Czech Mortality Predictions
focused on pension insurance
Martin Matějka, Pavel Zimmermann

MFF UK, Praha 2014
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

Expectancy growth (Faster in AUT)

Similar trends

Mortality crisis

Graph showing life expectancy trends from 1940 to 2020 for males in CZE and AUT, indicating a faster growth rate in AUT compared to CZE.
(in years)

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.
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

\[ y = 0.2906x - 506.69 \]
\[ y = 0.2986x - 526.43 \]
\[ y = 0.252x - 475.3 \]
\[ y = 0.1125x - 216.87 \]
\[ y = 0.2906x - 506.69 \]
\[ y = 0.2385x - 444.68 \]
\[ y = 0.2906x - 506.69 \]
\[ y = 0.1205x - 231.39 \]
\[ y = 0.1125x - 216.87 \]

Crosses in:
- 2467
- 2253
- 1857

Crossed in:
- 1857
Does CZE catch up???

Death probabilities \( \log(q_x) \)

25

65

45

75

CZE

AUT
Does CZE catch up???

![Graph showing the time theoretically necessary to reach the average life expectancy at the precise age of 65 years (in calendar years) and the time theoretically necessary to reach the average life expectancy at the precise age of 35 years (in calendar years).](image)
Mortality models
Phases of the modeling process

- Historical observations
- Prediction

Time

Adult ages

- Adult age models
- Extrapolation - adult ages

High ages

- High age models
- Extrapolation - high ages

Age
Phases of the modeling process

- Historical observations
- Prediction

Time

Adult ages
- Adult age models
- Extrapolation - adult ages

High ages
- High age models
- Extrapolation - high ages

Age
Basic Models

- Dynamics of $q_{xt}$ or $\mu_{xt}$ is modelled.
- Reduction factor
  \[ q_{xt} = q_{x0} R(t) \]
  \[ q_{xt} = q_{x0} R(x,t) \]
  e.g. for linear trend in log($q_{xt}$)
  \[ R(x,t) = \exp(-\delta_x t) \]

- Or...
  \[ R_x(t - t') = \alpha_x + (1 - \alpha_x)(1 - f_x) \frac{t - t'}{20} \]

\[
\begin{align*}
  f_x &= \begin{cases} 
    c & \text{if } x < 60 \\
    1 + (1 - c) \frac{x - 110}{50} & \text{if } 60 \leq x \leq 110 \\
    1 & \text{if } x > 110
  \end{cases} \\
  \alpha_x &= \begin{cases} 
    b & \text{if } x < 60 \\
    \frac{(110 - x) b + (x - 60) k}{50} & \text{if } 60 \leq x \leq 110 \\
    k & \text{if } x > 110
  \end{cases}
\end{align*}
\]
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.

\[ \hat{\mu}_{xt} = f(x; \hat{\Theta}_t) \]
Cairns – Blake – Dowd

Example: Cairns – Blake – Dowd (CBD)

Example: 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 \(t\)
- \(\alpha_x\) defines the shape of the age profile of mortality averaged over time
- \(\beta_x\) represents the pattern of deviations from the age profile of mortality
- \(\kappa_t\) describes the variation in the general level of mortality
Parameters of the LC model

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

- \( \alpha_x \) - age profile of mortality averaged over time
Parameters of the LC model

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

- \( \kappa_t \) - **general level** of mortality (independent on age)
Parameters of the LC model

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

- \( \beta_x \) - how rapidly or slowly mortality at each age varies when the general level of mortality (\( \kappa_t \)) 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

- Historical observations
- Prediction

Time

Adult ages

- Adult age models
- Extrapolation - adult ages

High ages

- High age models
- Extrapolation - high ages

Age
Modeling the high ages

observed $q_x$ - males 2011

High ages

age

0
0,1
0,2
0,3
0,4
0,5
0,6
0,7
0,8
0,9
1

80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120
Modeling the high ages

Modeling the high ages

- **Exponential models**
  - Gompertz-Makeham (Koschin)
    \[ \mu_x = a + b \cdot c^x \]
  - Kannistö
    \[ \ln \left( \frac{m_x}{1 - m_x} \right) = \theta_0 + \theta_1 (x - x_0) \]

- **Logistic models**

- **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

Death rates – males - 2011

age

Death rates
Choosing the age-time period

\[ \ln(q_x) - \text{males} \]

Historical observations

Prediction

1993

2011

15

90

\[ \ln(q_x) - \text{males} \]

-4.5

-4

-3.5

-3

-2.5

-2

-1.5

-1


45-49 50-54 55-59 60-64 65-69

Adult ages

High ages

Time

Age
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
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 \(\kappa_t\) is recommended to better fit the model and the observed deaths \((d_{xt})\)
Estimation of the LC model

2) Weighted least squares

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

$$w_{xt} = d_{xt}$$ 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 $\kappa_t$

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

\[ D_{xt} \sim \text{Poisson}(\lambda_{xt} = E_{xt} m_{xt} = E_{xt} \exp(\alpha_x + \beta_x x_t)) \]

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

\[ \alpha_x \]

\[ \beta_x \]

\[ \kappa_t \]
Phases of the modeling process

- Historical observations
- Prediction

- Adult age models
- Extrapolation - adult ages

- High age models
- Extrapolation - high ages

Time

Adult ages

High ages

Age
Prediction of the LC model

- Prediction of the $\kappa_t$ 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.
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 $\Delta$ 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)
$t_0 = \text{year from which the reduction starts}$
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

Graphs showing the reduction in life expectancy and mortality rates for males in 2060 and 2011 for different values of \( t_{1/2} \).
Sensitivity analysis

$k_t$ - males

Slower  Faster

Base

$e_x$ males (2060)

$\ddot{a}_x$ males (2011)
Phases of the modeling process

- Historical observations
- Prediction

Adult ages

- Adult age models

High ages

- High age models

Extrapolation - adult ages

Extrapolation - high ages
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

\[ q_x - \text{males - 2011} \]
Phases of the modeling process

- Historical observations
- Prediction

<table>
<thead>
<tr>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adult ages</td>
</tr>
<tr>
<td>Adult age models</td>
</tr>
<tr>
<td>Extrapolation - adult ages</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>High ages</td>
</tr>
<tr>
<td>High age models</td>
</tr>
<tr>
<td>Extrapolation - high ages</td>
</tr>
</tbody>
</table>

Age
Extrapolation of high ages

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

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

Kannistö model is calibrated in each year from 1993 to 2011, thus series of estimates of $\theta_1$ are obtained.

We need to extrapolate the $\theta_1$ up to 2060.
Local trend?

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

$q_x - males - 2011$

$e_x males (2060)$

$\ddot{a}_x males (2011)$
Application problems

Application problems – LC model

Estimates ($\beta_x$) 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 identifying 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 – $\beta_x$ linearization

![Graph showing the $\beta_x$ linearization for males and females over age]

- $t = 0$
- Age range: 15 to 90
- $\beta_x$ values for males and females
Application problems – $\beta_x$ linearization

t = 5

Age

Males

Females
Application problems – $\beta_x$ linearization

![Graph showing $\beta_x$ linearization for males and females at $t = 10$.]
Application problems – $\beta_x$ linearization

t = 15
Application problems – $\beta_x$ linearization

![Graph showing $t = 20$ for males and females over age]
Application problems – LC model

\[ \beta_x \]

\[ qx \]

<table>
<thead>
<tr>
<th>Males</th>
<th>Linearization/Raw approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>age</td>
<td>ex (2060)</td>
</tr>
<tr>
<td>45</td>
<td>0.988</td>
</tr>
<tr>
<td>65</td>
<td>0.992</td>
</tr>
<tr>
<td>85</td>
<td>1.000</td>
</tr>
</tbody>
</table>

(ex (2060))

(åx (2011))
Sensitivity analysis

\[ \beta_x \]

<table>
<thead>
<tr>
<th>age</th>
<th>ex (2060)</th>
<th>äx (2011)</th>
</tr>
</thead>
<tbody>
<tr>
<td>45</td>
<td>1,05</td>
<td>1,03</td>
</tr>
<tr>
<td>65</td>
<td>1,10</td>
<td>1,02</td>
</tr>
<tr>
<td>85</td>
<td>1,44</td>
<td>1,03</td>
</tr>
</tbody>
</table>
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.
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


Sources


Sources


For further information please contact us.

Pavel Zimmermann
zimmermann@tools4f.com
+420 777 019 267