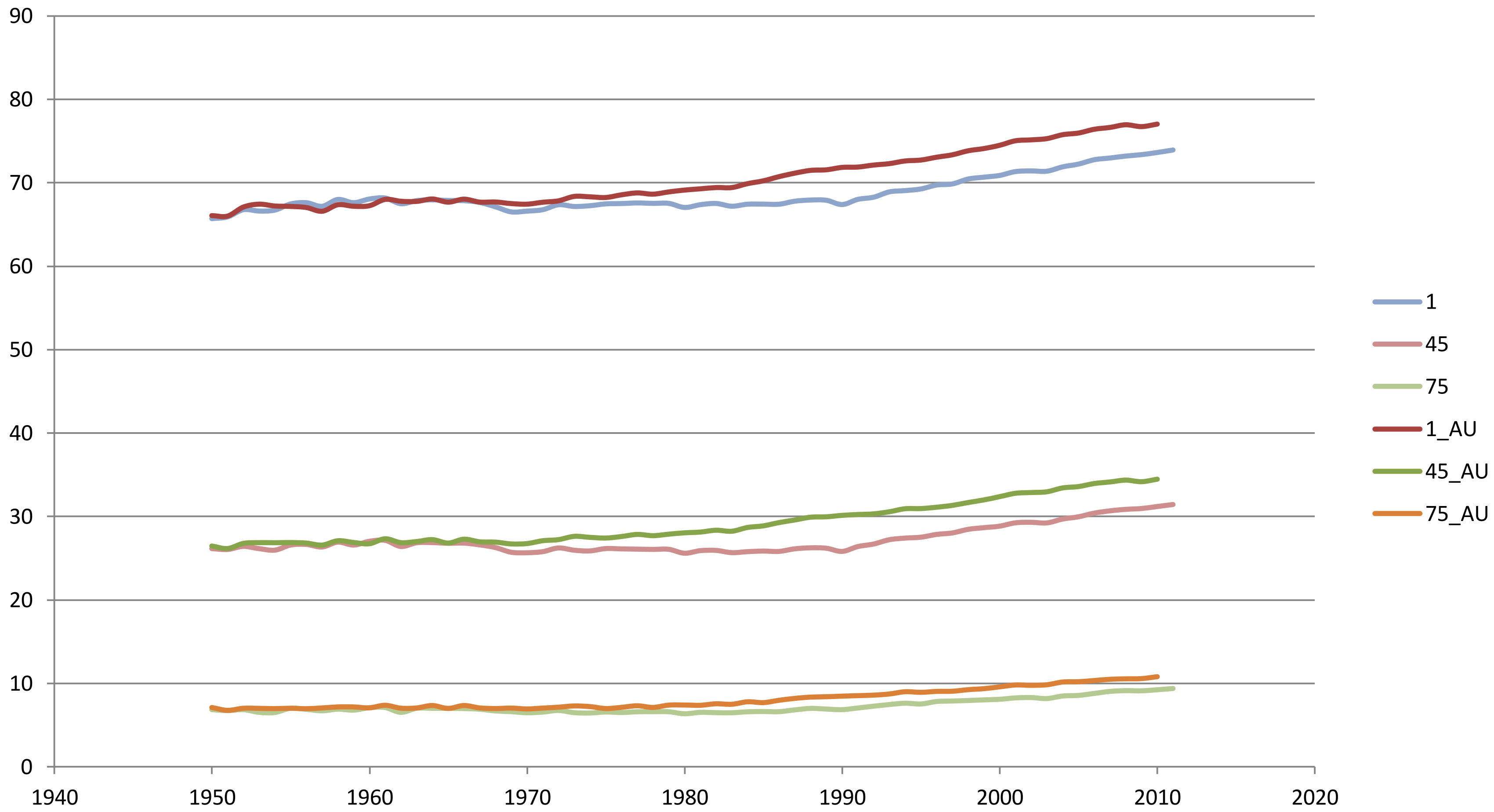
Historical Context

Historical Context

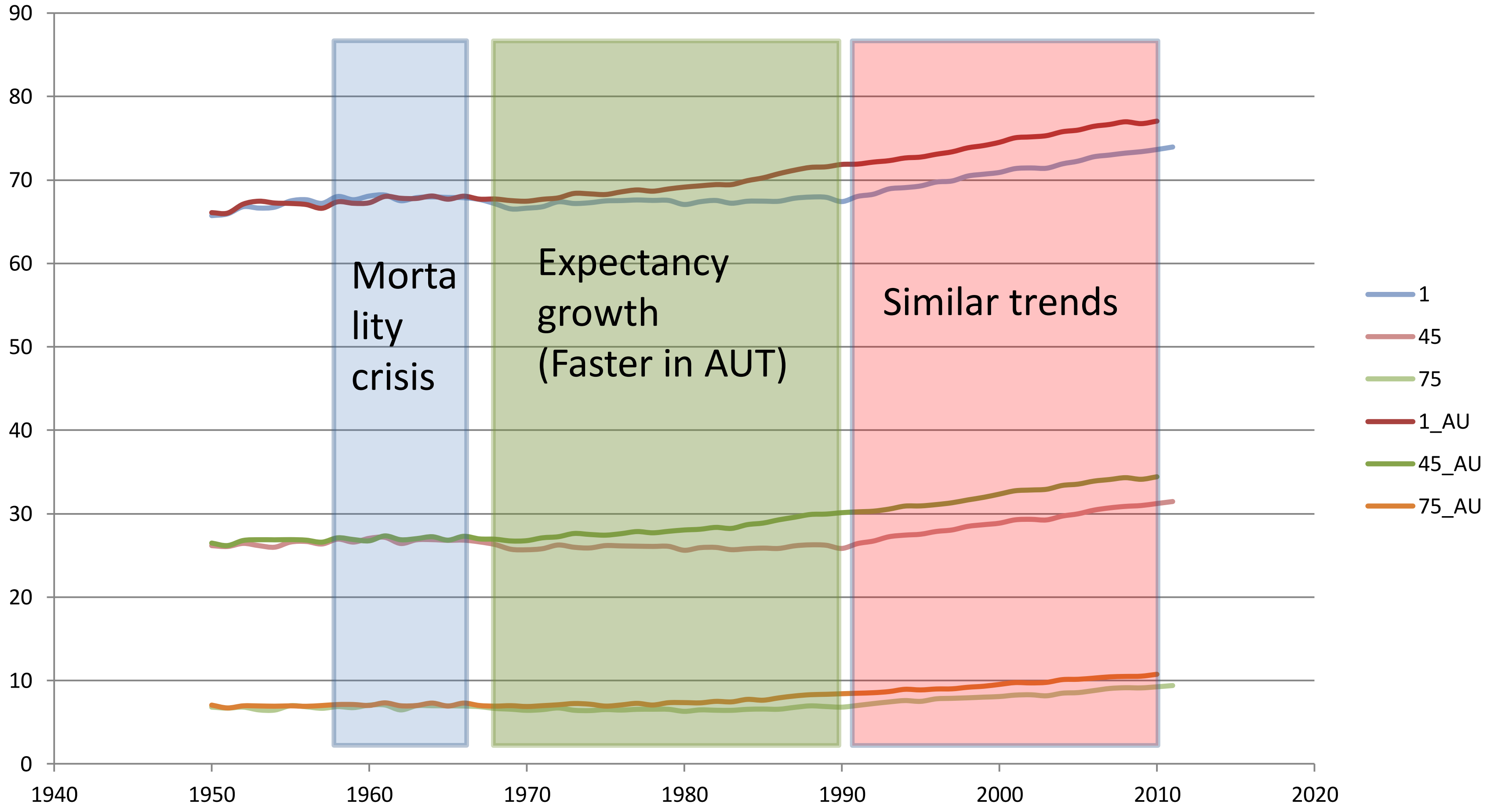
In western Europe and CZE after WWII the development of life expectancy was at first comparable

Then the difference started to grow…

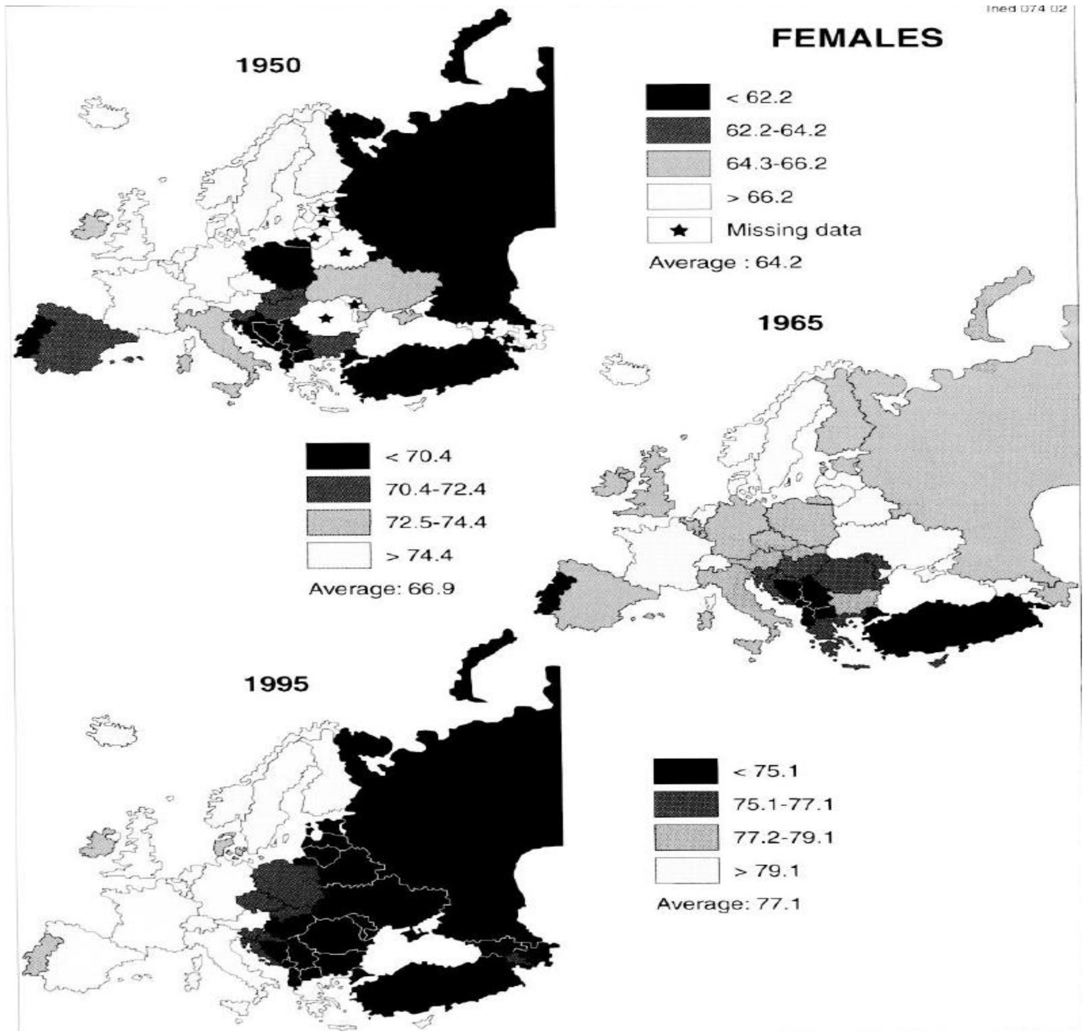
Life expectancies males CZE and AUT



Life expectancies males CZE and AUT



TOOLS 4F

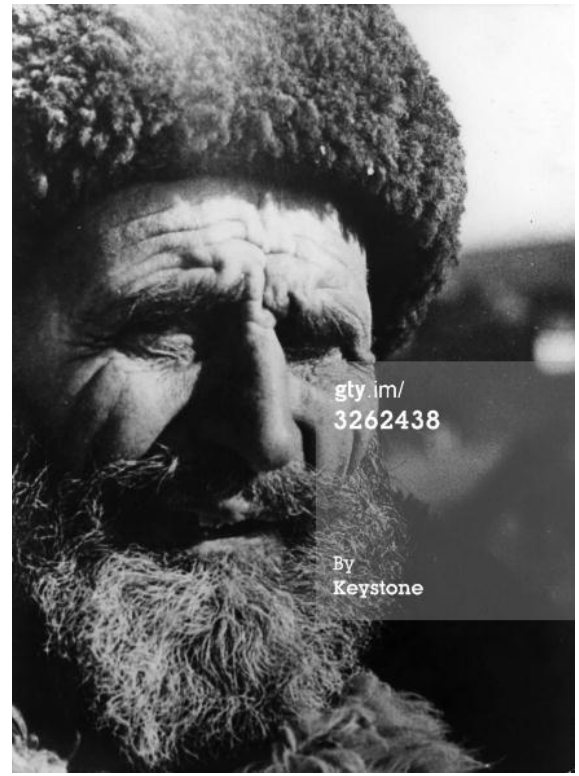


Map of life expectancies in Europe, in 1950, 1965 and 1995, by sex (in years)

Meslé France, Vallin Jacques. Mortality in Europe: the Divergence Between East and West. In: Population, 57e année, n°1, 2002 pp. 157-197.

Although…☺

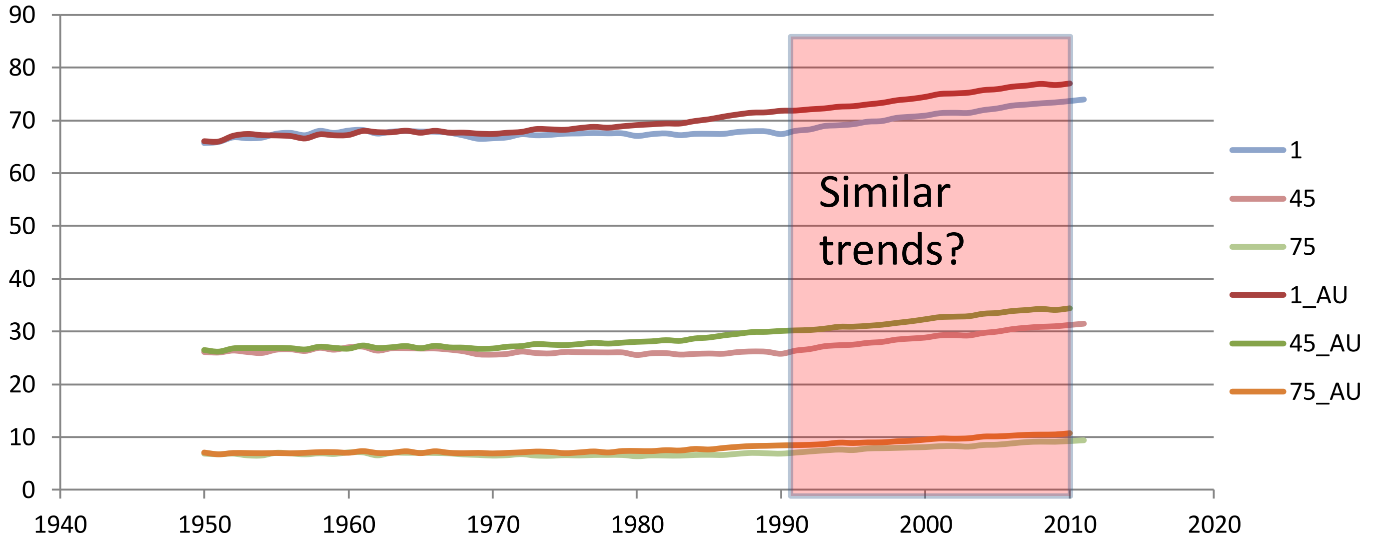
Mahmud Eyvazov commemorated with stamp in 1956 at the **age of 148**.



Shirali Muslimov credits his longevity to hard work. Here he was supposedly **over 160-years-old**. All photos were taken in 1963 or later.

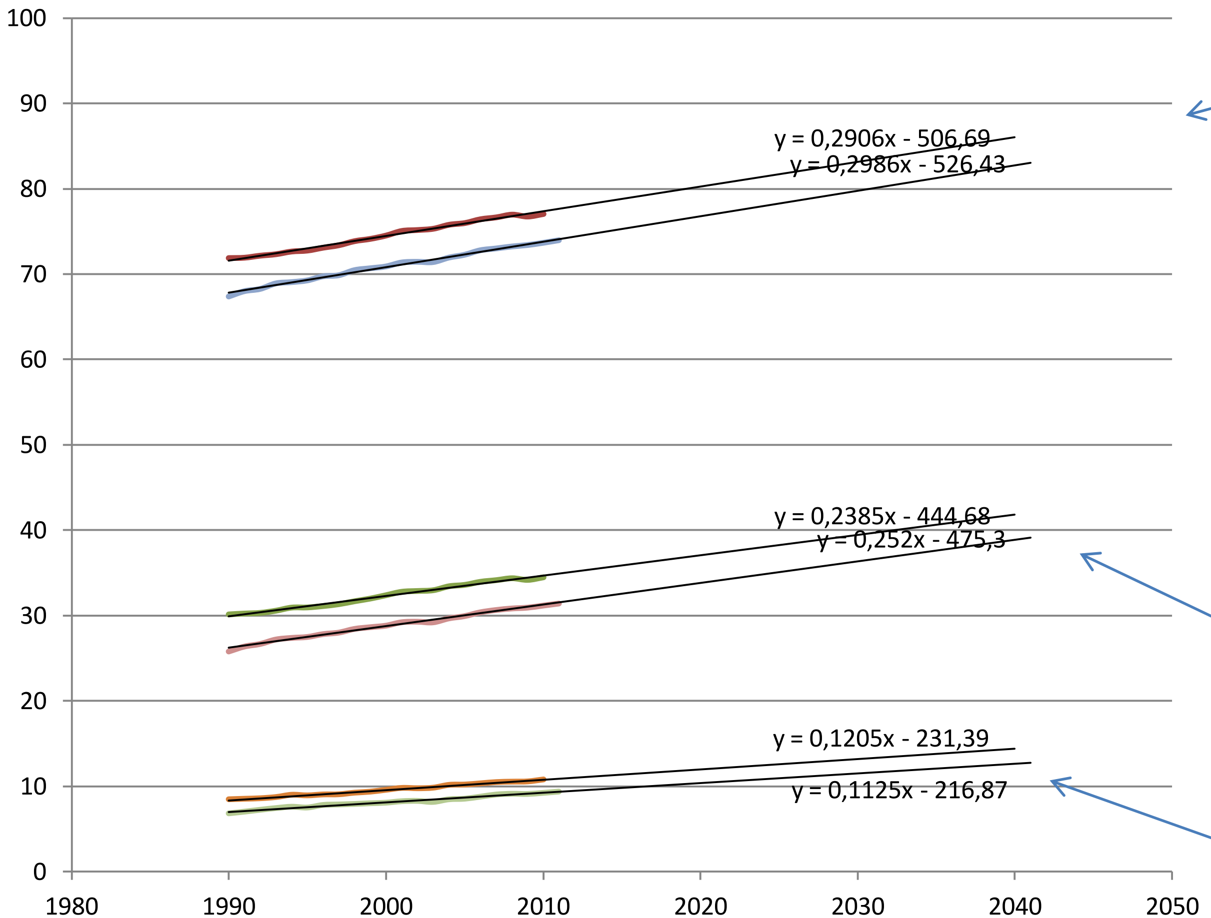
TOOLS 4F

Does CZE catch up???



- Obviously the trend has changed and the expectancy started to grow significantly after 1990
- But is it enough to catch up ?

Does CZE catch up??? - Expectancies



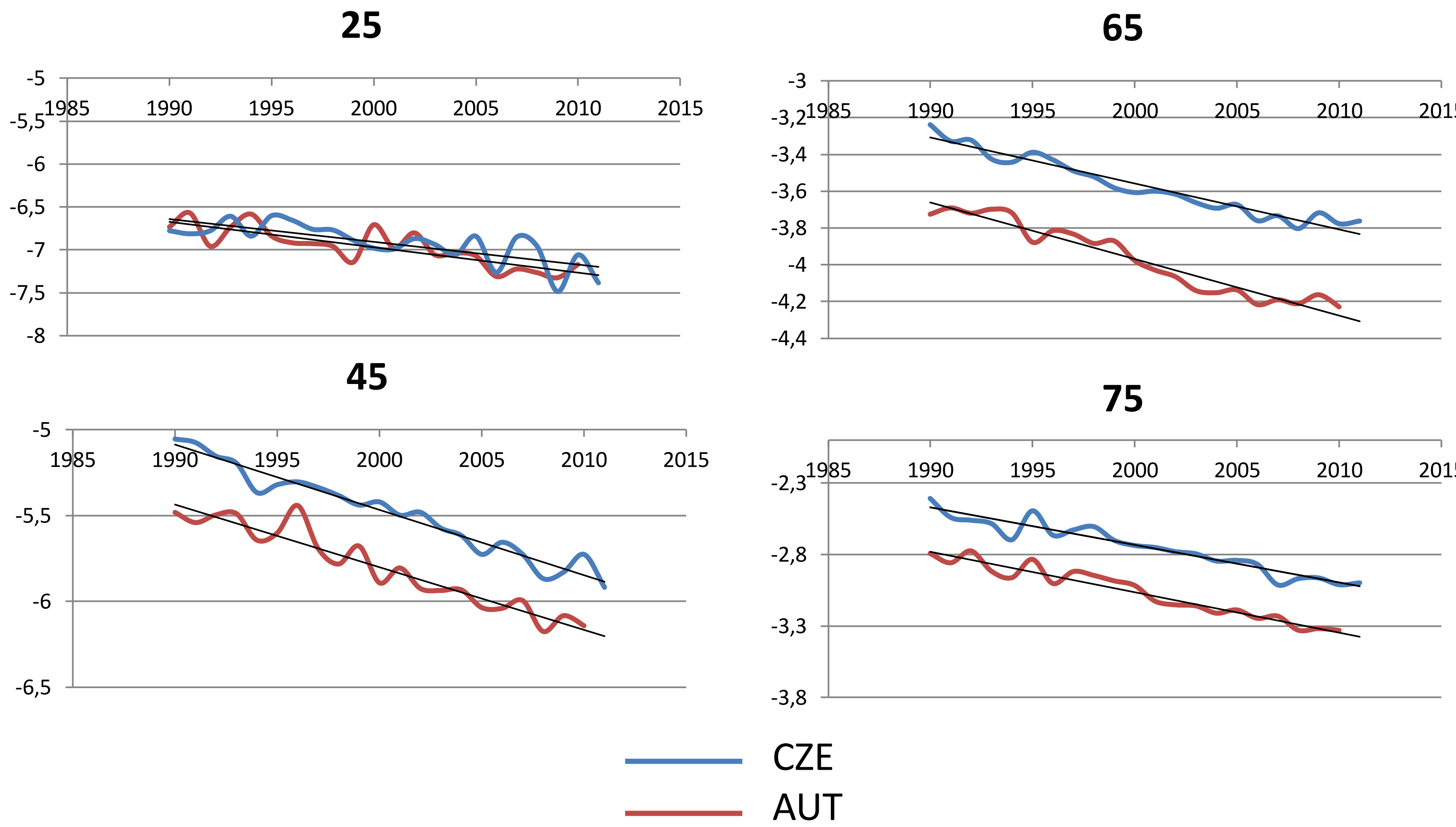
Crosses in:
2467

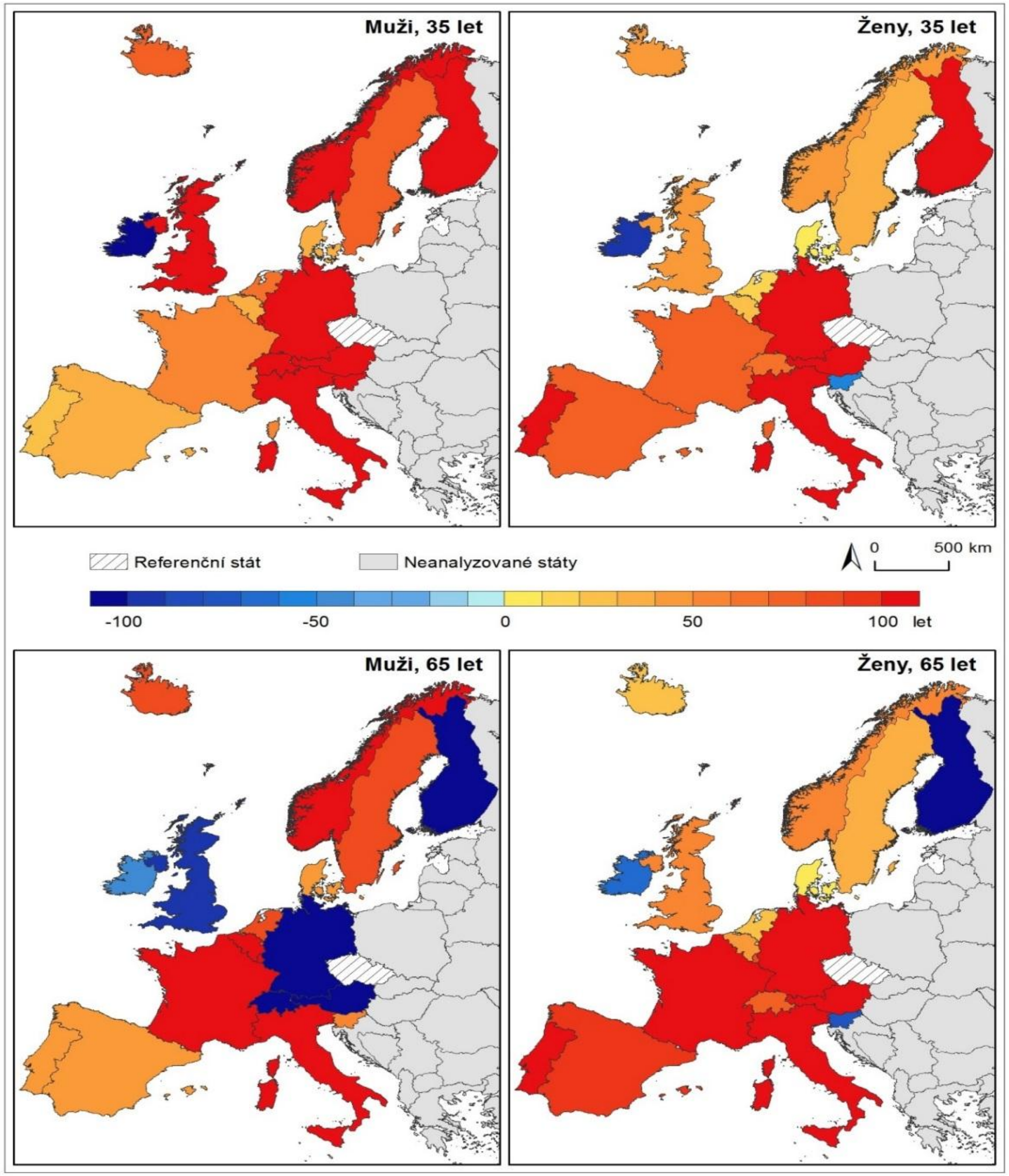
Crosses in:
2253

Crossed in:
1857

TOOLS 4F

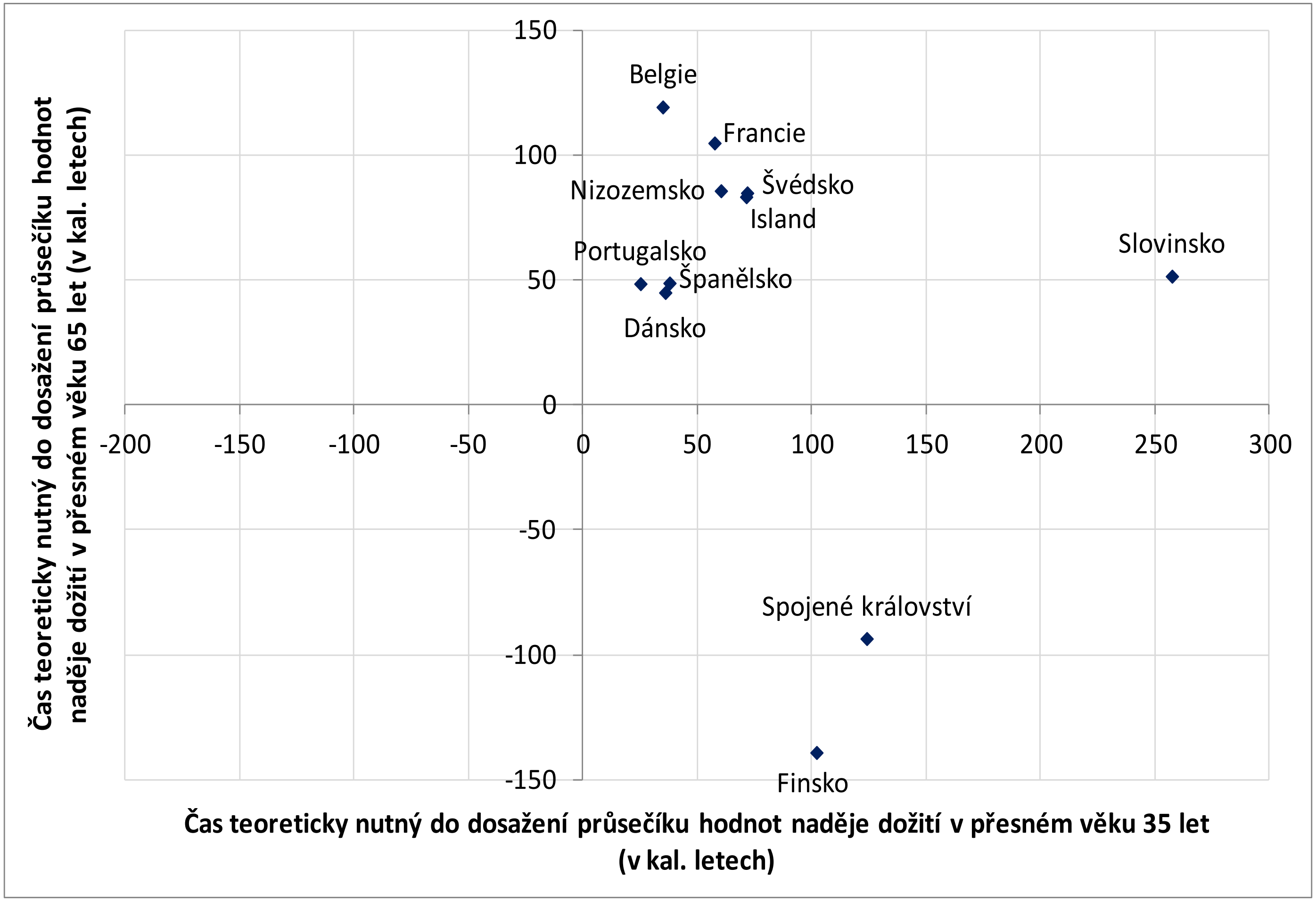
Does CZE catch up??? - Death probabilities Log(q_x)





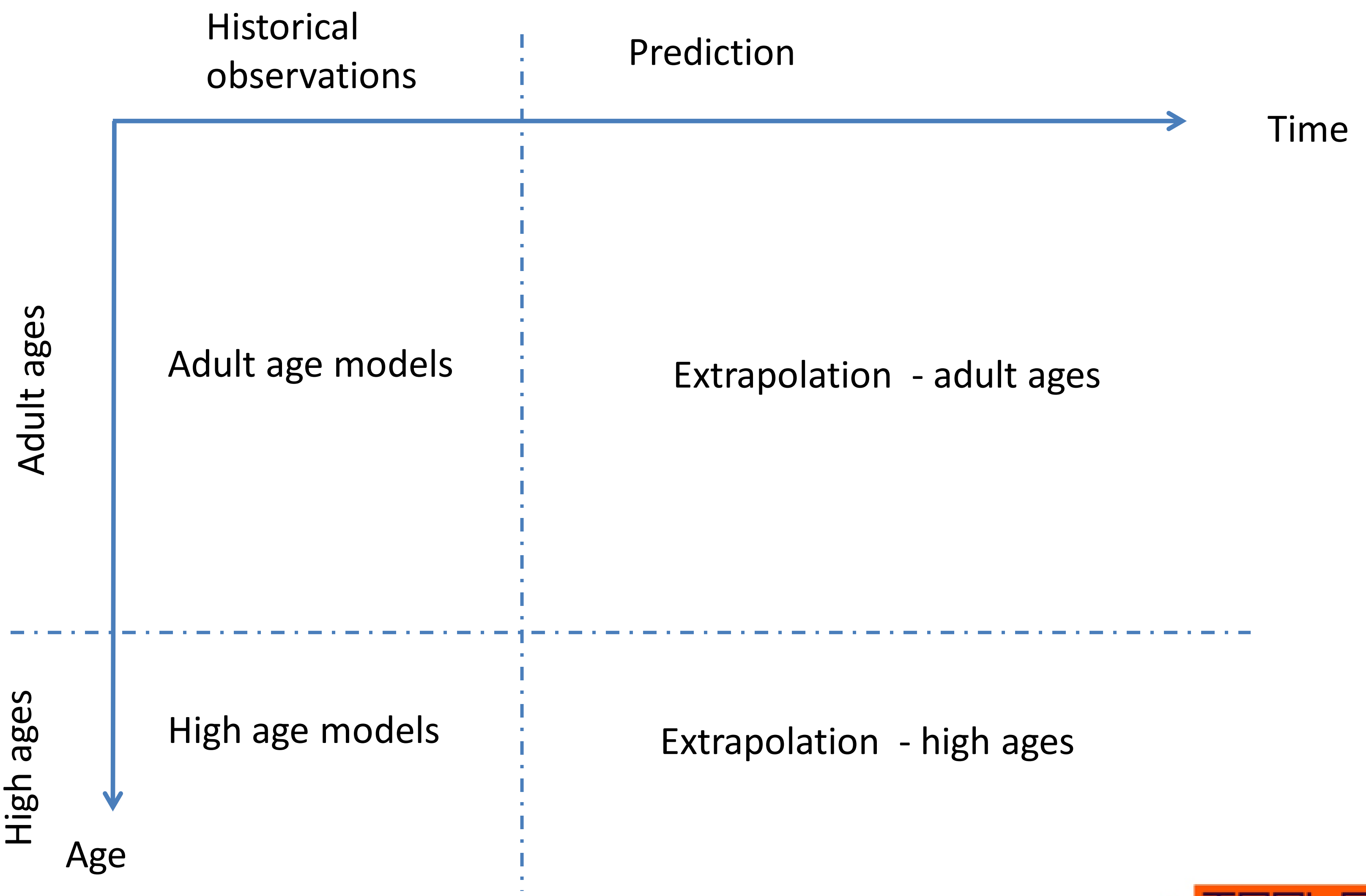
Muži, 35 let
Ženy, 35 let
Referenční stát
Neanalyzované státy
0
500 km
-100
-50
0
50
100
let
Muži, 65 let
Ženy, 65 let

Does CZE catch up???

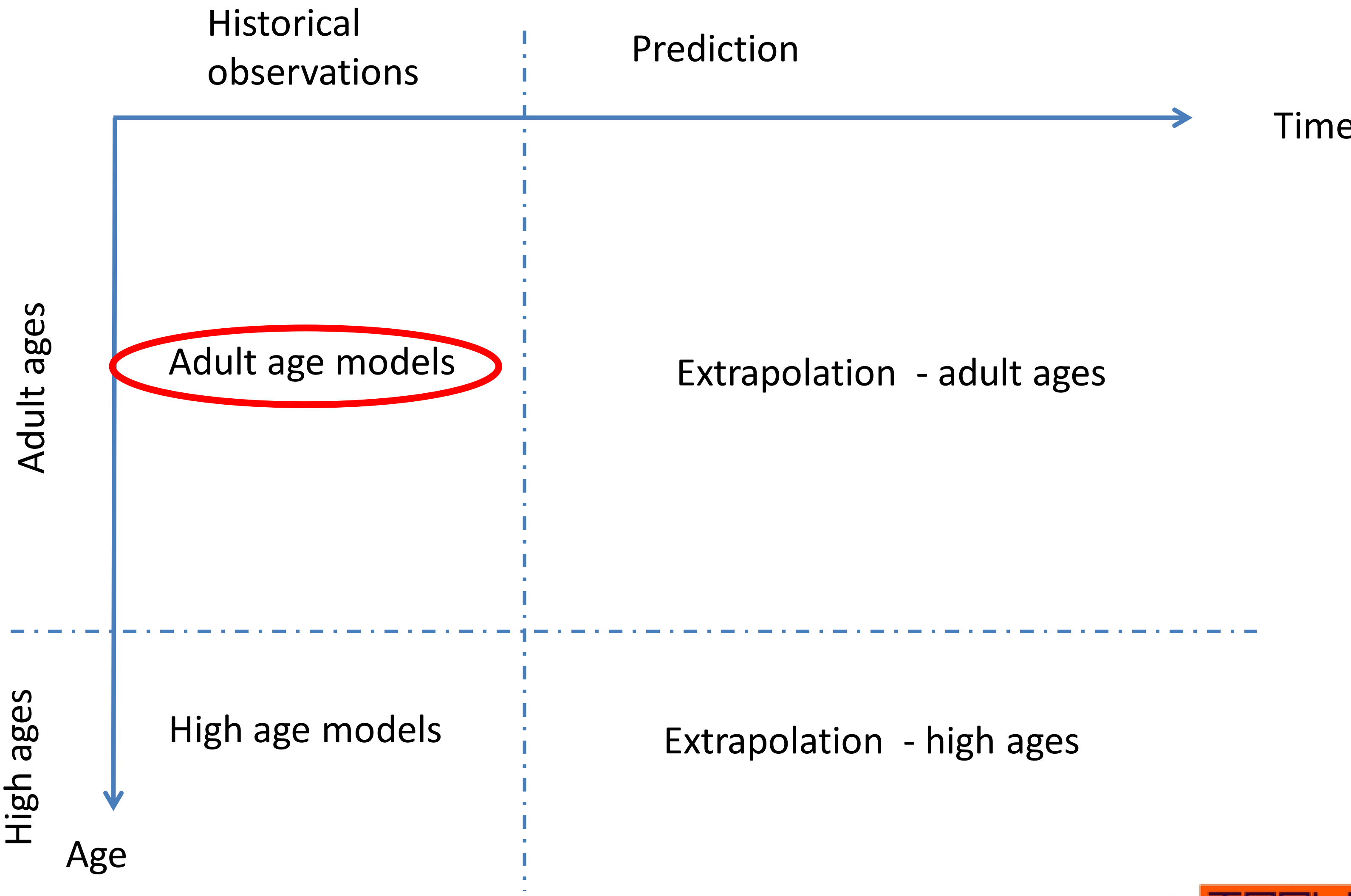


Mortality models

Phases of the modeling process



Phases of the modeling process



Basic Models

- Dynamics of q_{xt} *or* μ_{xt} *is modelled.*
- Reduction factor

$$q_{xt} = q_{x0} R(t) \qquad\qquad q_{xt} = q_{x0} R(x,t)$$

e.g. for linear trend in log(q_{xt})

$$R(x,t) = \exp(-\delta_x t)$$

- Or… $$R_x(t-t') = \alpha_x + (1-\alpha_x)(1-f_x)^{\frac{t-t'}{20}}$$

$$f_x = \begin{cases} c & \text{if } x < 60 \\ 1 + (1-c)\dfrac{x-110}{50} & \text{if } 60 \le x \le 110 \\ 1 & \text{if } x > 110 \end{cases} \qquad \alpha_x = \begin{cases} h & \text{if } x < 60 \\ \dfrac{(110-x)\,h + (x-60)\,k}{50} & \text{if } 60 \le x \le 110 \\ k & \text{if } x > 110 \end{cases}$$

Parametric Models

- Parametric models ("Mortality laws")
- A function ("law") is assumed to describe the dependence of mortality on age.

$$\mu_x = f(x;\Theta)$$

- The function is fitted in each year and time series of parameters are extrapolated to the future.

$$\hat{\mu}_{xt} = f(x;\hat{\Theta}_t)$$

TOOLS 4F

Cairns – Blake – Dowd

- Cairns – Blake – Dowd (CBD)
- Specification of the logistic regression with time dependent parameters

$$\log\left(\frac{q_{xt}}{1-q_{xt}}\right) = \alpha_t + \beta_t x$$

$$q_{xt} = \frac{\exp(\alpha_t + \beta_t x)}{1+\exp(\alpha_t + \beta_t x)}$$

Goal Tables

- Sometimes relevant data are lacking…
- …and there exist reliable forecast “next door”.
- It may be useful to avoid extrapolating local trend and
- Instead grow the local mortality to the goal table.

$$q(x; j+t) = q(x; j) \times \prod_{i=1}^{t} f(x; j) \times e^{i\alpha(x)}$$

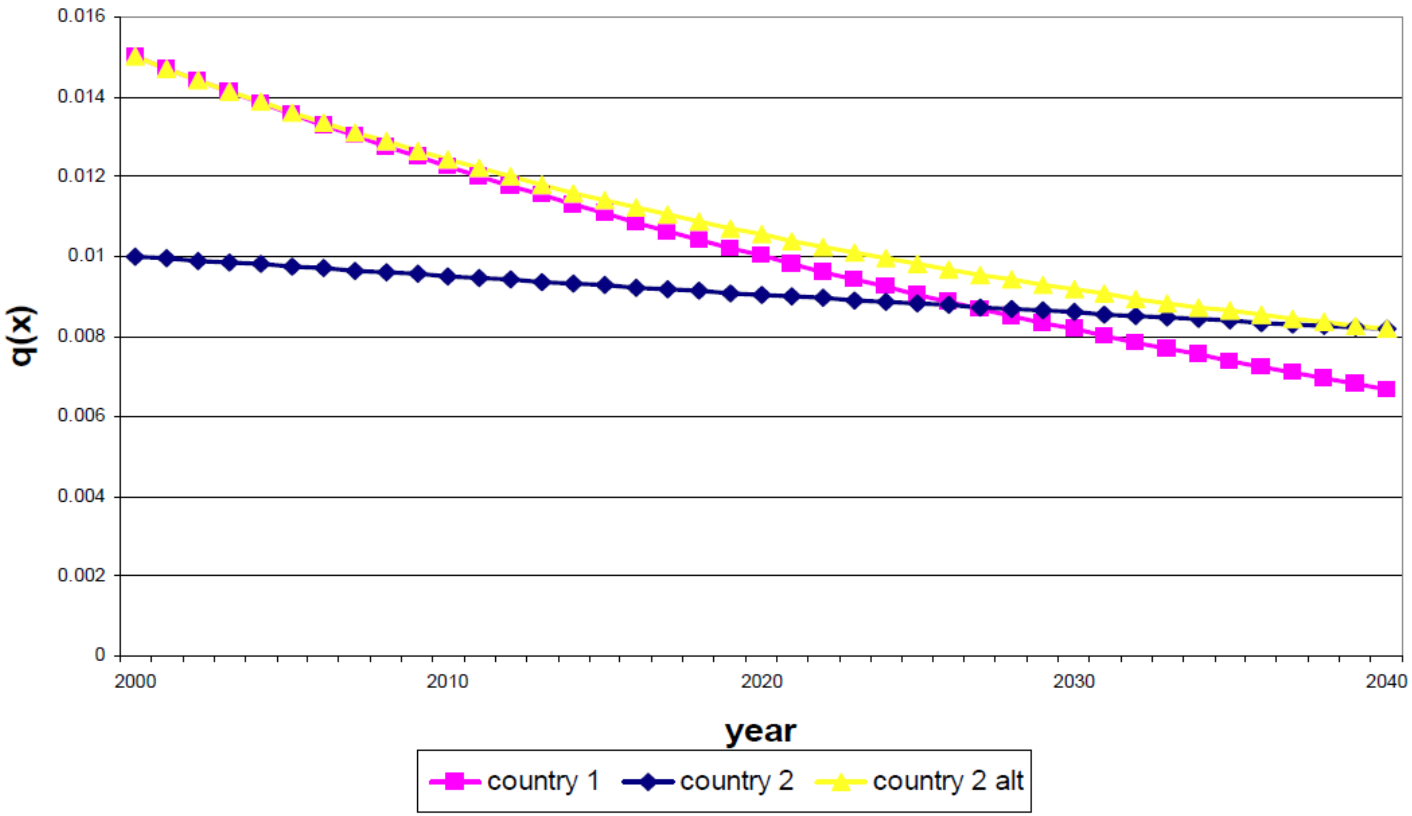
And:

$$q(x; j+t) = q(x; j) \times f(x; j)^t \times e^{\frac{\alpha(x)t(t+1)}{2}}$$

[Van Broekhoven 2010]

Goal Tables

- At the start following the local trend
- At the end reaching "the goal"



[Van Broekhoven 2012]

Lee-Carter Model

- log-bilinear model defined as follows:

$$\ln(m_{xt}) = \alpha_x + \beta_x \kappa_t + \varepsilon_t$$

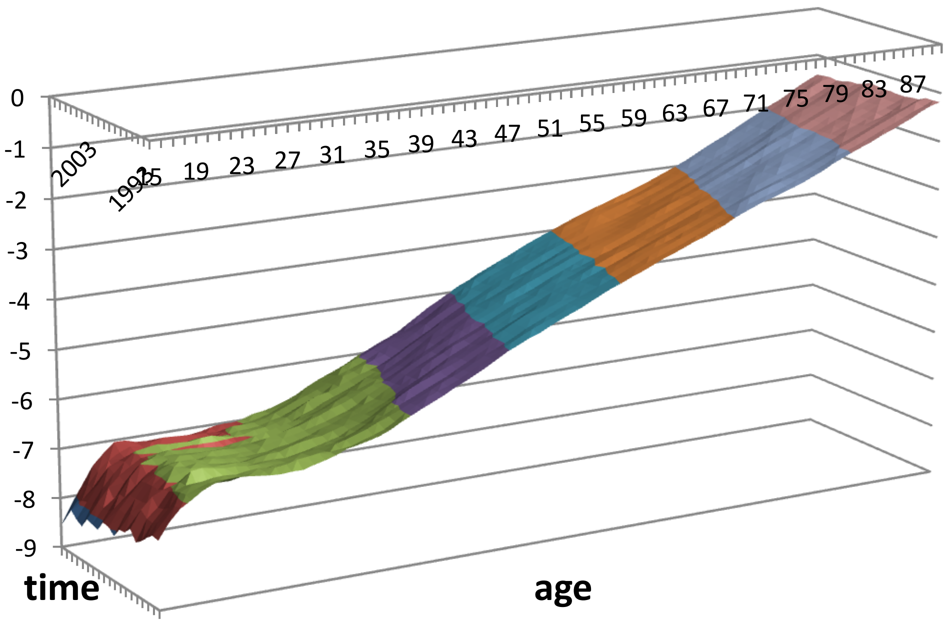
  – m_{xt} is the specific death rate at age *x* and year *t*
  – α_x defines the shape of the **age profile** of mortality averaged over time
  – β_x represents the pattern of **deviations from the age profile** of mortality
  – κ_t describes the variation in the **general level** of mortality

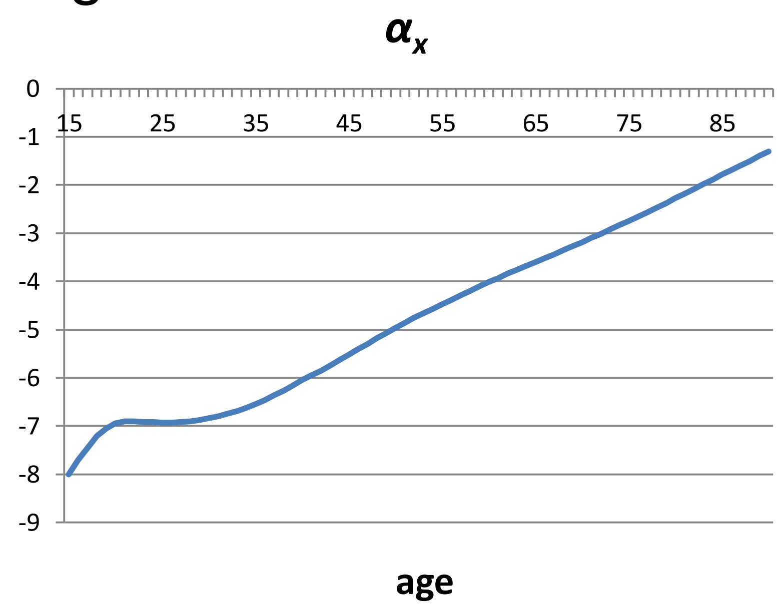
TOOLS 4F

Parameters of the LC model

$$\ln(m_{xt}) = \alpha_x + \beta_x \kappa_t + \varepsilon_{xt}$$

- α_x - **age profile** of mortality averaged over time

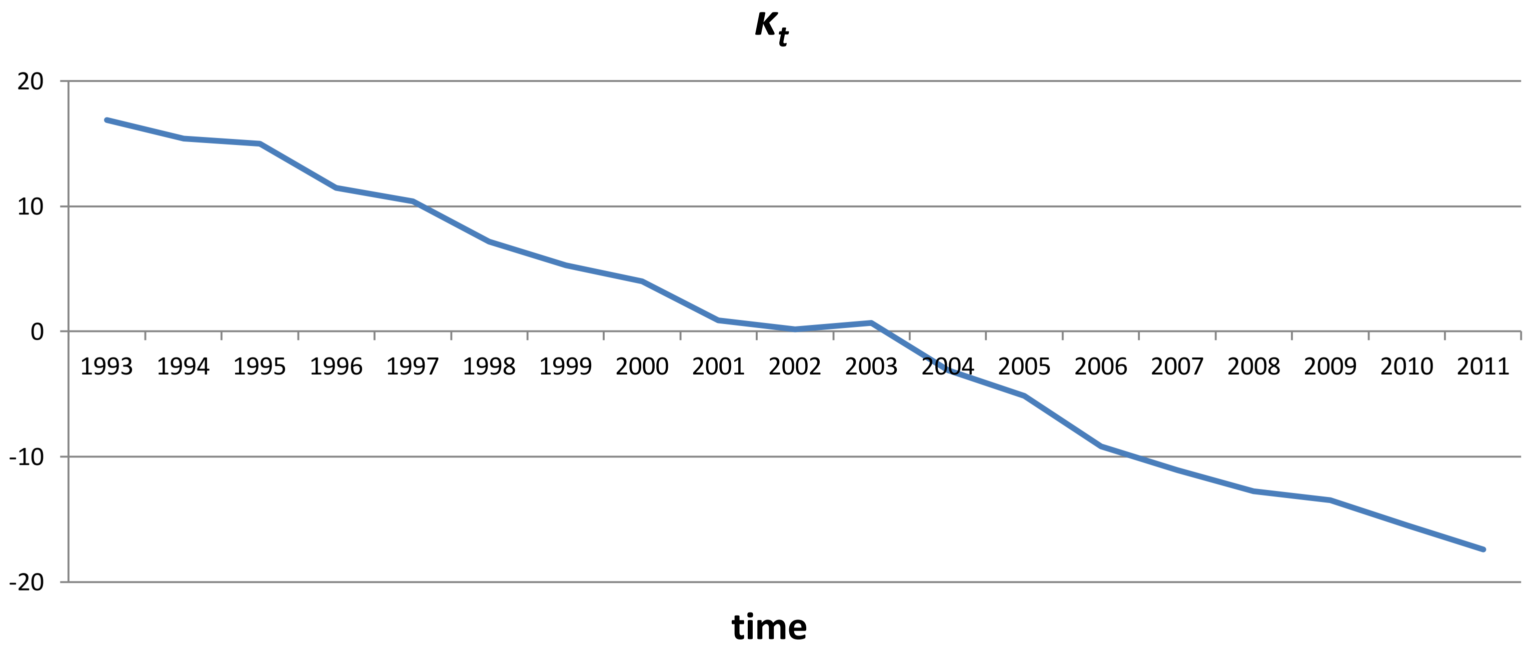




Parameters of the LC model

$$\ln(m_{xt}) = \alpha_x + \beta_x \kappa_t + \varepsilon_{xt}$$

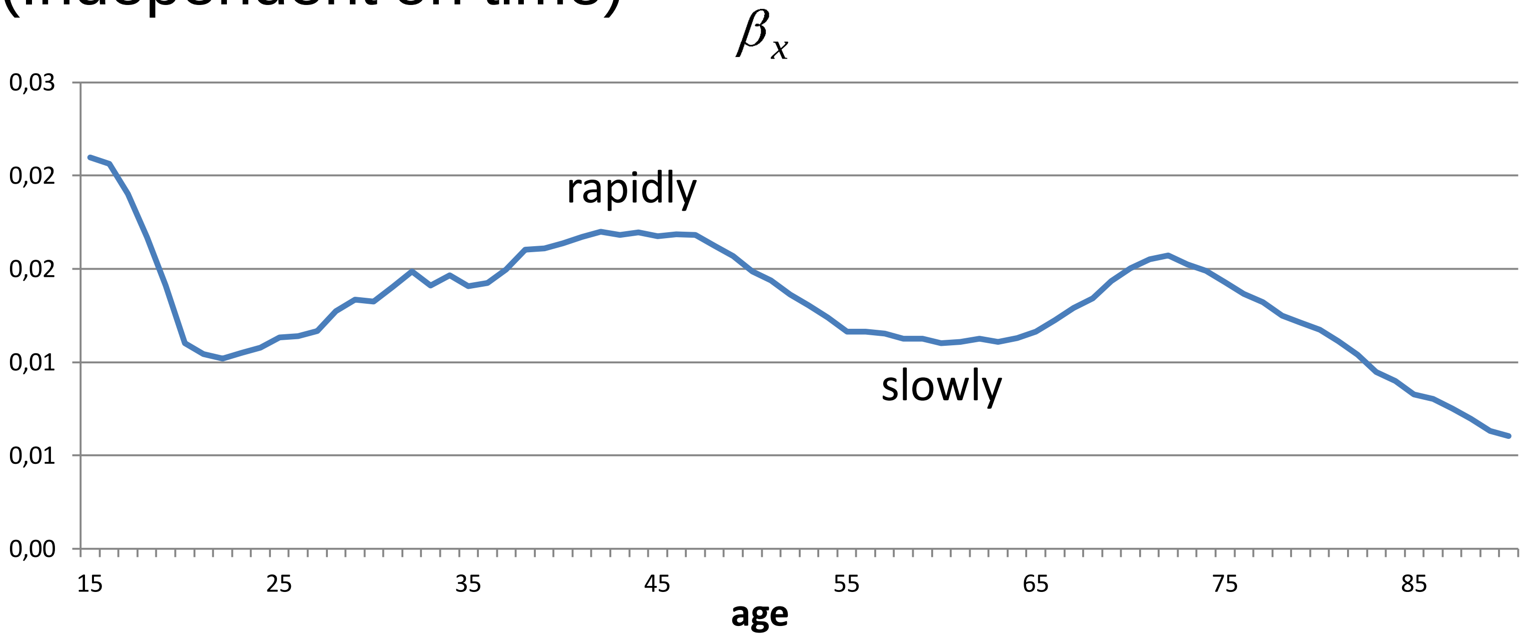
- κ_t - **general level** of mortality (independent on age)



Parameters of the LC model

$$\ln(m_{xt}) = \alpha_x + \beta_x \kappa_t + \varepsilon_{xt}$$

- β_x - how rapidly or slowly mortality at each age varies when the general level of mortality (κ_t) changes (Independent on time)



Lee-Carter Model

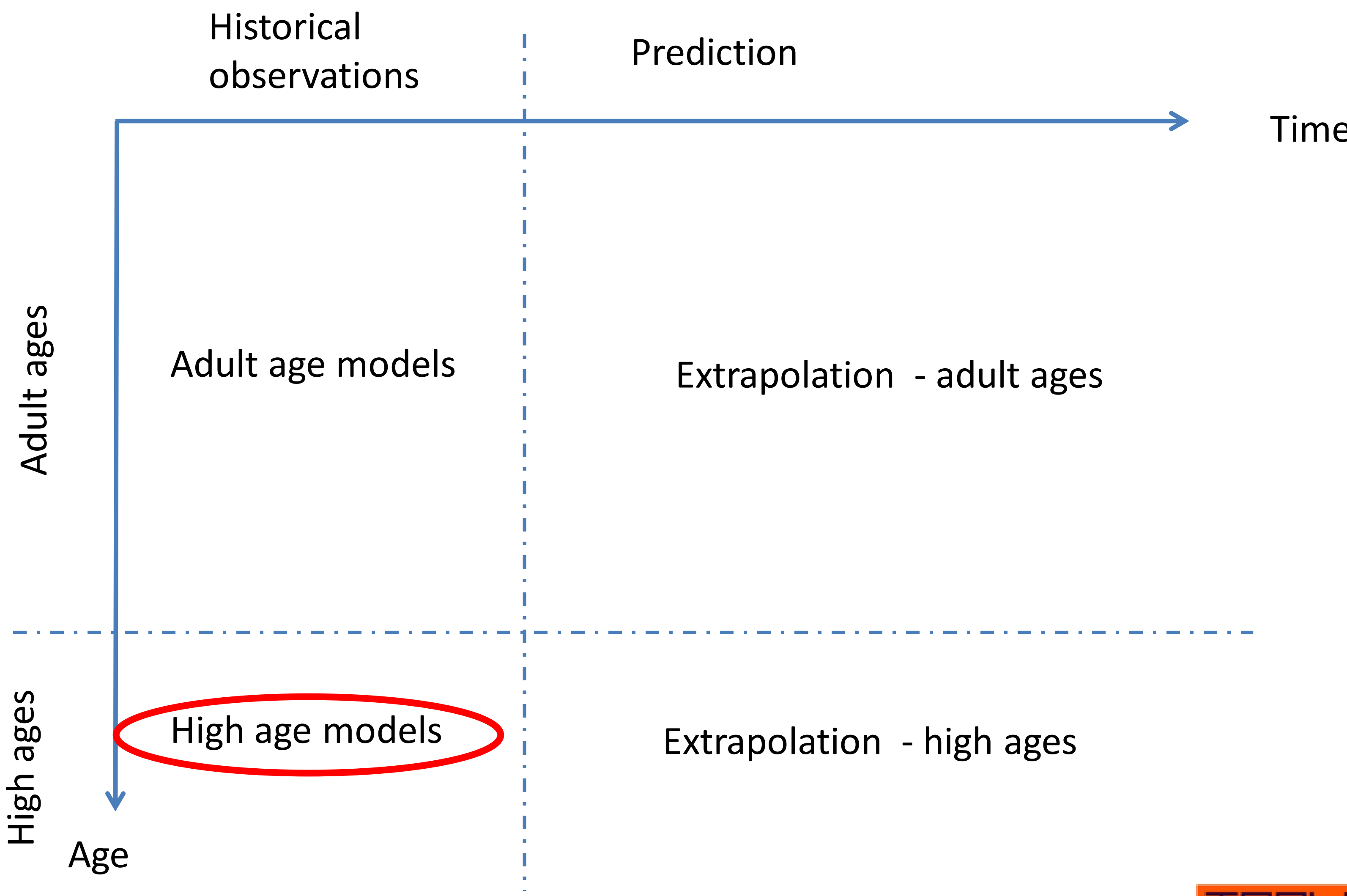
- LC model is identified by the constraints

$$\sum_t \kappa_t = 0 \quad \sum_x \beta_x = 1$$

- Further extensions possible (cohort effect)

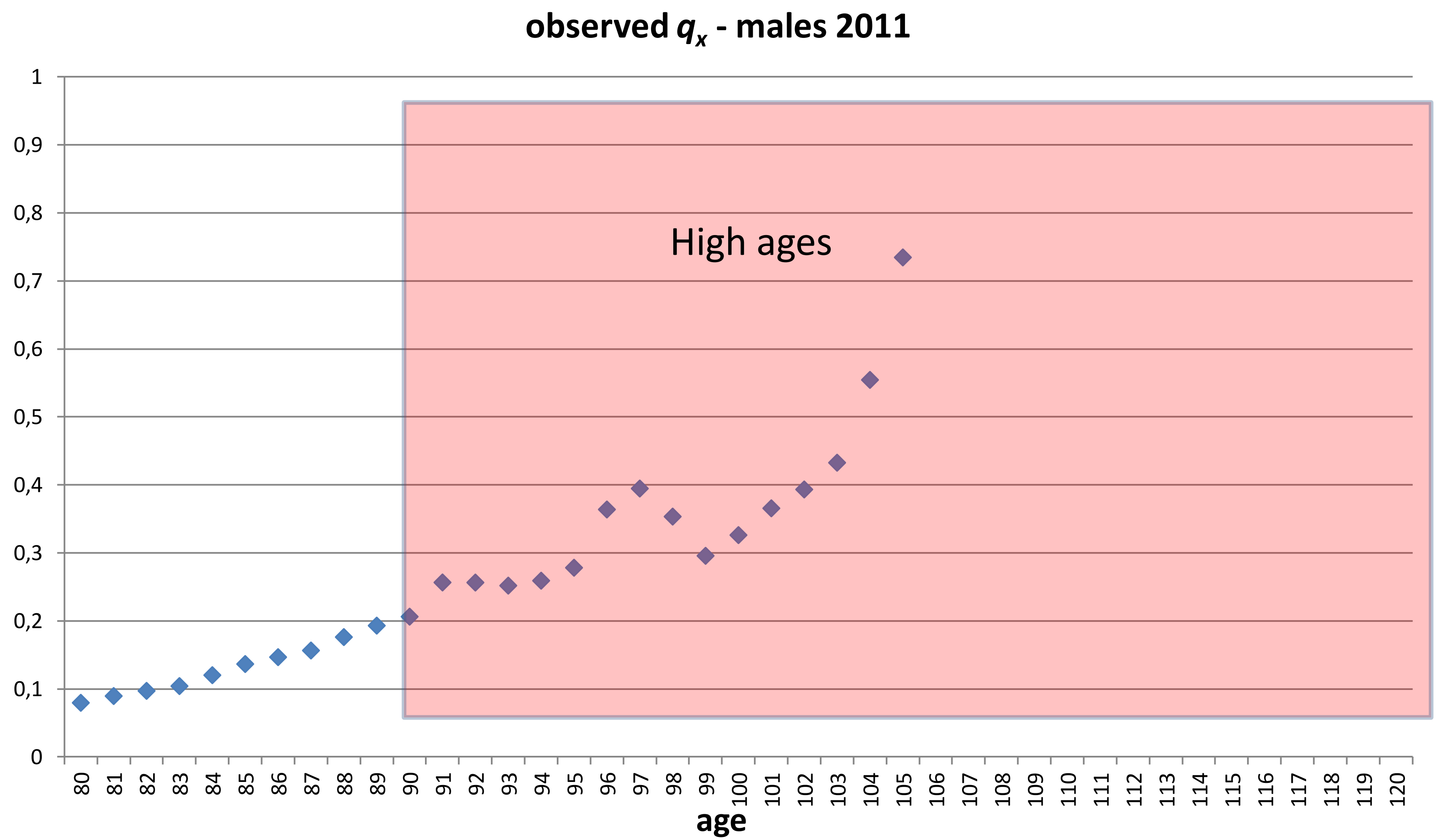
$$\ln(m_{xt}) = \alpha_x + \beta_x \kappa_t + \gamma_x \tau_{t-x}$$

Phases of the modeling process

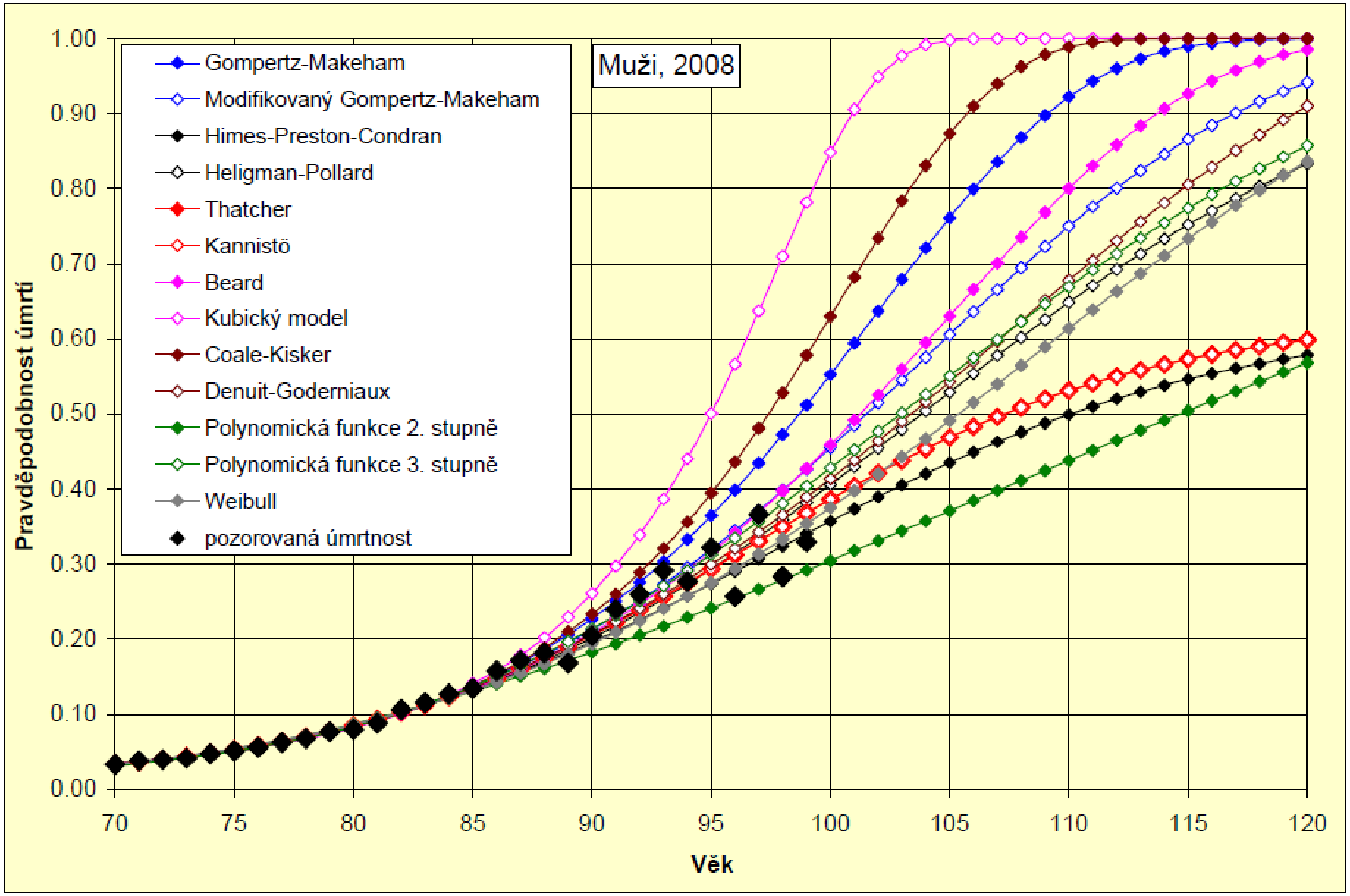


TOOLS 4F

Modeling the high ages



Modeling the high ages



BURCÍN, Boris; TESÁRKOVÁ, Klára; ŠÍDLO, Luděk. Nejpoužívanější metody vyrovnávání a extrapolace křivky úmrtnosti a jejich aplikace na českou populaci). Revue pro výzkum populačního vývoje, 52:77-89, 2010.

Modeling the high ages

- Exponencial models
  - Gompertz-Makeham (Koschin)

$$\mu_x = a + b \cdot c^x \qquad \mu_x = a + b \cdot c^{x_0 + \frac{\ln(d \cdot (x - x_0) + 1)}{d}}$$

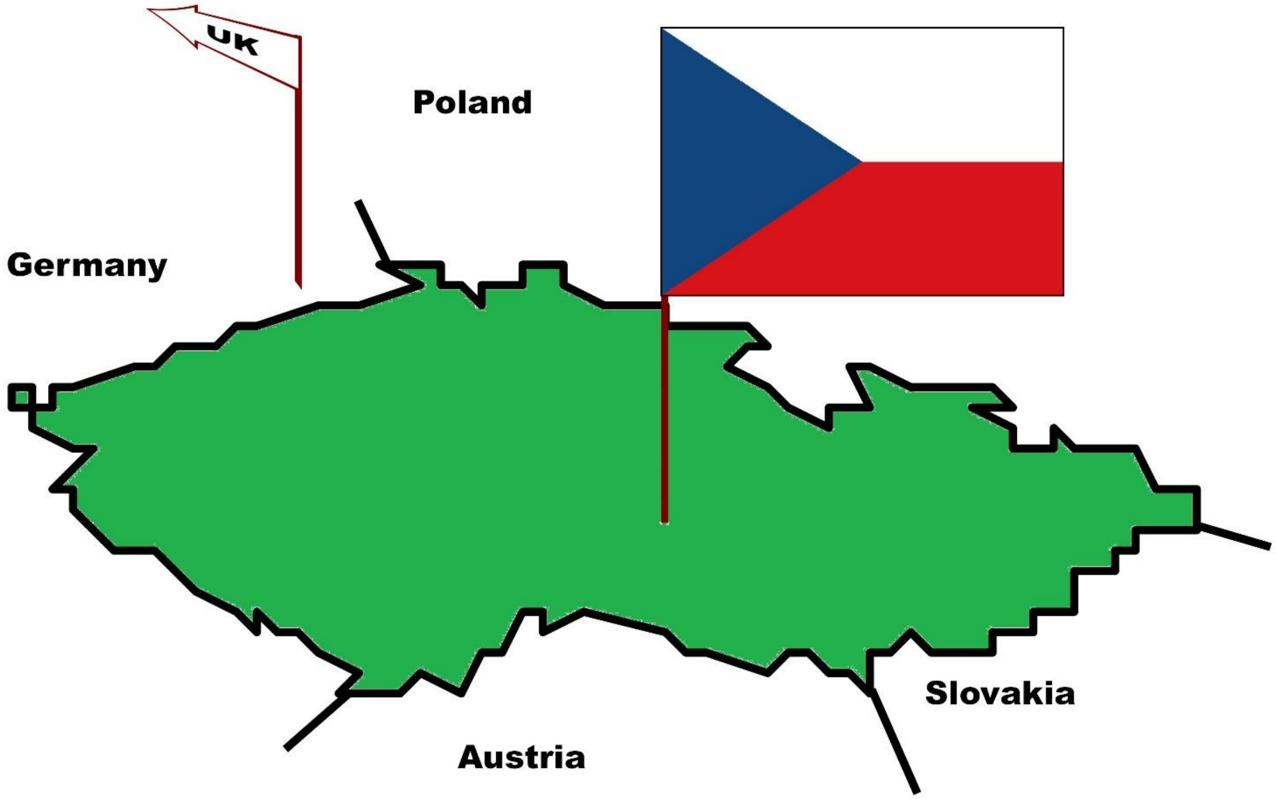
- Logistic models
  - Kannistö

$$\ln\left(\frac{m_x}{1 - m_x}\right) = \theta_0 + \theta_1 (x - x_0)$$

- Other models
  - Coale-Kisker

$$m_x = \exp(a \cdot x^2 + b \cdot x + c)$$

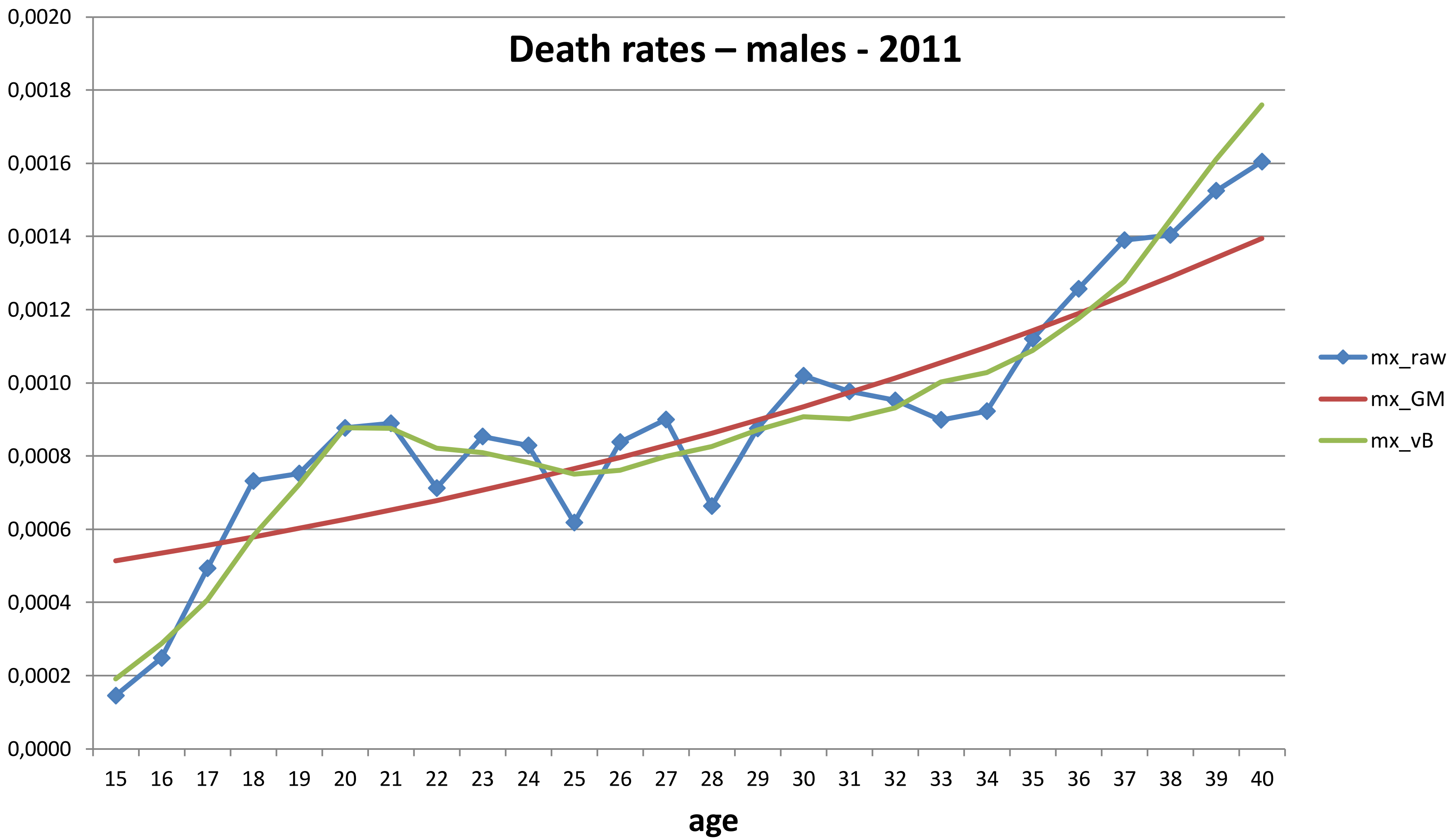
Application to the Czech mortality data



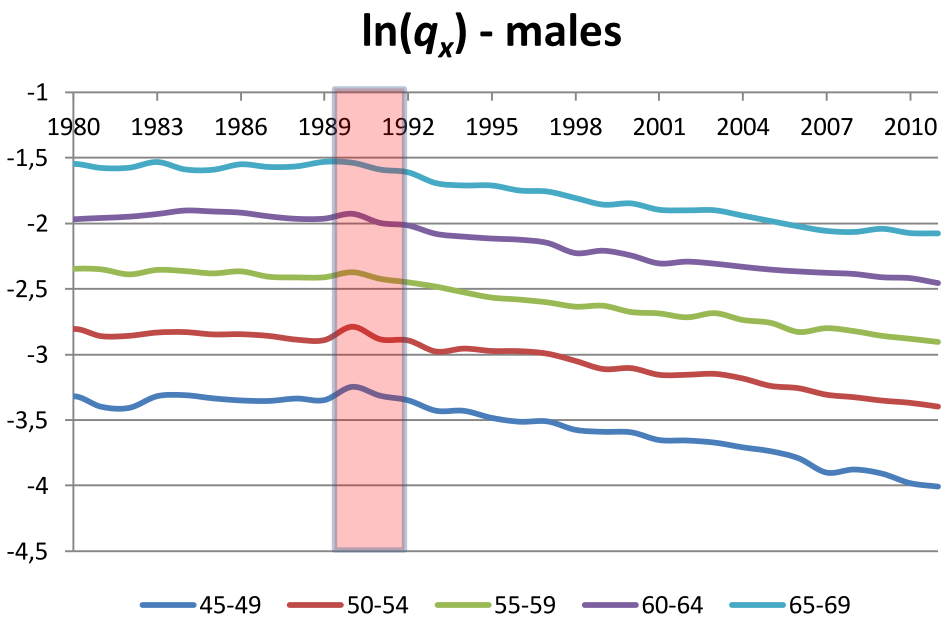
Application to the Czech mortality data

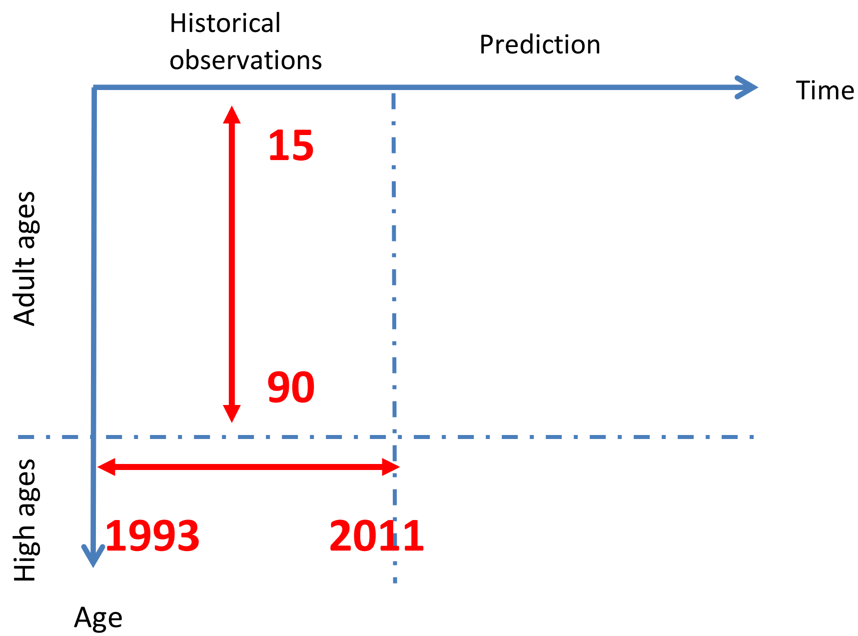
- Data sources
  - Public data from the Human Mortality Database (HMD) was used
  - HMD is the project of American and German researchers
  - www.mortality.org
- Data smoothing
  - Gompertz-Makeham method
  - Adaptive techniques
    - Moving averages
    - Van Broekhoven algorithm

Data smoothing

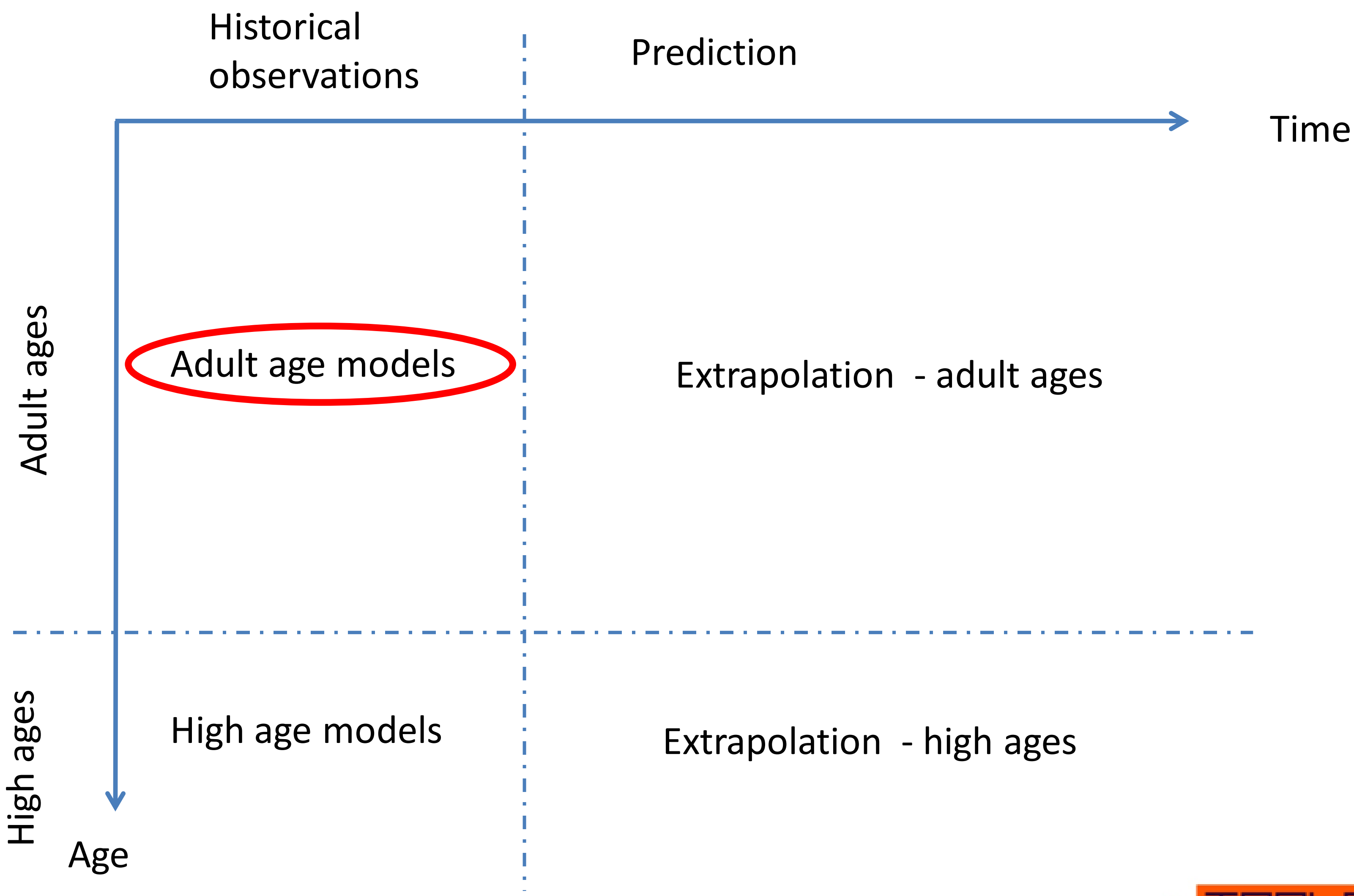


Choosing the age-time period





Phases of the modeling process



Lee-Carter Model

$$\ln(m_{xt}) = \alpha_x + \beta_x \kappa_t + \varepsilon_{xt}$$

- Estimation of the LC model
  1) Ordinary least squares
  2) Weighted least squares
  3) Maximum likelihood estimation

Estimation of the LC model

1) Ordinary least squares

$$\sum_{x,t}(\ln(\hat{m}_{xt})-\alpha_x-\beta_x\kappa_t)^2 \rightarrow \min$$

$$\hat{m}_{xt}=\frac{d_{xt}}{E_{xt}}$$

where d_{xt} is the number of deaths and E_{xt} is the exposure to risk

- Singular value decomposition method can be used to find a least squares solution
- Second stage estimation of κ_t is recommended to better fit the model and the observed deaths (d_{xt})

Estimation of the LC model

2) Weighted least squares

$$\sum_{x,t} w_{xt} (\ln(\hat{m}_{xt}) - \alpha_x - \beta_x \kappa_t)^2 \rightarrow \min$$

$w_{xt} = d_{xt}$ This ensures that predicted death rates will be close to observed values with highest number of deaths

No need to make the second-stage estimation of κ_t

Under the Least squares method the errors are assumed to be homoskedastic (but $\ln(m_x)$ is much more volatile at higher ages). This assumption is often not realistic, thus it is suggested to use Maximum likelihood estimation

Estimation of the LC model

3) Maximum likelihood estimation (MLE)

MLE on the Poisson number of deaths allowing heteroskedasticity

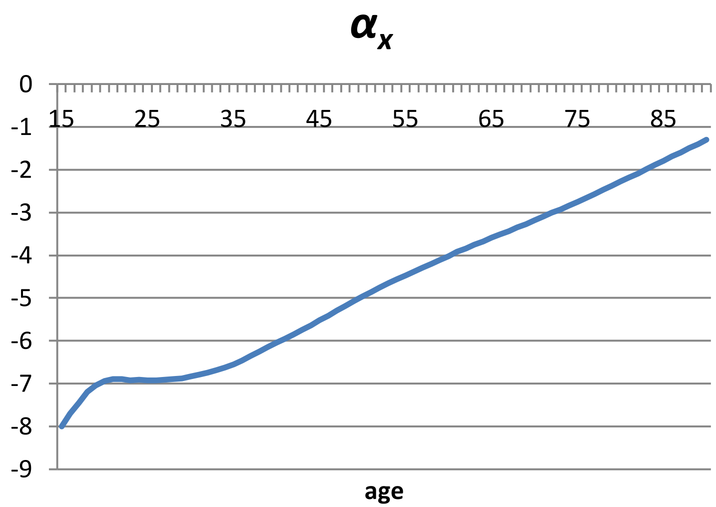
Poisson number of deaths

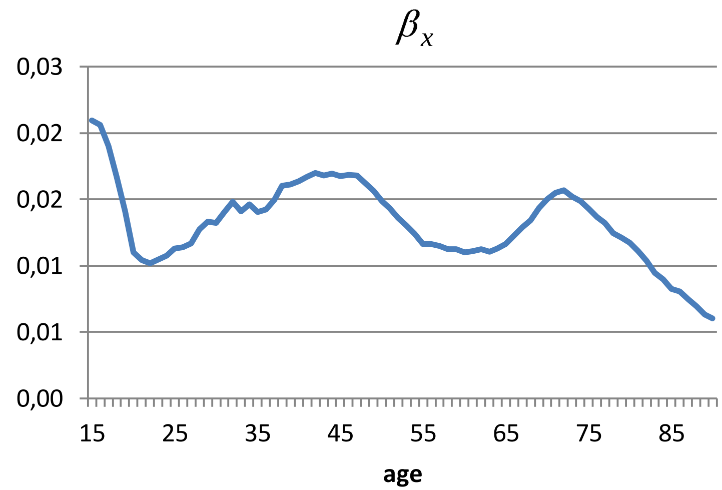
$$D_{xt} \sim \text{Poisson}(\lambda_{xt} = E_{xt} m_{xt} = E_{xt} \exp(\alpha_x + \beta_x \kappa_t))$$

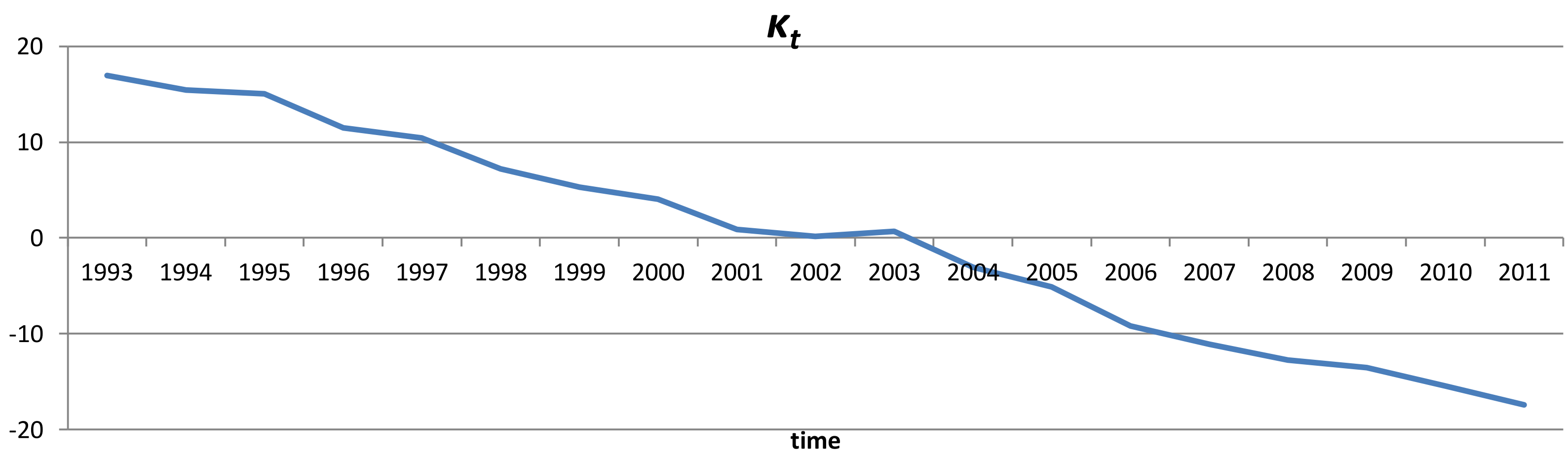
Parameters of the LC model are estimated by maximizing the likelihood function (Newton iterative method is used)

$$\prod_{x,t} \frac{(\lambda_{xt})^{d_{xt}}}{d_{xt}!} \exp(-\lambda_{xt})$$

Estimation of the LC model

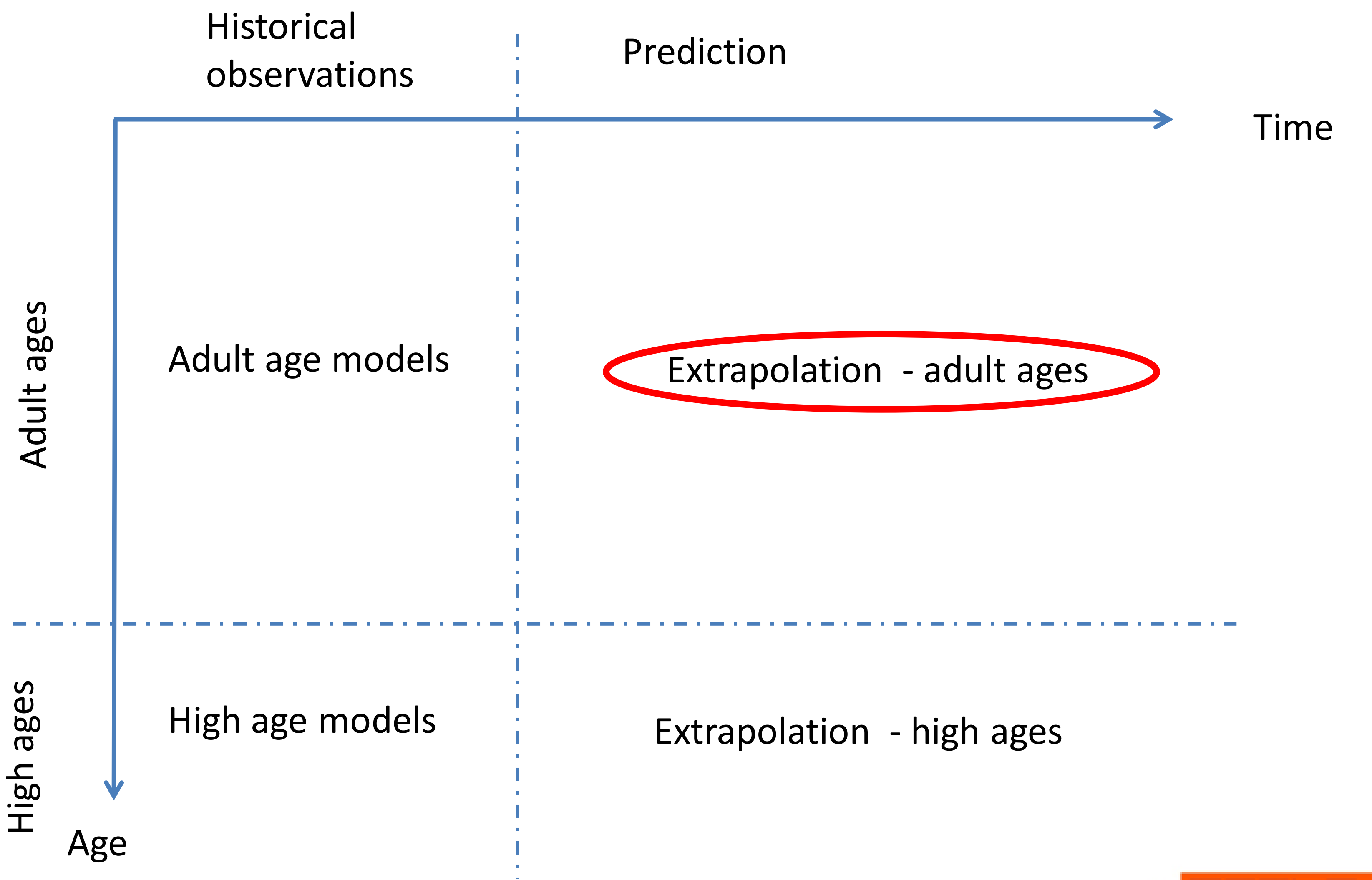






TOOLS 4F

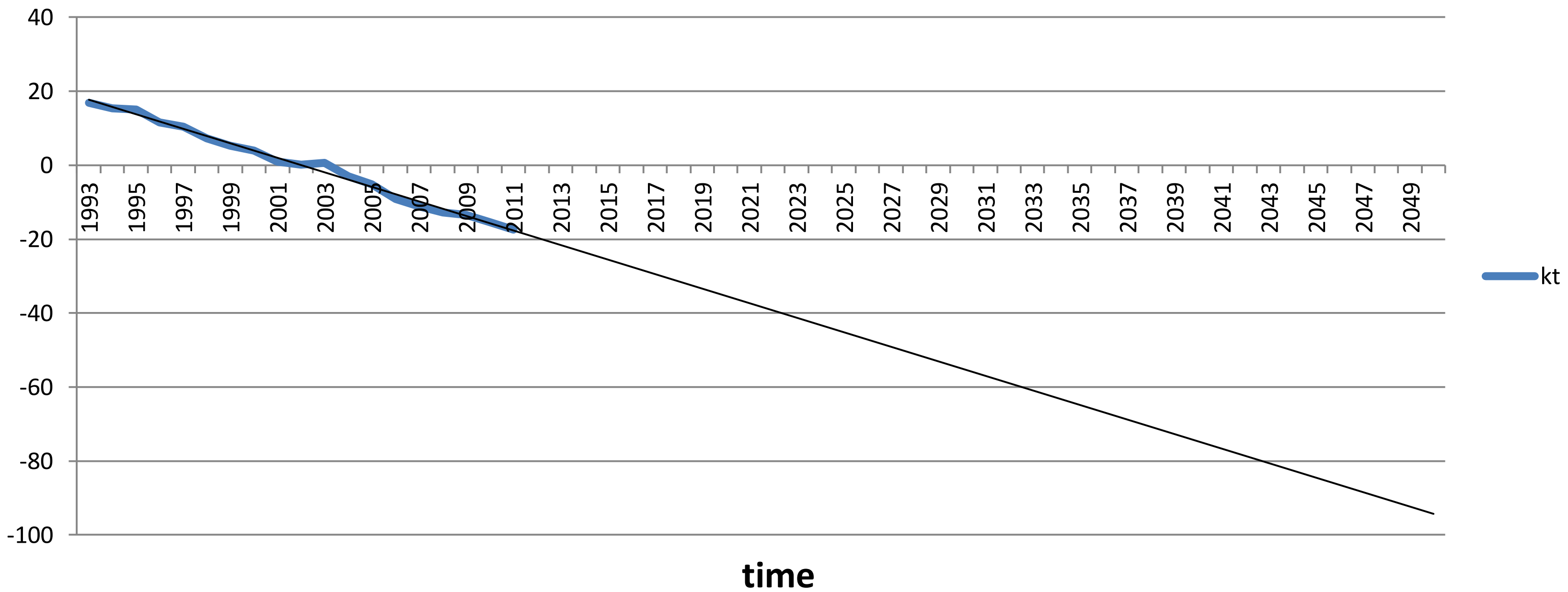
Phases of the modeling process



TOOLS 4F

Prediction of the LC model

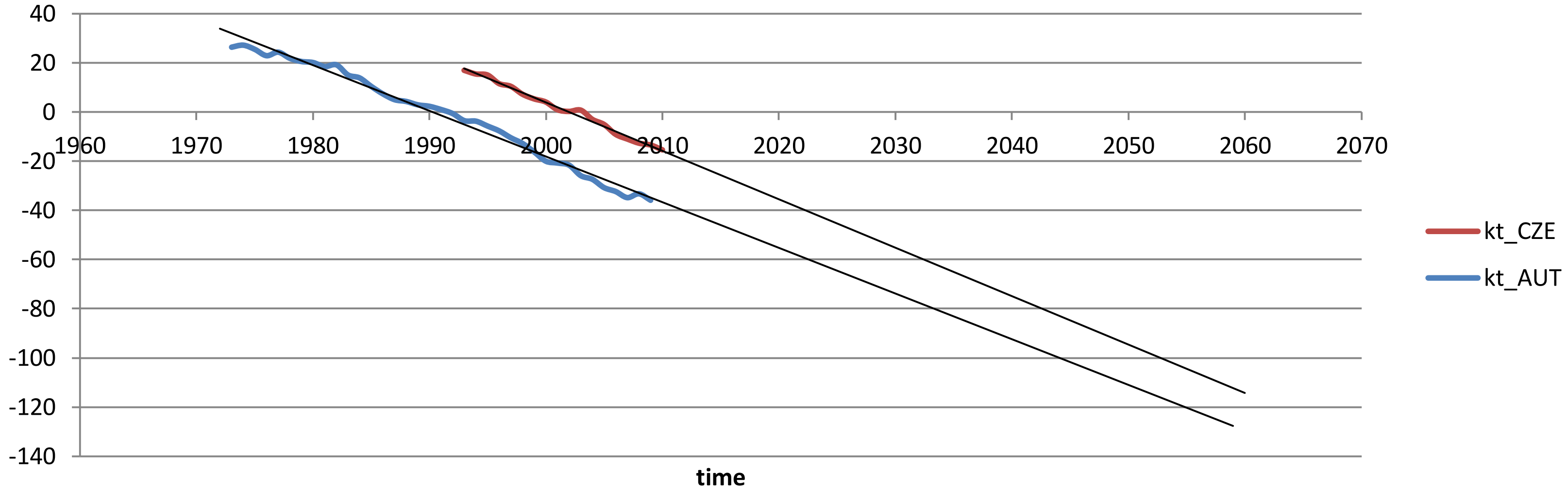
- Prediction of the κ_t using a random walk with drift: $$\kappa_t = \kappa_{t-1} + \Delta + \varepsilon_{xt}$$
- But linear trend will not last forever…



Local trend?

- As we only consider short history, there is always a danger that a ‘local trend’ is extrapolated for a long period
- It is necessary to compare the short term trend with surrounding countries which did not experience the trend change in 1990 and hence are forecasting their trend based on longer history

Mid term trend AUT vs short term trend CZE



- We can conclude that the CZE short term trend is similar to mid term AUT trend
- And hence we can assume that the CZE short term trend would be similar to CZE mid term trend in the case there was no socialism

“Bio-demographic” limit

- Mortality has been decreasing for a long time
- But it can not decrease to zero (or under)
- There is always minimum level of mortality – “bio-demographic limit”
- The trend will slow down when approaching the limit

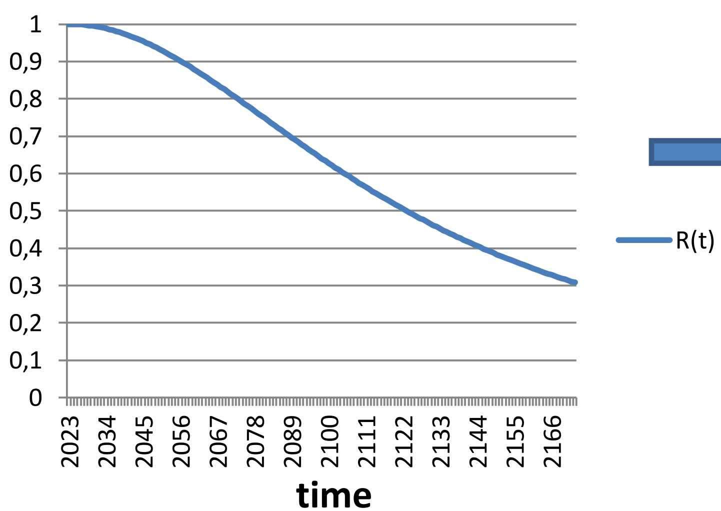
Trend Reduction

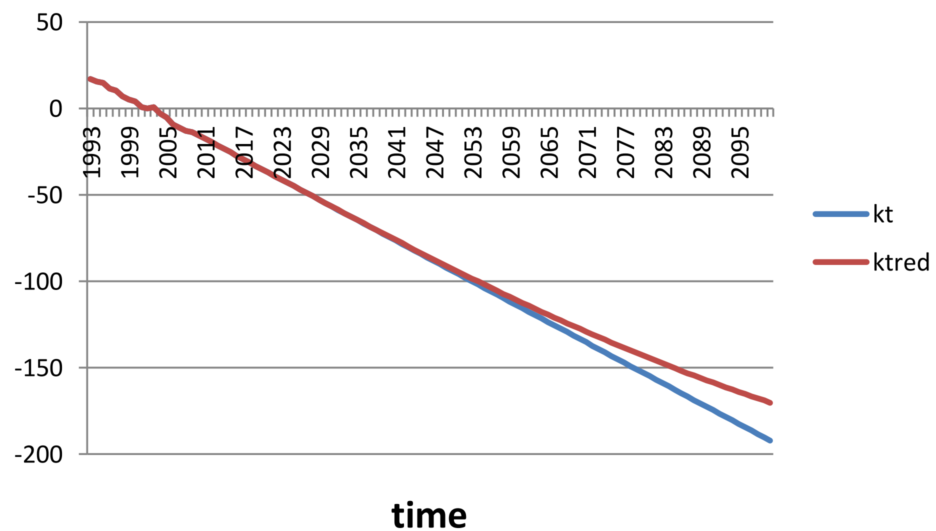
Reduction factor *R(t)* non-linearly reduces the difference Δ to zero as time tends to infinity

$$R(t) = \frac{1}{1 + \frac{t - t_0}{t_{1/2}}}$$

$t_{1/2}$ = 100 years (half-time)

t_0 = year from which the reduction starts





Convergence

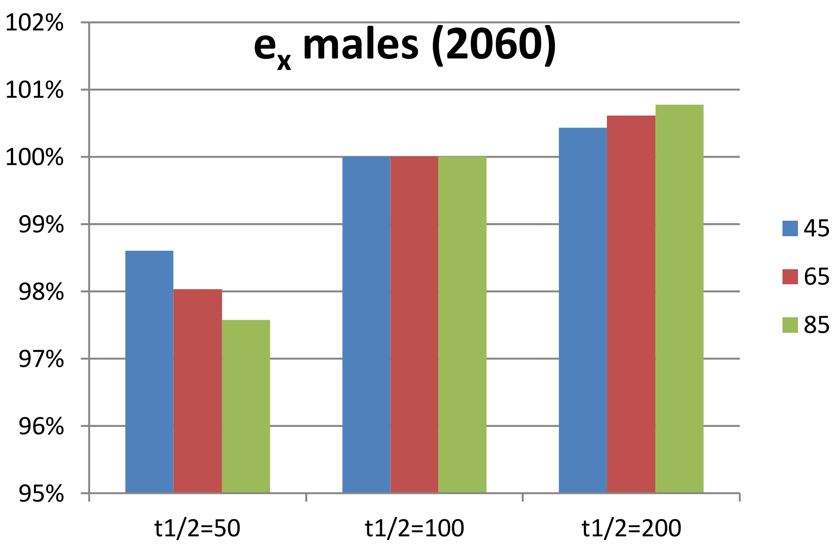
- This means that when the minimum is approached, countries with worse mortality will start to catch up
- Based on the life expectancy extrapolation, CZE will be at present AUT level approximately in 2023
- Reduction should be delayed from AUT reduction by approximately 12 years

Sensitivity analysis

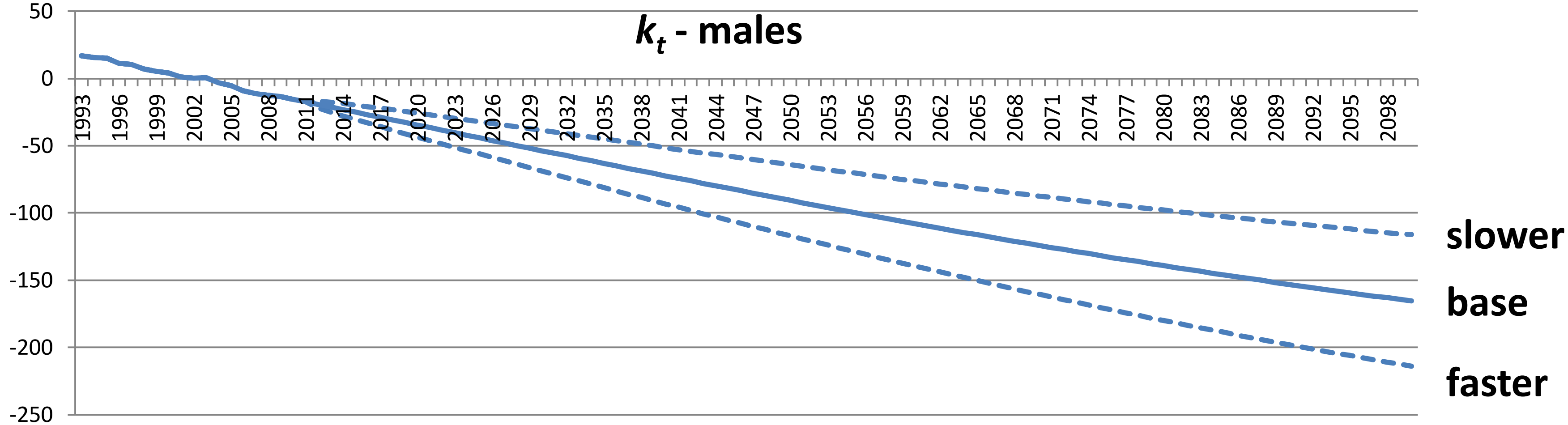
Speed of the long-term reduction

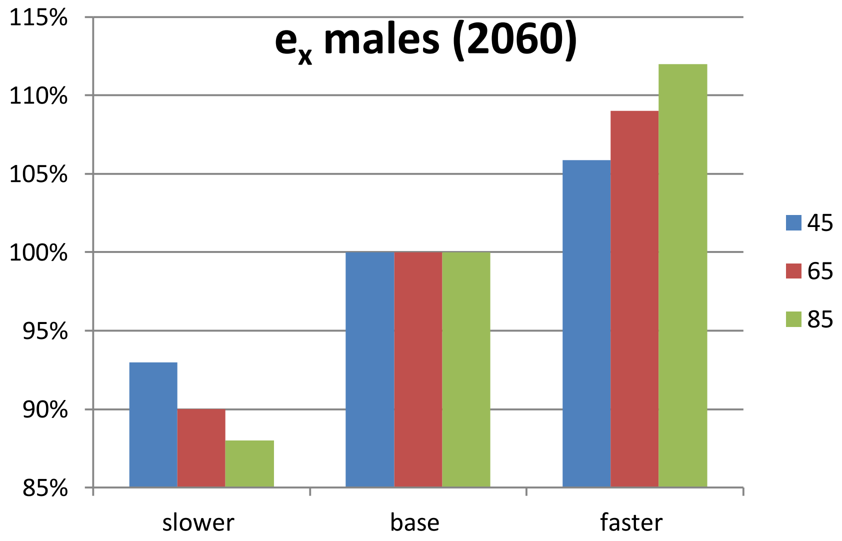
$$R(t) = \frac{1}{1 + \frac{t - t_0}{t_{1/2}}}$$

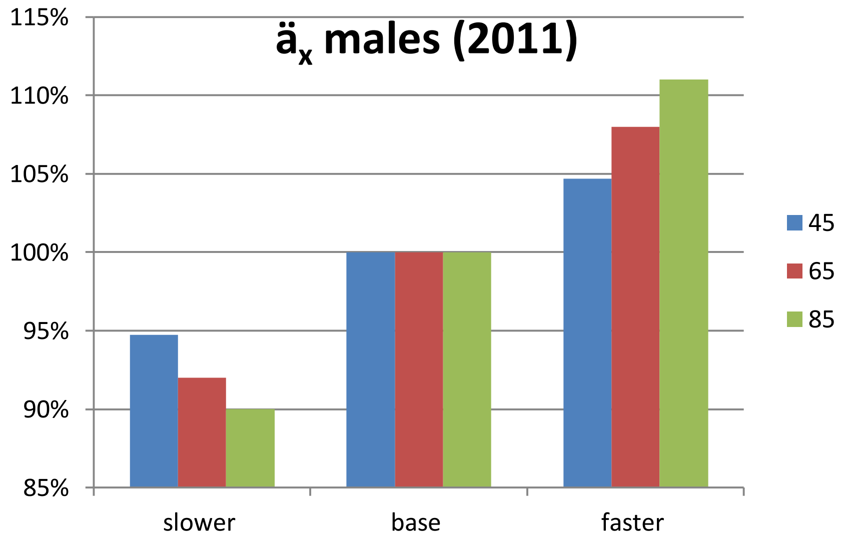
$t_{1/2}$ = 50 years
$t_{1/2}$ = 100 years (base scenario)
$t_{1/2}$ = 200 years



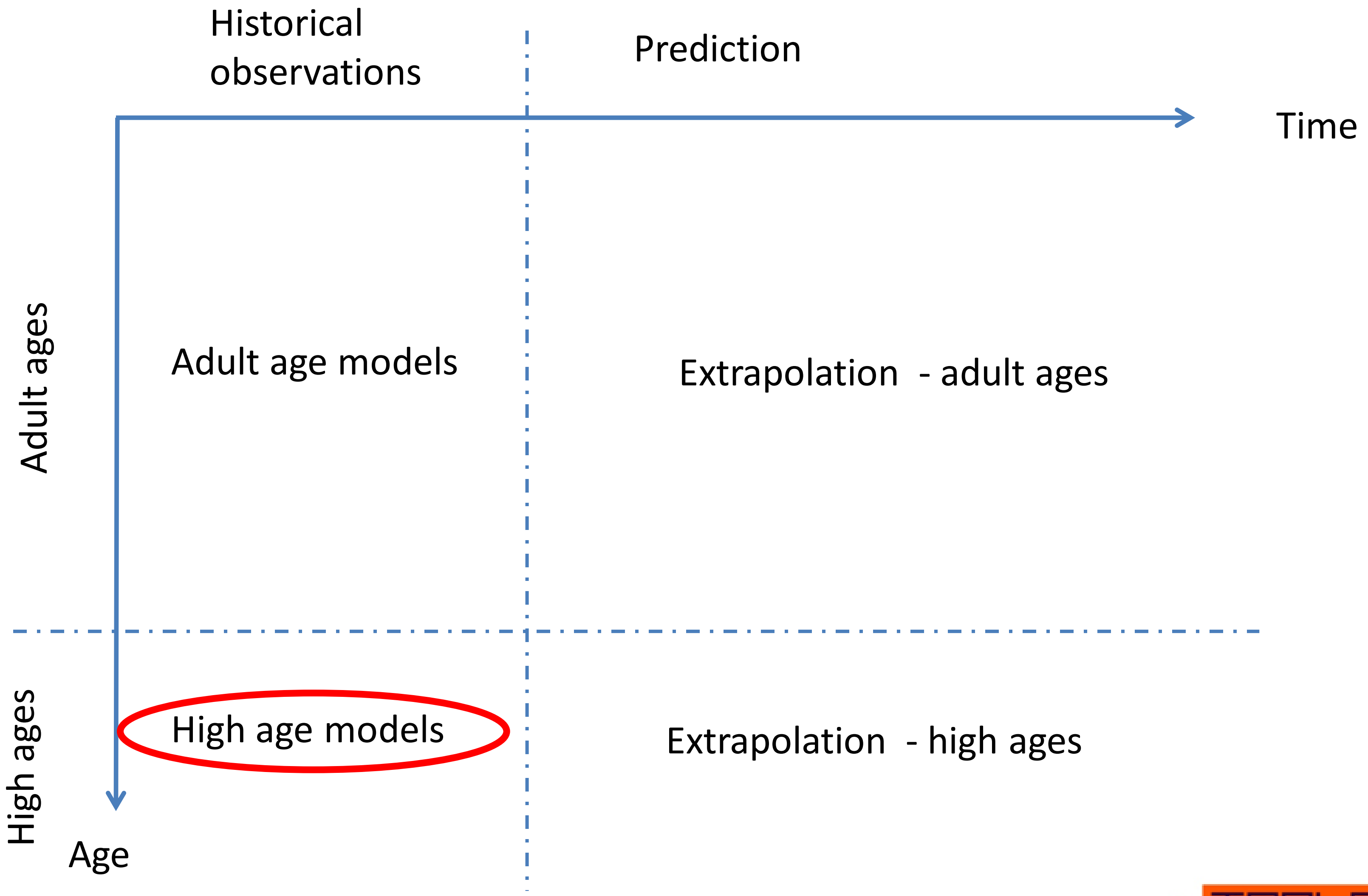
Sensitivity analysis







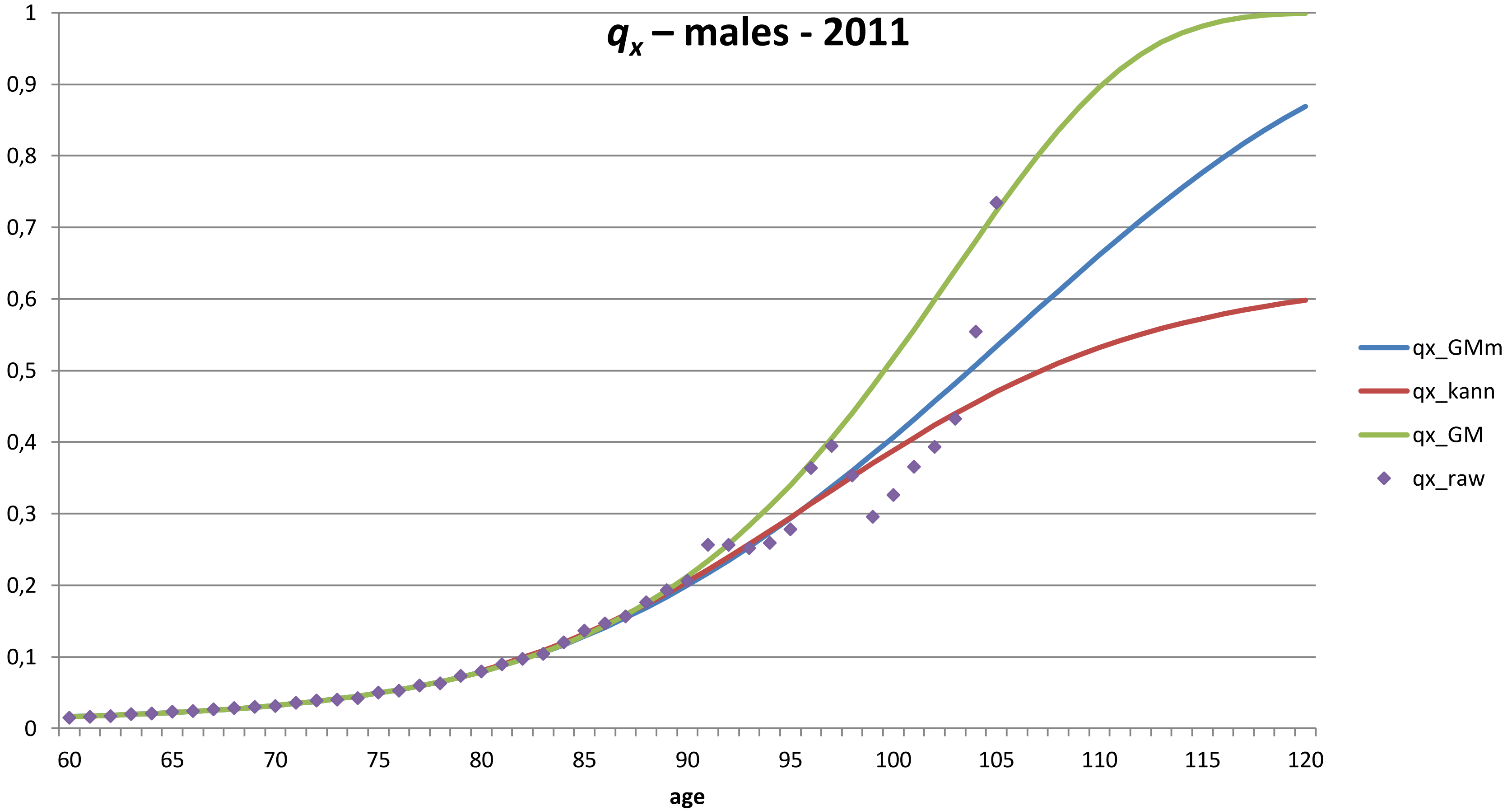
Phases of the modeling process



Modeling the high ages

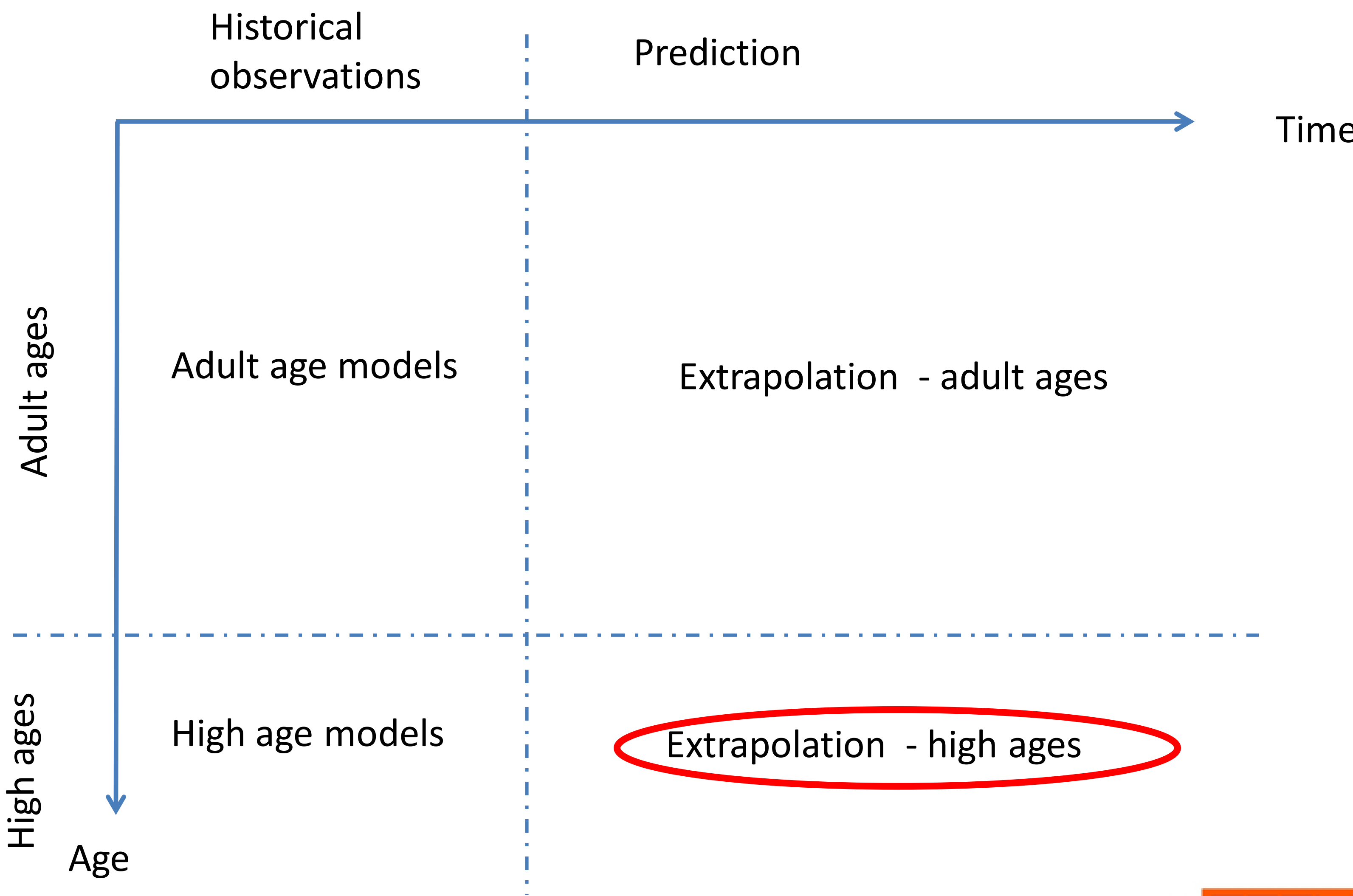
- Kannistö model $\ln\left(\frac{m_x}{1-m_x}\right) = \theta_0 + \theta_1(x - x_0)$
  - The logit transformation of death rates is expressed as a linear function of age
  - The model is considered as one of the most relevant for describing the mortality at the end of life
  - It is used in Human Mortality Database
  - Robust estimates – Same parameter estimates (MLE) for ages 80 – 90 and 80 – 95
  - "S-curve" shape
  - Forecasts best the ages 95 – 105

Modeling the high ages



TOOLS 4F

Phases of the modeling process



Extrapolation of high ages

- Smooth connection to real (predicted) data at the age $x_0 = 90$ has to be ensured.

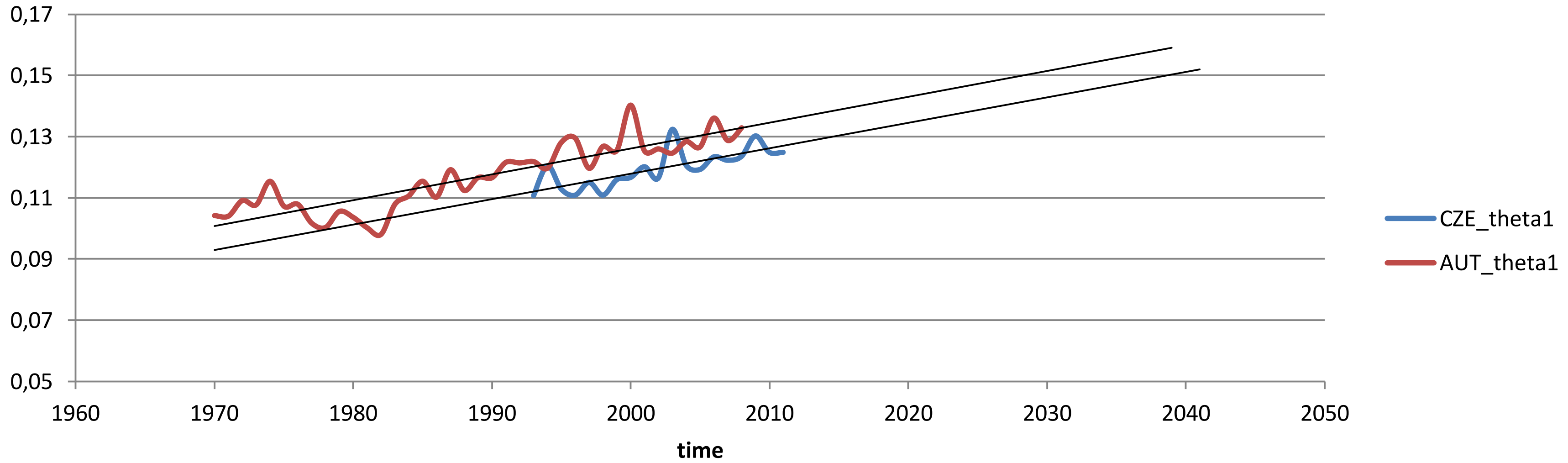
$$\ln\left(\frac{m_x}{1-m_x}\right) = \theta_0 + \theta_1 (x - x_0) \longrightarrow \ln\left(\frac{m_x}{1-m_x}\right) = \ln\left(\frac{m_{x_0}}{1-m_{x_0}}\right) + \theta_1 (x - x_0)$$

- Kannistö model is calibrated in each year from 1993 to 2011, thus series of estimates of θ_1 are obtained

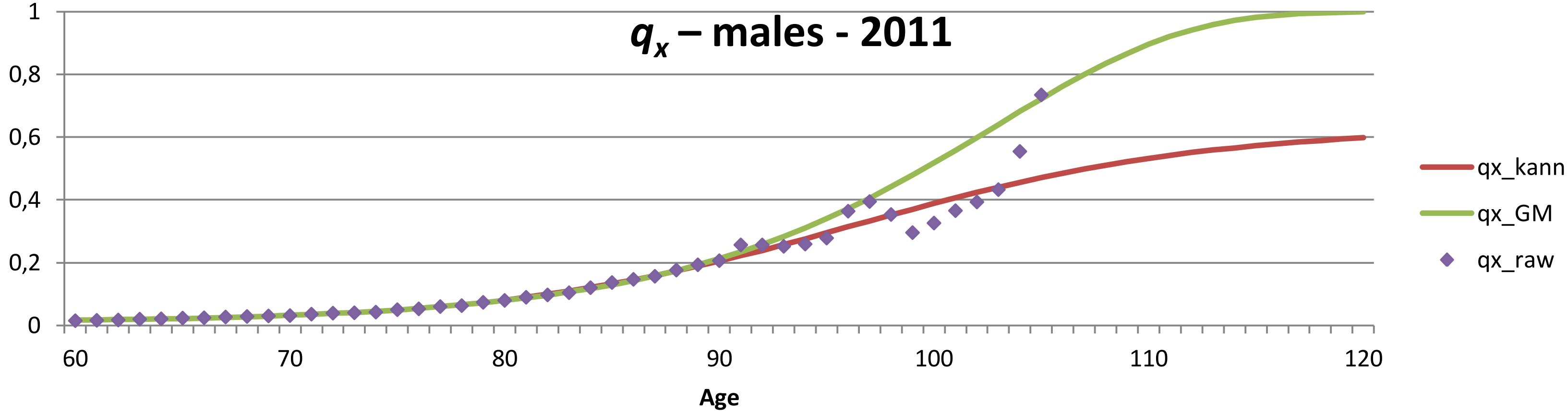
- We need to extrapolate the θ_1 up to 2060

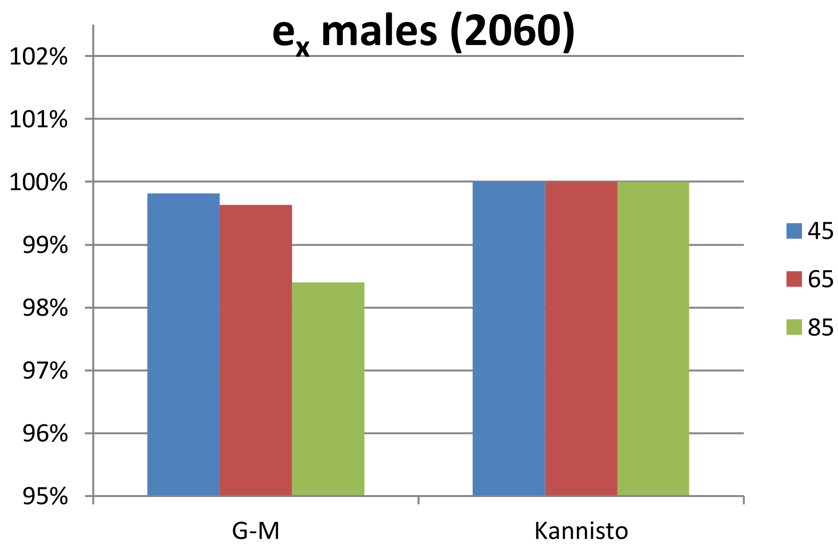
Local trend?

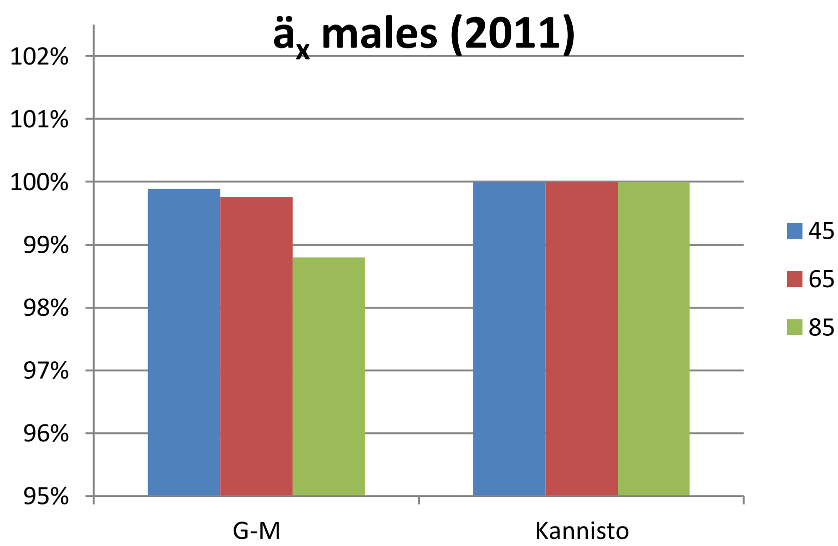
- The same situation as in the case of extrapolation of the parameter κ_t in the LC model
- CZE short term trend is similar to mid term AUT trend
- The reduction factor is also applied



Sensitivity analysis







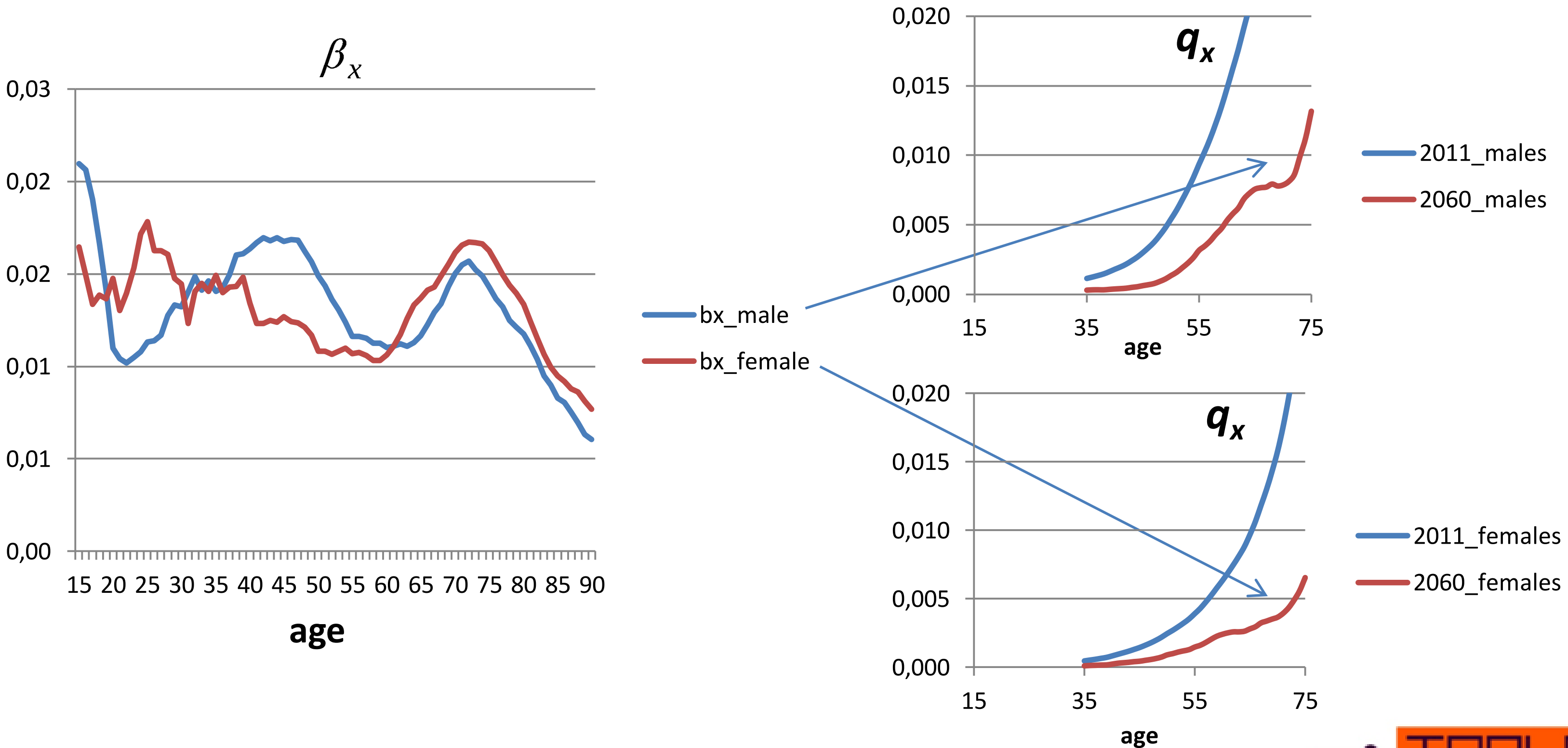
TOOLS 4F

Application problems

[http://www.helpingpsychology.com/problem-solving-in-cognitive-psychology]

Application problems – LC model

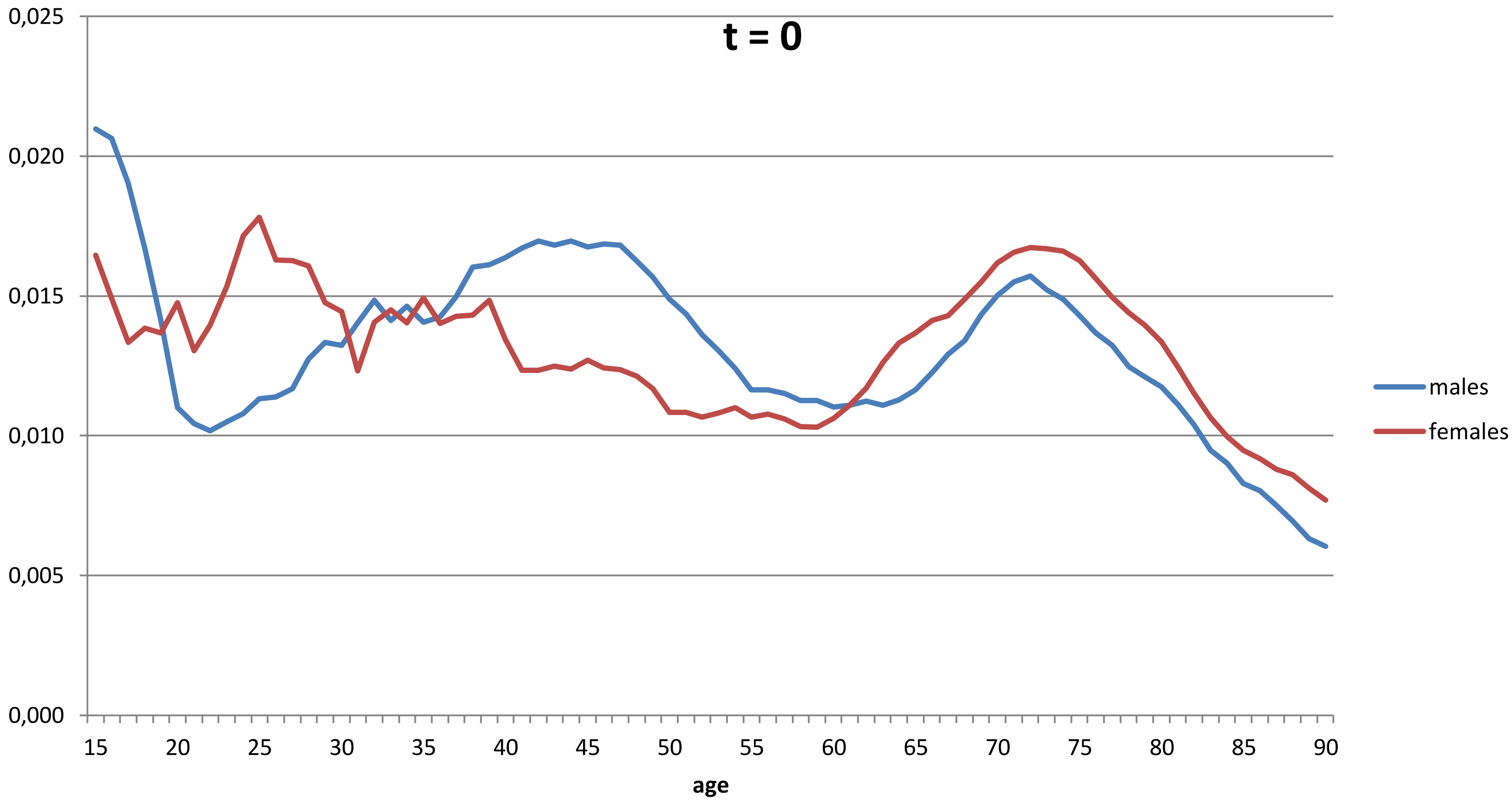
- Estimates (β_x) of the LC model does not ensure monotone predicted death probabilities



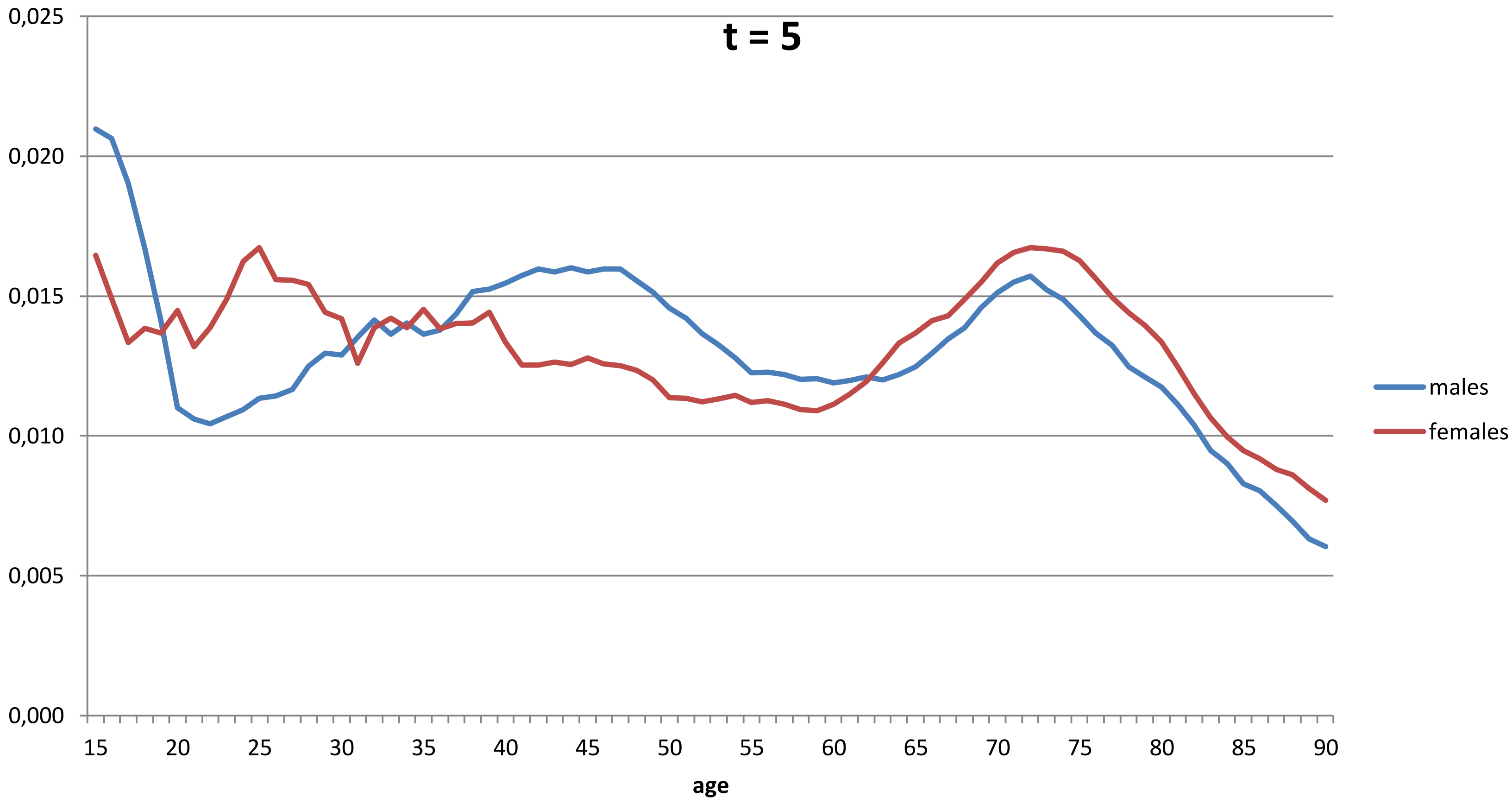
Application problems – LC model

- Monotone predicted death probabilities are ensured by “linearization” of betas
- In addition the indentifying constraint of the LC model $\sum_x \beta_x = 1$ must be met
  - Which implies the linearization for ages:
    for males at the age of 20 to 71 years
    for females at the age of 19 to 63 years
- The linearization is complete in 20 years (dependent on time)

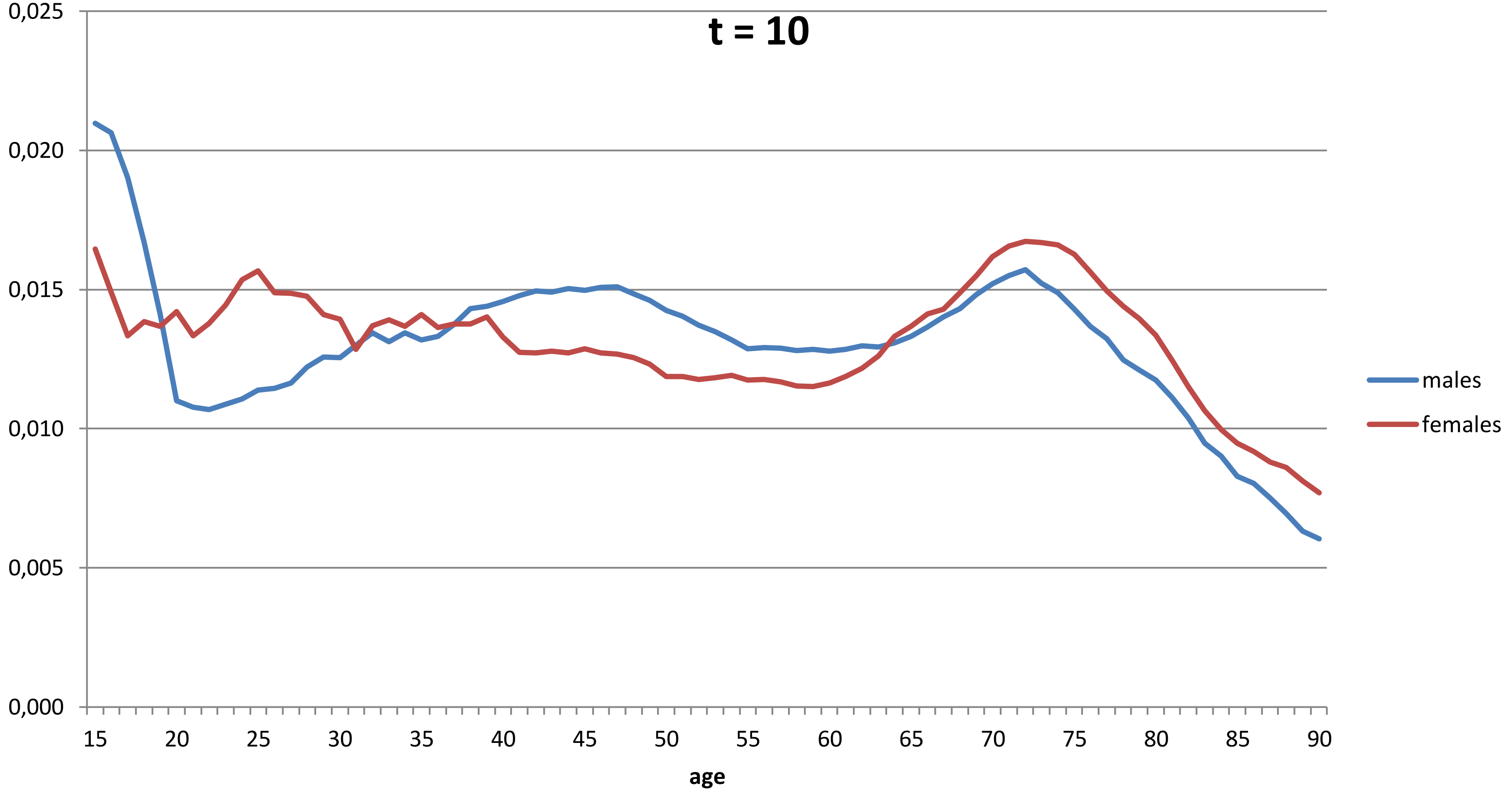
Application problems – β_x linearization



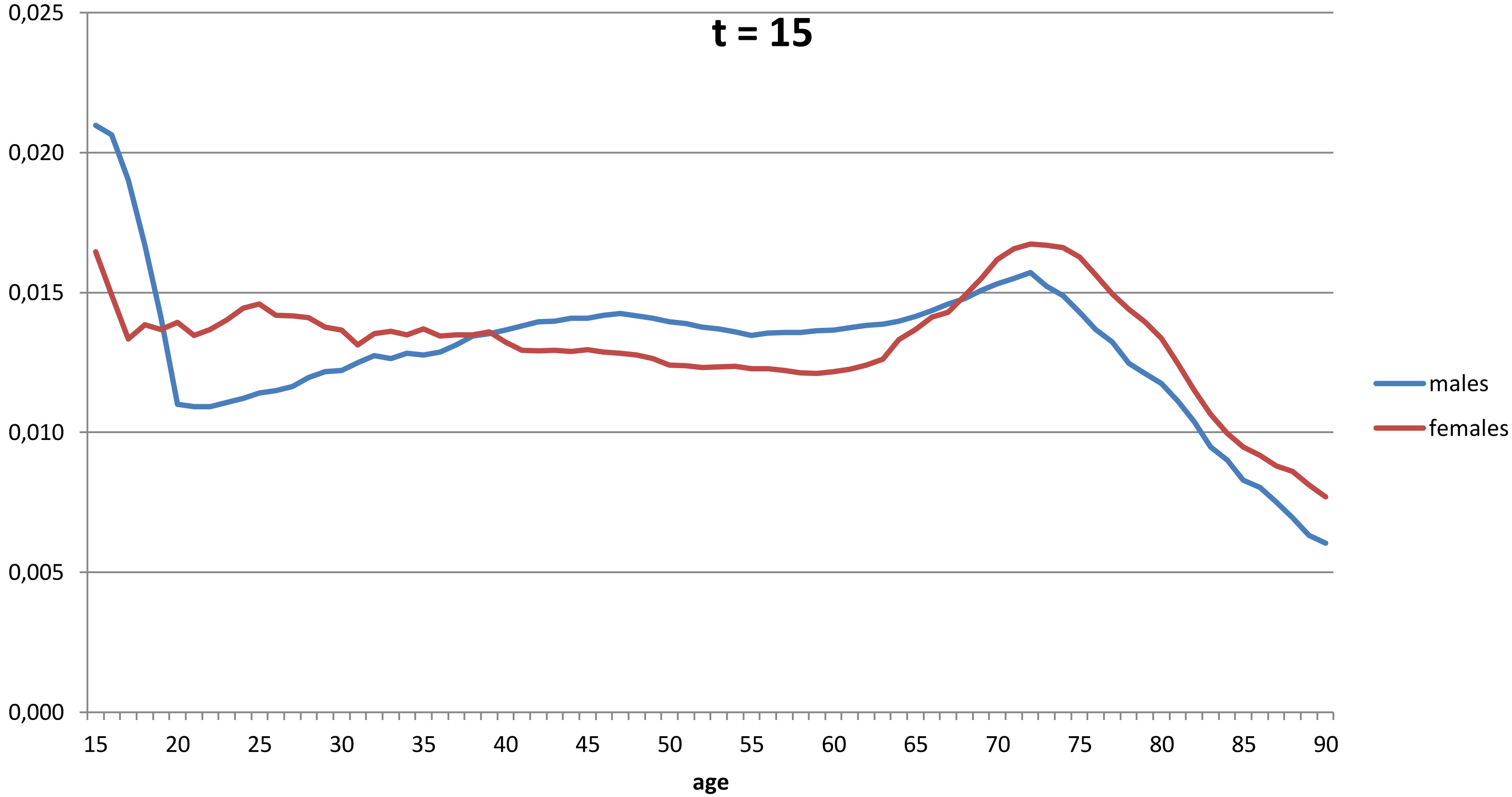
Application problems – β_x linearization



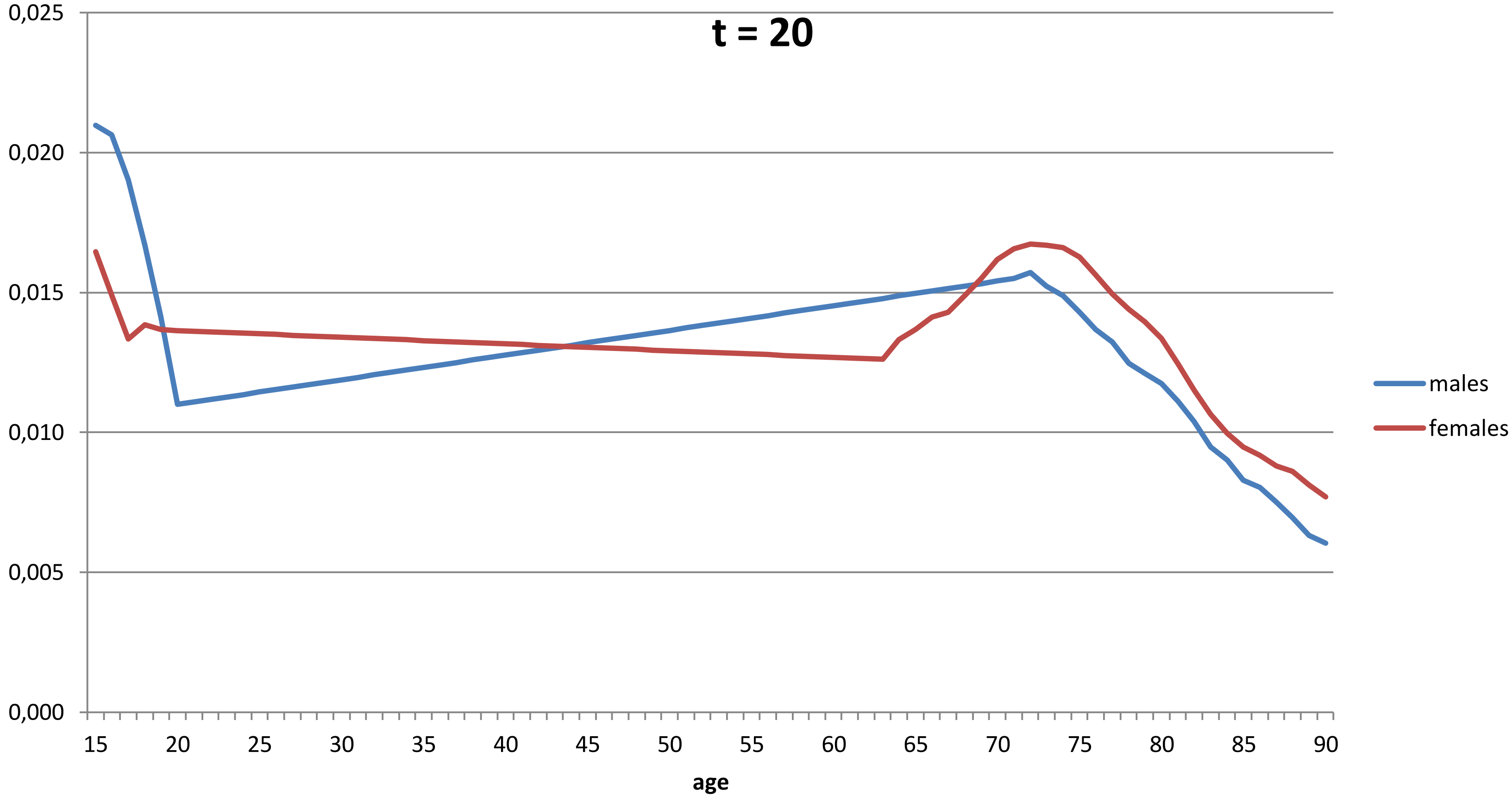
Application problems – β_x linearization



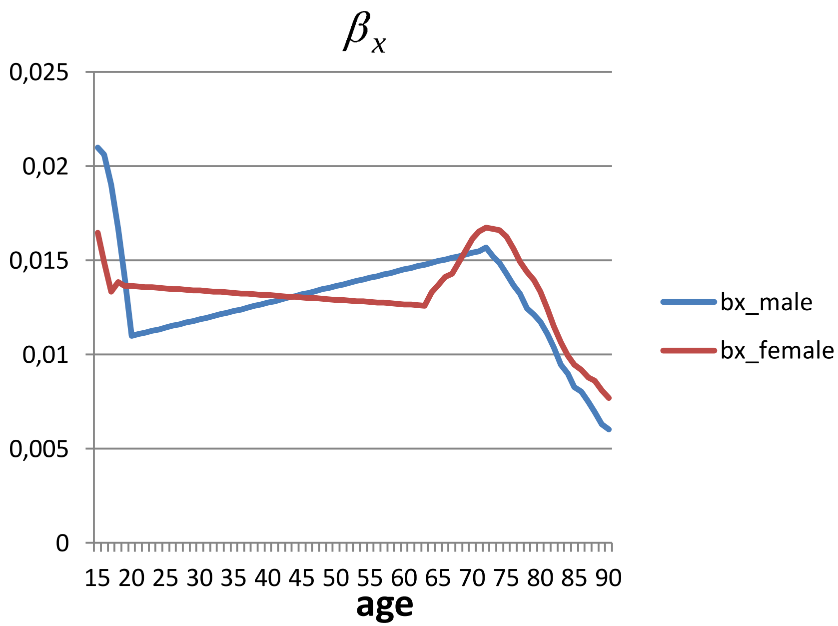
Application problems – β_x linearization

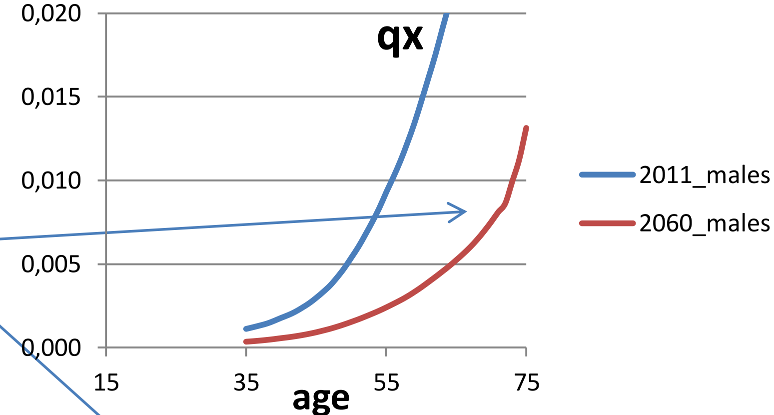


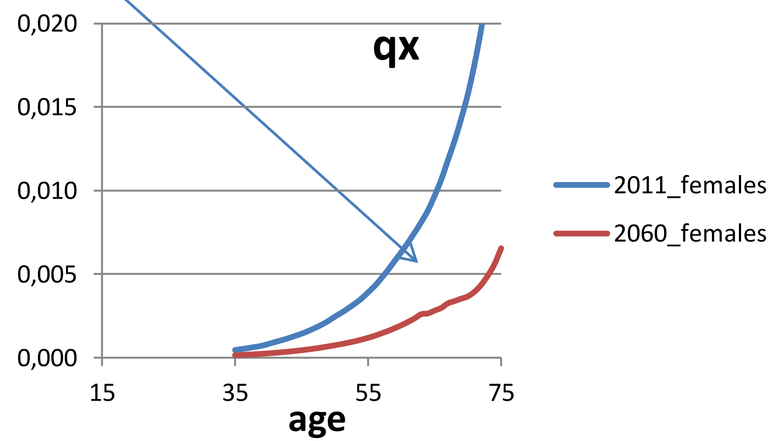
Application problems – β_x linearization



Application problems – LC model

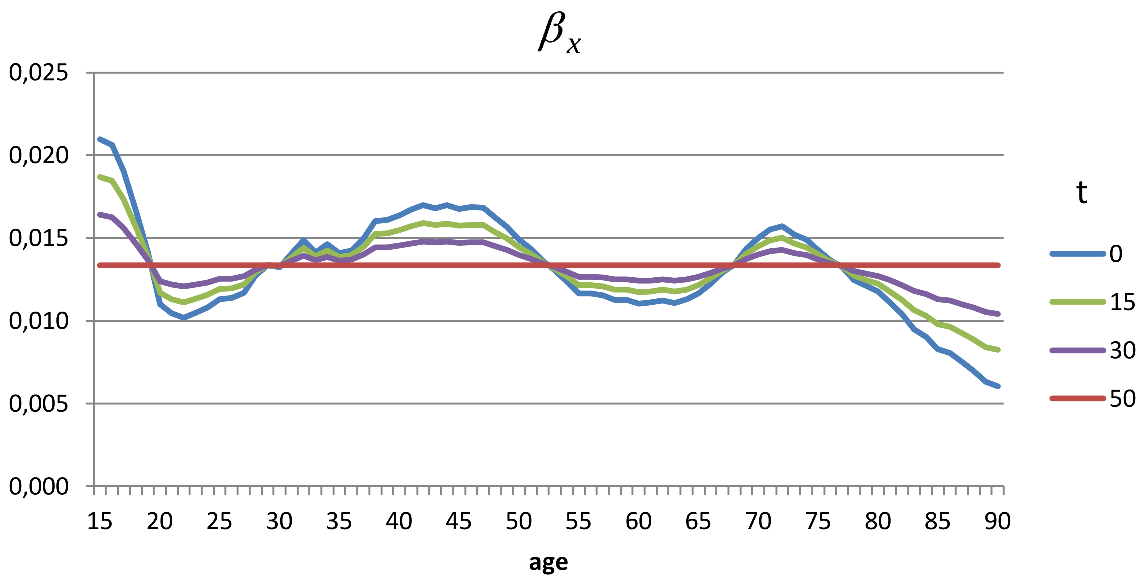




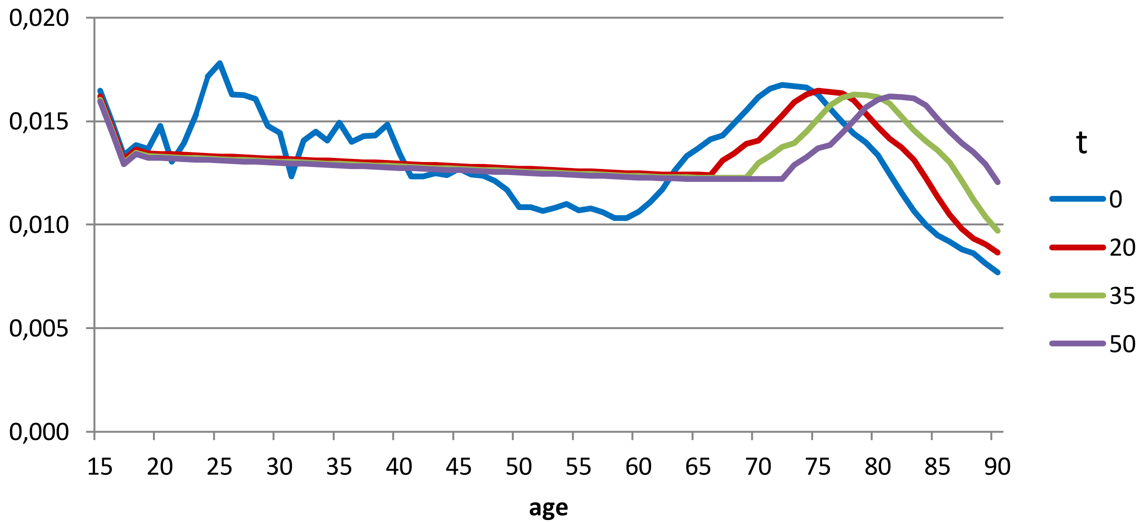


Males Linearization/Raw approach		
age	ex (2060)	äx (2011)
45	0,988	0,991
65	0,992	0,990
85	1,000	0,996

Sensitivity analysis



Males Alternative/Base approach		
age	ex (2060)	äx (2011)
45	1,05	1,03
65	1,10	1,02
85	1,44	1,03



Males Alternative/Base approach		
age	ex (2060)	äx (2011)
45	1,03	1,03
65	1,06	1,02
85	1,19	1,00

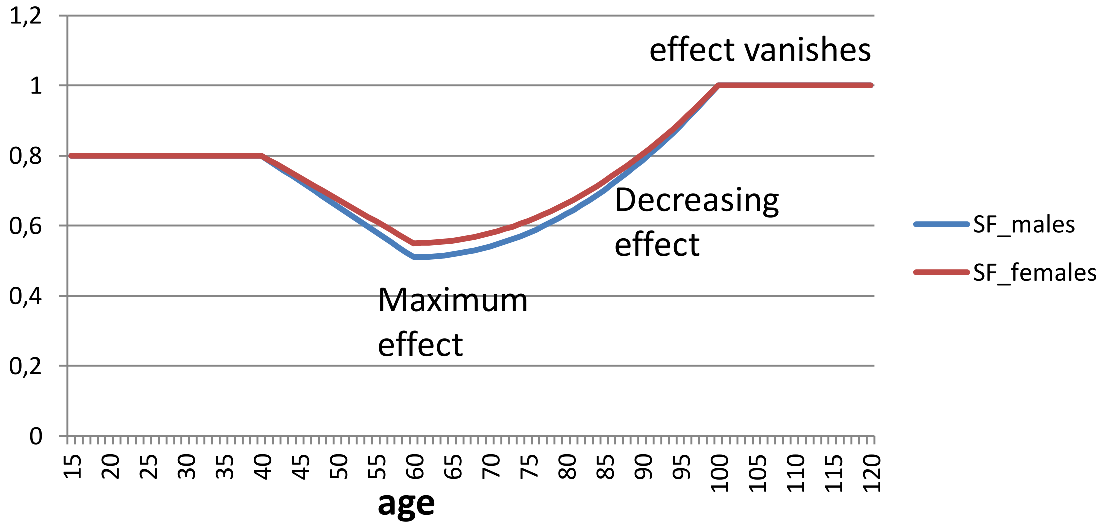
TOOLS 4F

Selection factors

- We have to take into account the different mortality of annuitants compared to the whole population
- The different social and health status structure of the group of annuitants
  - higher income
  - healthier people
- There are no data whatsoever to calibrate the impact of the different health status on the mortality of the annuitants in CZE

Selection factors

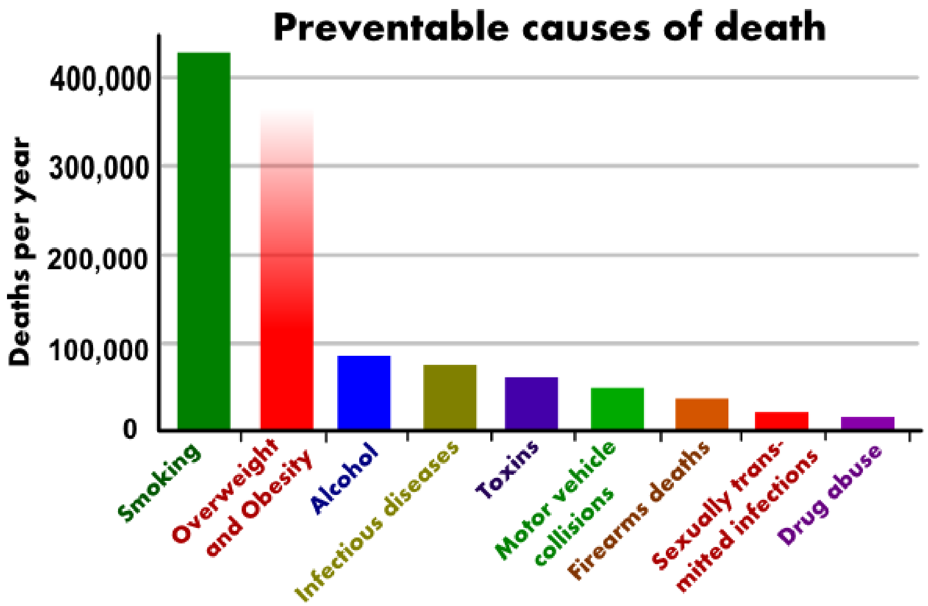
- German and Austrian selection factors were calculated from the data pooled by Gen Re and the Munich Re Group from more than 20 German insurance companies (period 1995-2002) for the purpose of The German table DAV 2004-R.



Further extensions

Causes of Death?

- Causes of death differ substantially
- Different dynamics
- Several categorizations e.g.:
  - Preventable
  - Amenable
  - Non-avoidable



Factors

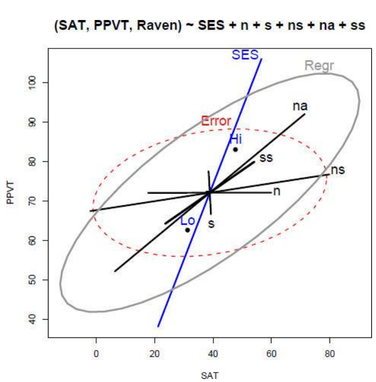
- Usually country, age and sex are used as factors
- But there are other significant factors
  - Education
  - Marital status
  - Address (city, altitude…)

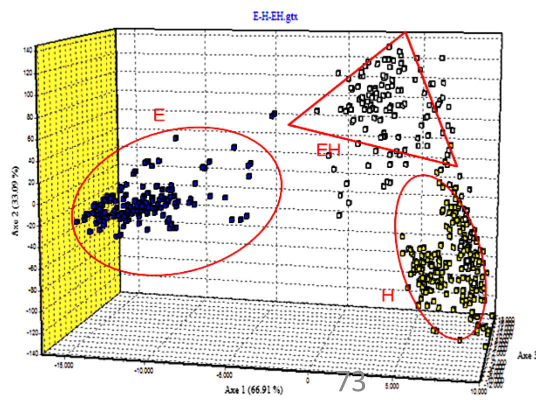
- Segmentation of the portfolio
  - Targeting new clients
  - Improve estimates on existing portfolio if the drivers (or its proxies) are available.

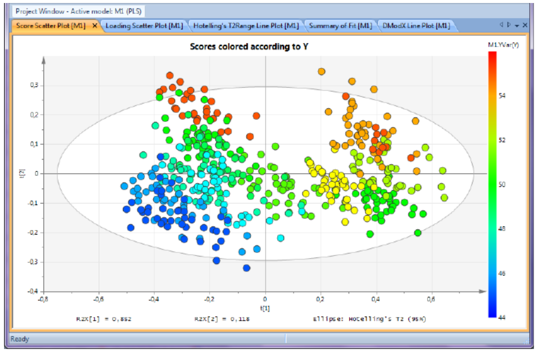
Statistical techniques

Several statistical techniques are available for modeling. For example correspondence analysis or multinomial logistic regression.

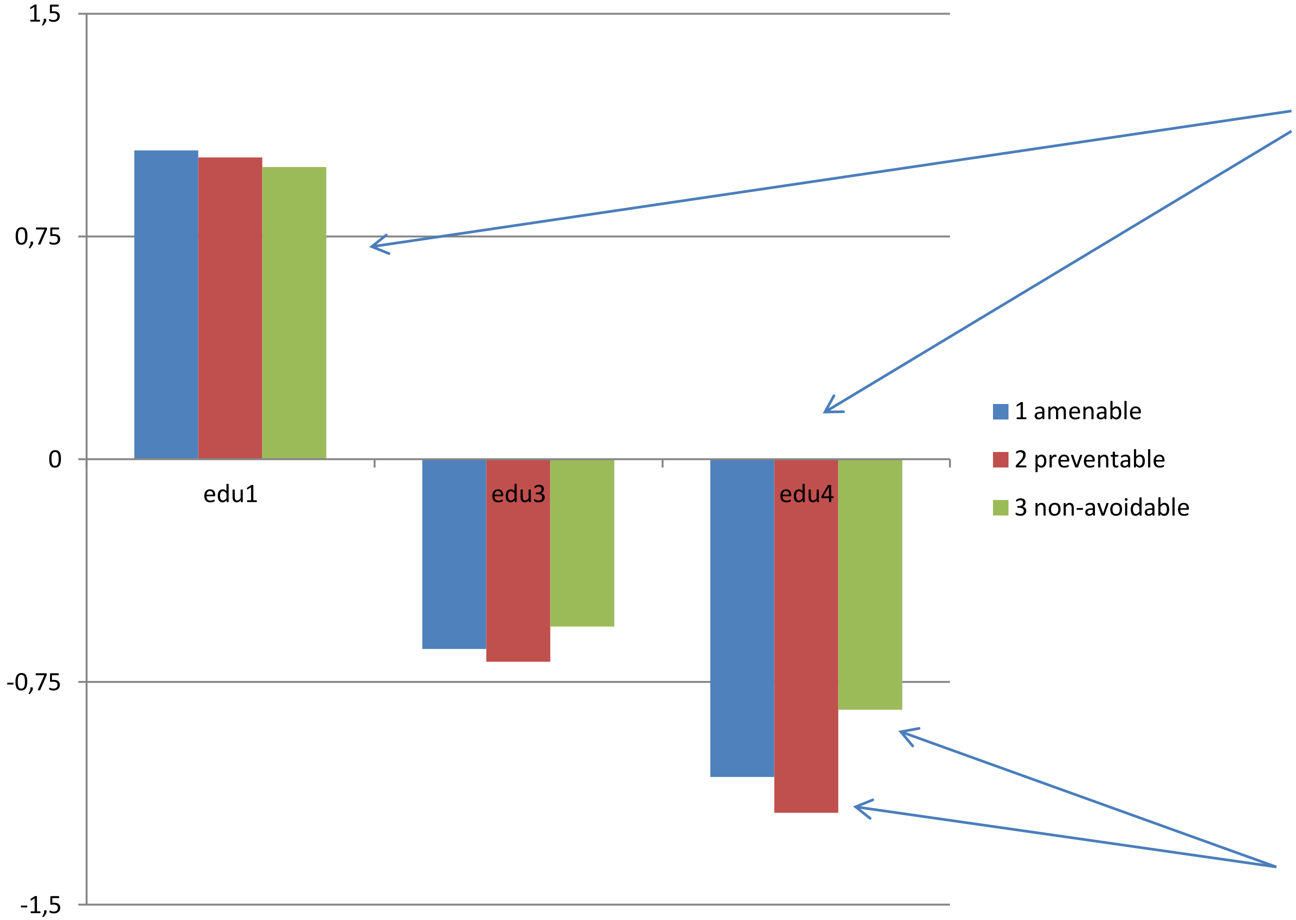
Some illustrative findings are presented on the following slides.







Impact of education level on different causes of death – multinomial reg.



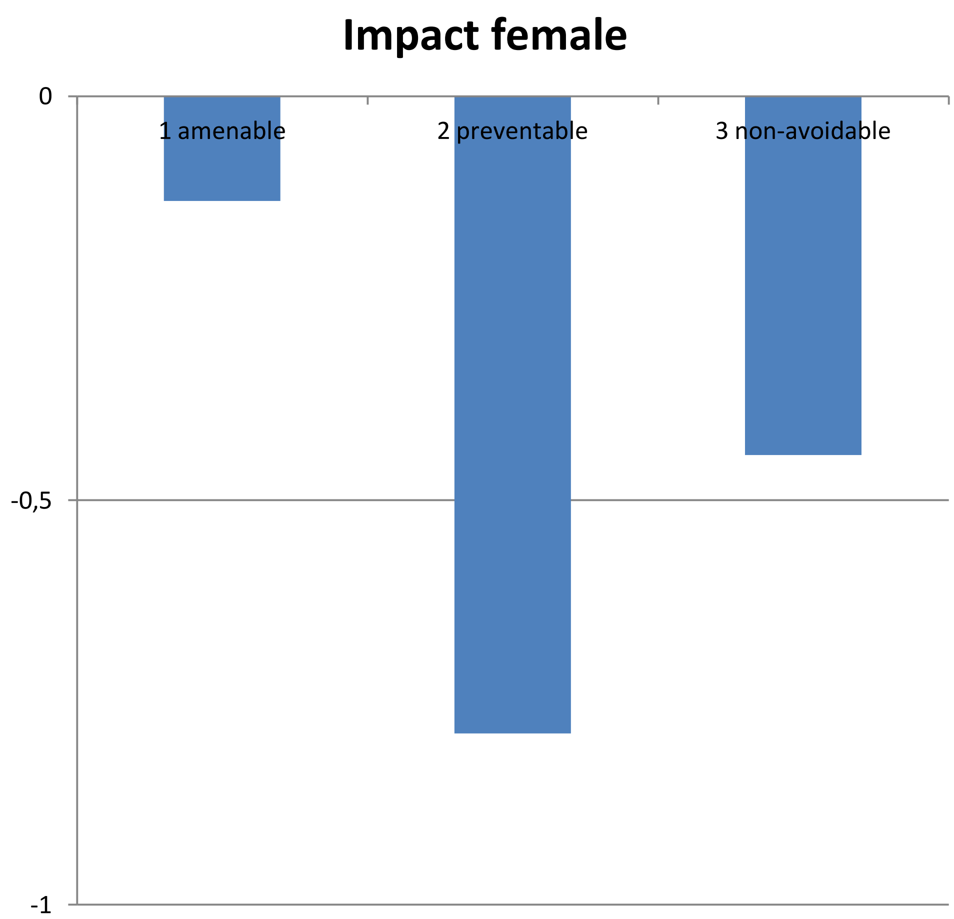
In general, **higher education decreases the mortality significantly**

Diagnoses are clustered as:

1. Amenable
2. Preventable
3. Non-avoidable

Higher education has even higher impact on preventable and amenable causes

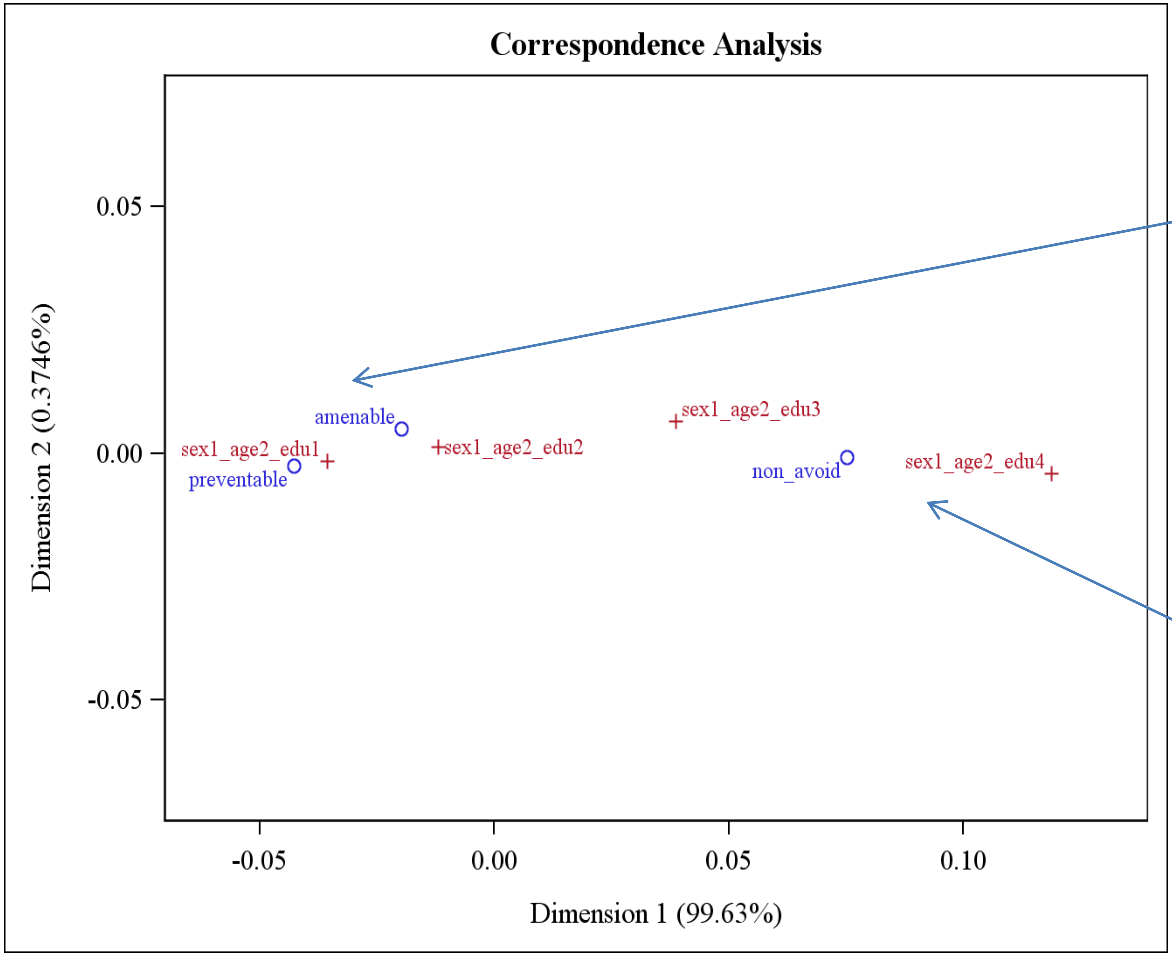
Impact of sex on different causes of death – multinomial reg.



It is widely known that women have lower mortality than men.

The **difference however varies through different causes**

Correspondence analysis



It is obvious that death on **preventable and amenable causes correspond** mostly with **lower education** levels

While as **non-avoidable causes correspond** mostly **with higher education** levels

Sources

BROUHNS, Natacha and DENUIT, Michel and VERMUNT, Jeroen K. A Poisson log-bilinear regression approach to the construction of projected lifetables. Insurance: Mathematics and Economics, 2002, 31.3: 373-393.

BURCÍN, Boris and TESÁRKOVÁ, Klára and ŠÍDLO, Luděk. Nejpoužívanější metody vyrovnávání a extrapolace křivky úmrtnosti a jejich aplikace na českou populaci). Revue pro výzkum populačního vývoje, 52:77-89, 2010.

CAIRNS, Andrew JG, BLAKE, David and DOWD, Kevin. Pricing death: Frameworks for the valuation and securitization of mortality risk. Astin Bulletin, 36(1):79, 2006.

CALDUCH VERDIELL, Núria, et al. Old age mortality: a state of the art. 2008. Bachelor's Thesis

Sources

CIPRA, Tomáš: Generační úmrtnostní tabulky pro důchodové pojištění a penzijní připojištění v České republice. Pojistné rozpravy 3 (1998), 31-57.

ČERNAYOVÁ PETRA, The Lee-Carter model for forecasting mortality with various extensions: Application to the Dutch mortality data, MSc thesis MFF UK, Praha 2008.

KAINHOFER, Reinhold; PREDOTA, Martin; SCHMOCK, Uwe. The new Austrian annuity valuation table AVÖ 2005R. 2006.

LEE, Ronald D and CARTER, Lawrence R. Modeling and forecasting US mortality. Journal of the American statistical association, 87(419):659{671,
1992.

Sources

MESLÉ, France; VALLIN, Jacques. Mortality in Europe: the divergence between east and west. Population (english edition), 2002, 57.1: 157-197

PITACCO, Ermanno, et al. Modelling longevity dynamics for pensions and annuity business. Oxford University Press, 2009.

VAN BROEKHOVEN, Henk. Market value of liabilities mortality risk: a practical model. North American Actuarial Journal, 2002, 6.2: 95-106

For further information please contact us.

Pavel Zimmermann

zimmermann@tools4f.com

+420 777 019 267

WE UNDERSTAND YOUR JOB