COS-D407. Scientific Modeling and Model Validation

Week 3
Model Validation Example: Validation framework for demographic mortality forecast models

Guest Lecturer: Ricarda Duerst[1]

UNIVERSITY OF HELSINKI, FINLAND

01.11.2021 – 15.12.2021

[1]PHD STUDENT AT UNIVERSITY OF HELSINKI ANDMAX PLANCK INSTITUTE FOR DEMOGRAPHIC RESEARCH, ROSTOCK, GERMANY

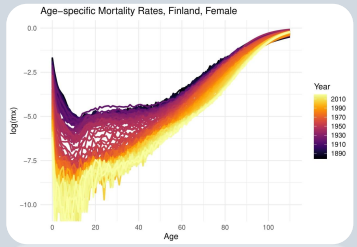
Agenda

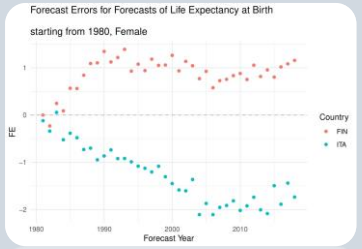
TODAY

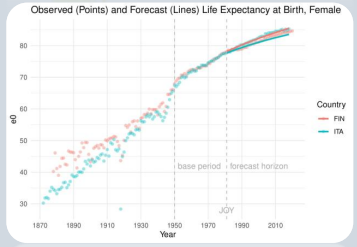
- Human Mortality
- Mortality Forecasting
- Forecast Validation
- Our Validation Framework
- !! Preparations for Wednesday !!

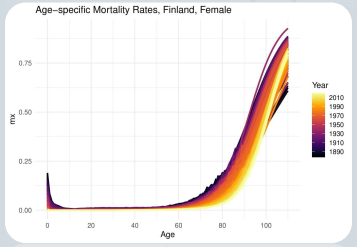
WEDNESDAY

Hands-on exercises on mortality forecast validation









Human Mortality

- Many demographic measures to talk about human mortality, e.g.:
- **Crude death rate**: $CDR = \frac{Number\ of\ deaths}{Number\ of\ person{-}years\ lived} = \frac{D}{N}$
- **Age-specific mortality rate**: $_n m_x = \frac{Number\ of\ deaths\ in\ age\ range\ x\ to\ x{+}n}{Number\ of\ person{-}years\ lived\ in\ age\ range\ x\ to\ x{+}n} = \frac{_n D_x}{_n N_x}$
- Distribution of the **age at death** D_x
- **Life expectancy at birth**: e_0 calculated using the Life Table,
  a summary measure of the current mortality experience of a population
- **Life span disparity at birth**: $e_0^\dagger$ additional step of the Life Table,
  a summary measure for the variability of the age at death

Human Mortality:
Finland

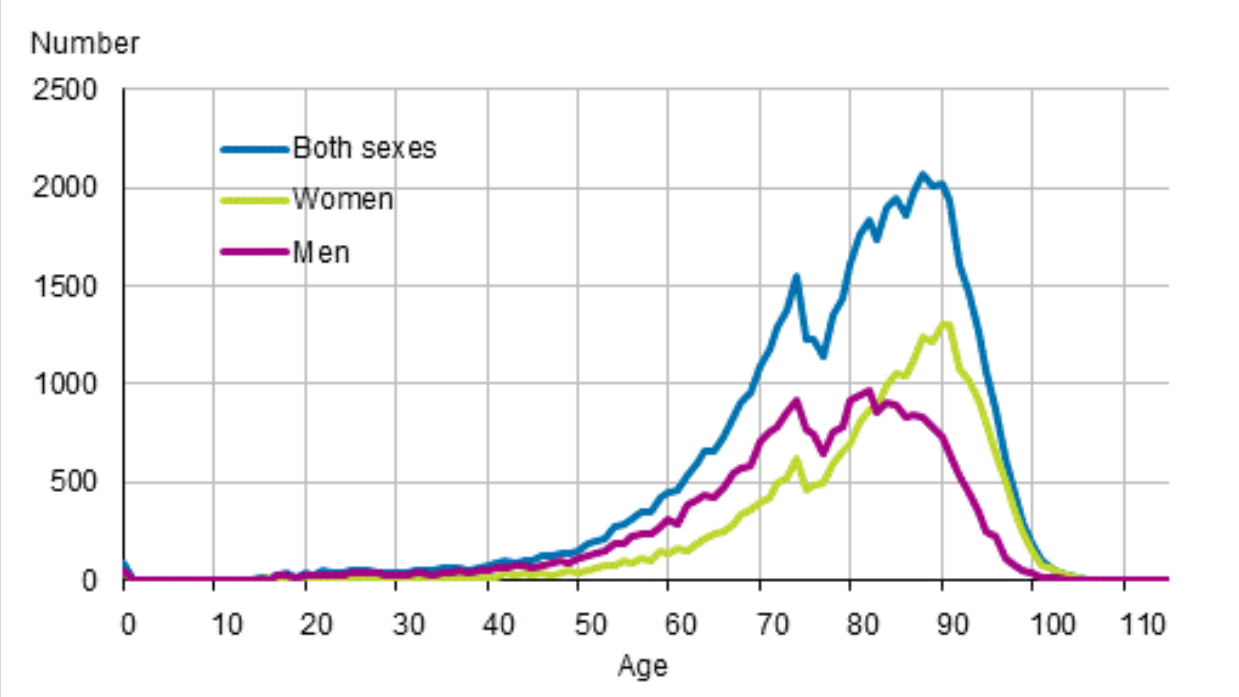
- Data source: Statistics Finland https://www.stat.fi/til/kuol/2020/kuol_2020_2021-04-23_tie_001_en.html
- Year 2020:

CDR = 9.87 deaths per 1000 inhabitants

Mean age at death: 75.3 years for men, 82.1 years for women

Life expectancy at birth: 79 years for boys, 84.6 years for girls

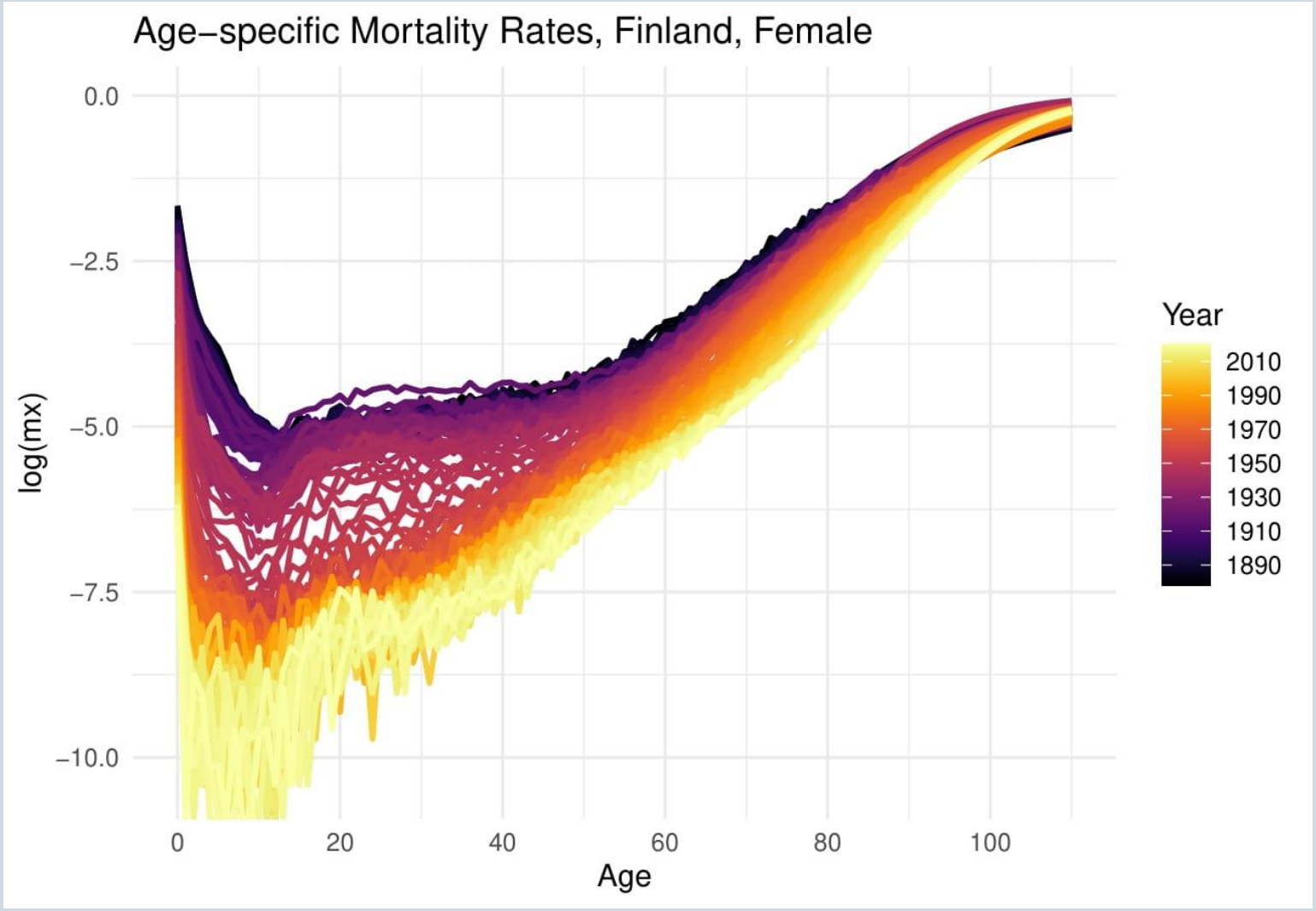
Age distribution at the time of death by sex in 2020



Source: Statistics Finland https://www.stat.fi/til/kuol/2020/kuol_2020_2021-04-23_tie_001_en.html

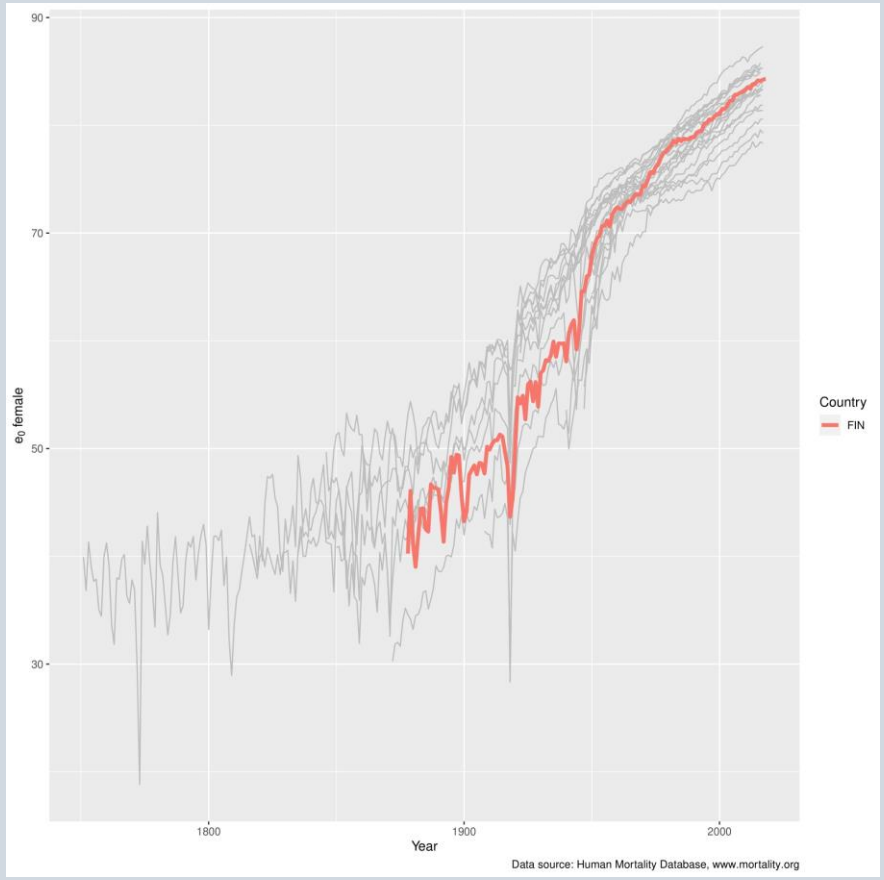
Question for you!

At which ages happened the strongest mortality improvements in Finland
a) in the past
b) and recently?

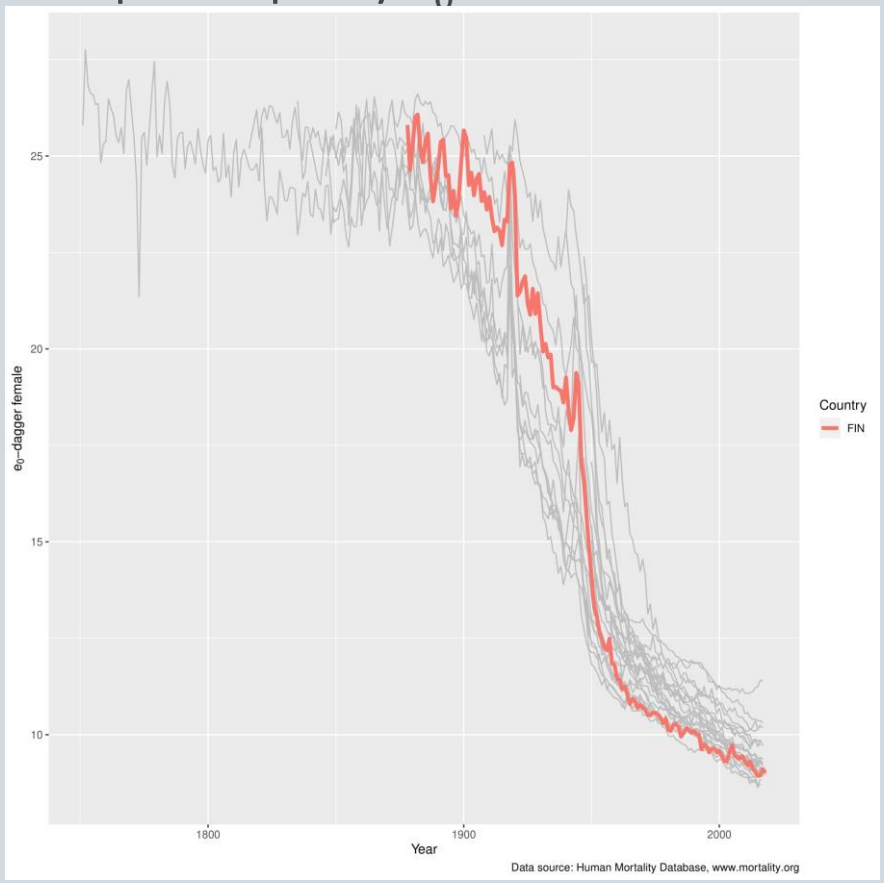


Human Mortality:
Finland

Life expectancy at birth e_0



Life span disparity $e_0^\dagger$



Mortality Data

Human Mortality Database (HMD)

- www.mortality.org
- by Max Planck Insitute for Demographic Research, Rostock and University of California, Berkeley
- detailed population and mortality data for 41 countries
- data access with free account → see preparations for Wednesday!

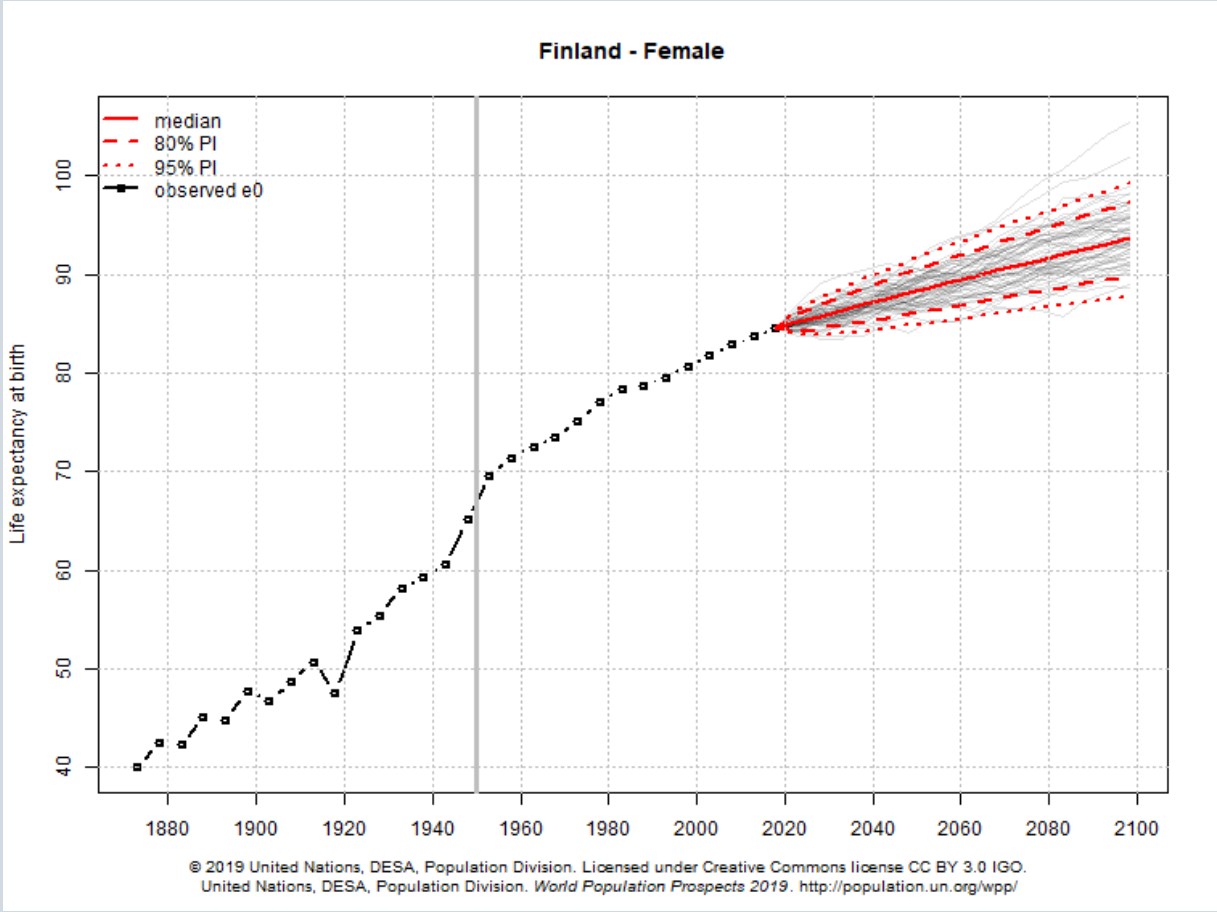
UN World Population Prospects (UN WPP)

- https://population.un.org/wpp/
- by Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat
- global population estimates and projections for key demographic indicators including mortality
- free data access

Mortality Forecasting

- Different approaches:
  - Extrapolation (continuing observed trends), e.g. Lee-Carter model (1992)
  - Explanation (including causes of deaths with known determinants), e.g. lung cancer mortality
  - Expectation (based on expert opinions)
  - Deterministic vs. Probabilistic (point forecasts vs. intrduction of uncertainty with prediction intervals)
  - Forecasts of age-specific mortality rates or life expectancy at birth directly
- See e.g.: Booth, H., & Tickle, L. (2008). Mortality modelling and forecasting: A review of methods. *Annals of actuarial science*, *3*(1-2), 3-43. https://www.actuaries.org/CTTEES_TFM/Documents/Booth_Tickle.pdf

Mortality Forecasting:
Probabilistic Projection of e_0 for Finnish Females



Mortality Forecasting:
Lee Carter Model (1992)

- Golden standard in mortality forecasting → widely used and approved
- Methodologically simple extrapolation model → basic assumption: age-specific trends in the past will continue in the future
- Age parameter b_x and time-varying parameter k_t → estimated using singular value decomposition, then fixed
- Extrapolation of k_t using time series model (random walk with drift)

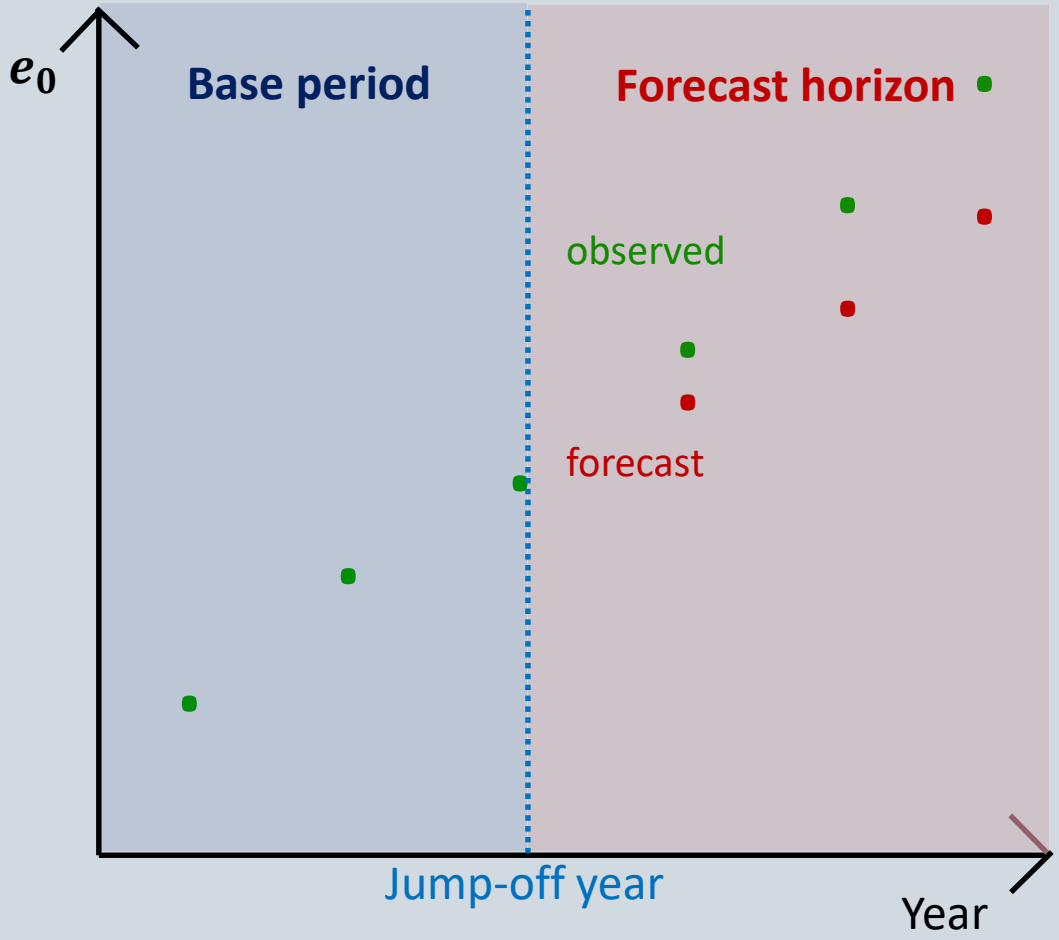
$$\ln(m_{x,t}) = a_x + b_x * k_t + \varepsilon_{x,t}$$

t	Year
x	Age
$\ln(m_{x,t})$	Log death rates
a_x	Average of $\ln(m_{x,t})$ in base period by age
b_x	Change in age-specific mortality
k_t	Overall level of mortality
$\varepsilon_{x,t}$	Error term

Forecast Validation

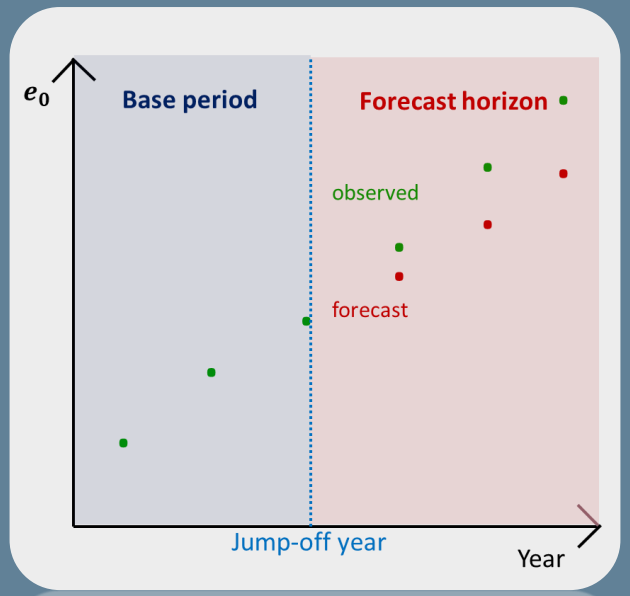
Ex-post validation (or out-of sample validation):

- Checking performance in hind-sight
- Principal: withholding **observed** data and „simulating“ a forecast for the respective period to compare the **forecast** results with the observed values
- **Base period** (BP): years with input data for forecast
- **Jump-off year** (JOY): last year of BP
- **Forecast Horizon** (FH): forecast years



Question for you!

Why is forecast validation important?



- Researchers can compare the ex-post performance of their newly developed forecast methods to others

→ **Forecast validation helps to improve methodology**

- Policy-makers rely on forecasts for policy decisions, e.g.:
  - COVID-19 deaths and infections → lockdowns, distancing rules
  - Life expectancy → Age of retirement
  - Population age structure → social security systems
  - Births → child care and education

→ **Forecast validation helps to choose forecasts that are likely to be accurate and unbiased**

Forecast Validation: Error Measures

- Forecast error: $FE_t = \boldsymbol{F_t} - \boldsymbol{Y_t}$
- Percentage error: $PE_t = 100 * \frac{\boldsymbol{F_t} - \boldsymbol{Y_t}}{Y_t}$
- Forecast feature and example measure:
  - Accuracy (closeness to observed values), e.g. PE
  - Bias (under- or overestimation of observed values), e.g. ME
  - Uncertainty (assessment of the probability distribution), e.g. Prediction Interval Coverage
- Characteristics, e.g.:
  - Absolute vs. signed measures (APE vs. PE)
  - Crude vs. relative measures (FE vs. PE)
  - Mean vs. median mesaures (MAE vs. MdE)
  - Yearly (or other time unit) vs. summary measures (FE vs. ME)

t	Year
F	Forecast value in year t
Y	Observed value in year t
n	Number of forecast years

Error Measure **Examples**	Formula
Mean Error	$ME = \underset{t=1,n}{\text{mean}}(FE_t)$
Median Error	$MdE = \underset{t=1,n}{\text{median}}(FE_t)$
Mean Absolute Error	$MAE = \underset{t=1,n}{\text{mean}}(\lvert FE_t \rvert)$
Absolute Percentage Error	$APE_t = \lvert PE_t \rvert$
Mean Absolute Percentage Error	$MAPE = \underset{t=1,n}{\text{mean}}(APE_t)$

Question for you!

Can you think of advantages / disadvantages of the different kinds of error measures? Regarding:

- Absolute vs. signed measures
- Crude vs. relative measures
- Mean vs. median measures

- **Crude** measures: easy interpretation of measure (same scale as forecast value, e.g. years)
- **Relative** measures: comparability of forecasts between countries
- **Absolute** measures: for assessment of forecast accuracy
- **Signed** measures: for assessment of forecast bias
- **Median** measures are less sensitive to outliers compared to **mean** measures

See e.g.: Shcherbakov, M, Brebels, A, Shcherbakova, N, Tyukov, A, Janovsky, T & Kamaev, V (2013). A survey of forecast error measures. World Applied Sciences Journal, 24(24). https://www.researchgate.net/profile/Adriaan-Brebels/publication/281718517_A_survey_of_forecast_error_measures/links/56f43b2408ae81582bf0a1a9/A-survey-of-forecast-error-measures.pdf

Our Validation Framework

- **Project title:**

Validation and Mortality Forecasting: Establishing Method Validation Prior to Forecasting

- **Authors:**

Ricarda Duerst, Christina Bohk-Ewald

- **Research question:**

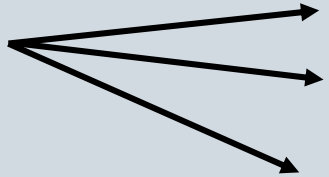
How can forecast validation serve as a test to decide whether a method is suitable to forecast mortality in a country of interest?

- **Research Design:**

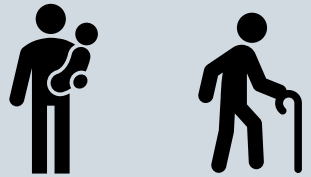
Develepment of a validation framework & exemplary application to Lee-Carter forecasts

Our Validation Framework:
Study Design

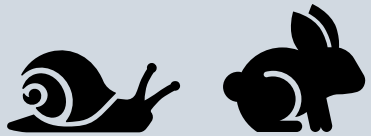
Different
mortality measures



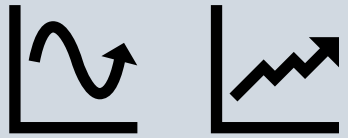
In different
mortality settings



Mean level &
distinct age patterns



Level of mortality
development

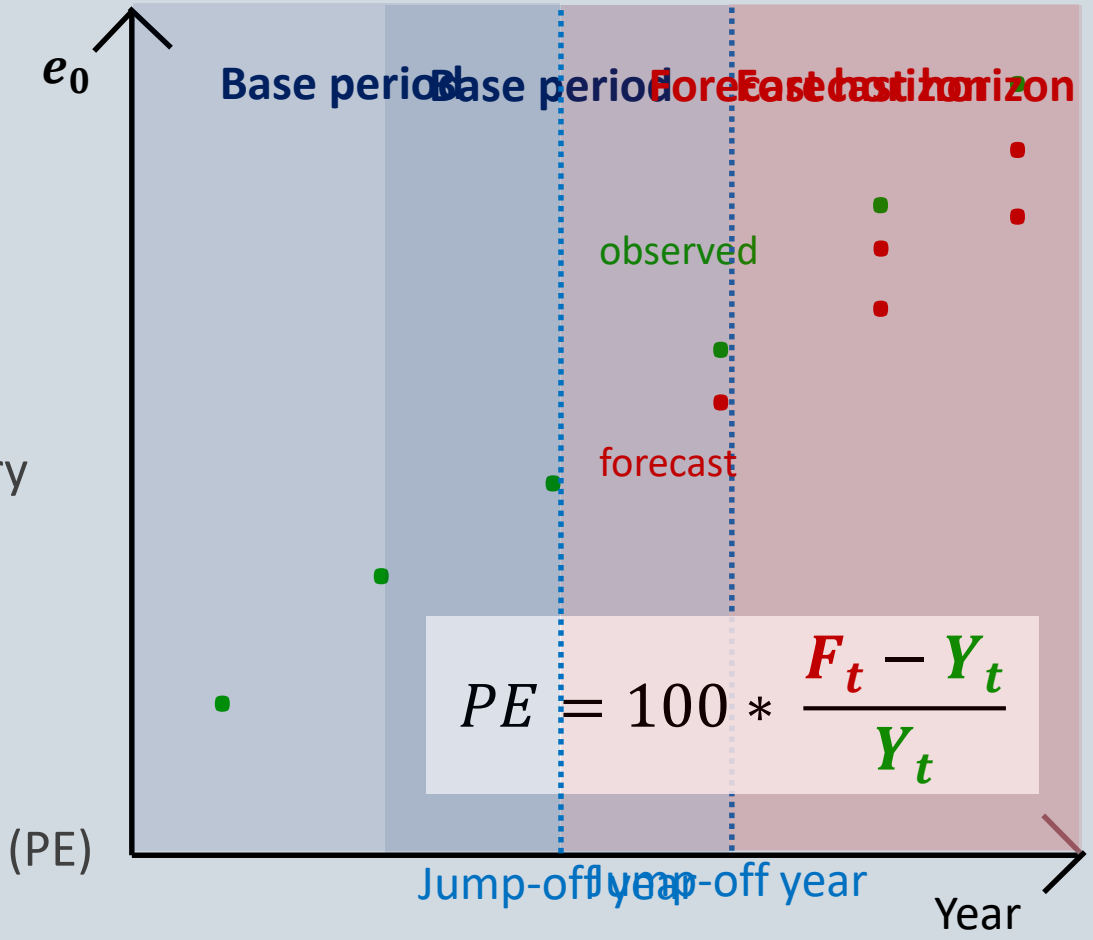


Patterns in mortality
development

→ Large data basis & detailed validation framework

Our Validation Framework:
Making Most of the Data

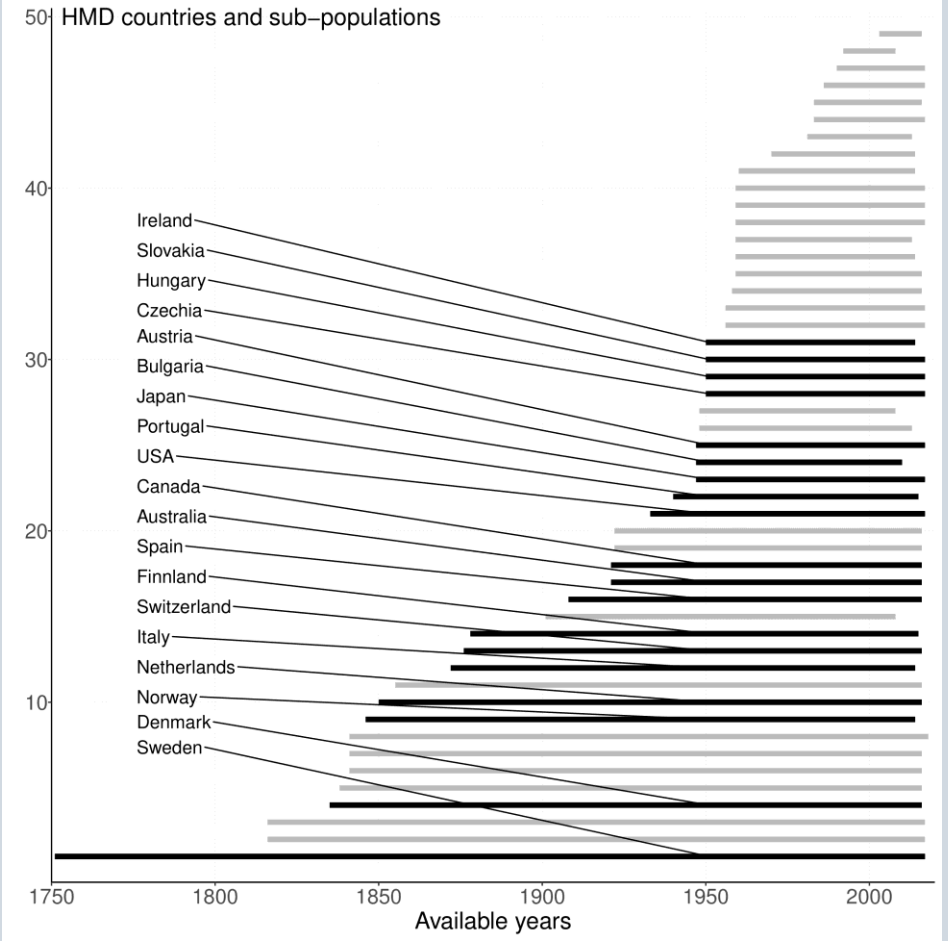
- Out-of-sample validation
- Forecast horizon: 1 to 30 years
- Base period: always 30 years
- As much jump-off years as data for each country allows
- Forecast of m_x → e_0 and $e_0^{\dagger}$
- Accuracy & bias with forecast percentage error (PE)



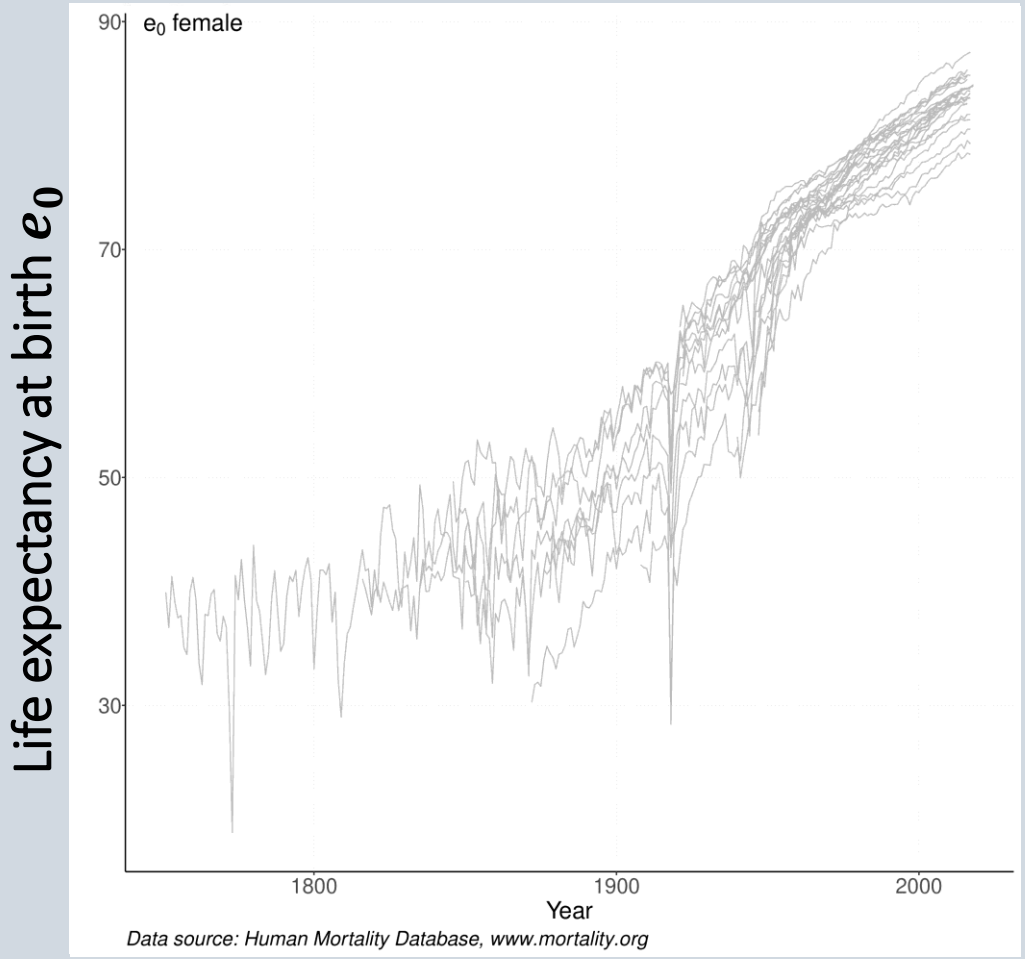
Our Validation Framework:
Data Source HMD

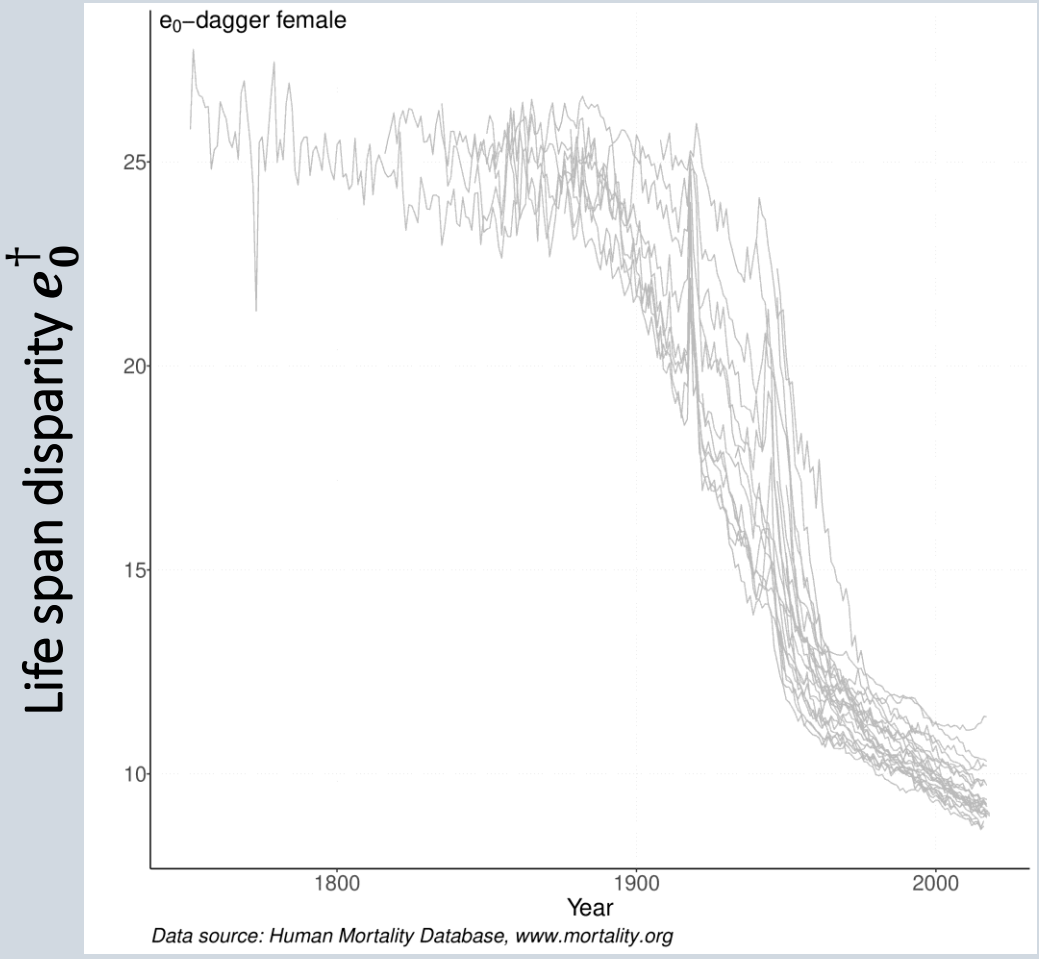
Human Mortality Database:

- 41 countries available
- m_x from life tables
- Ages 0 to 100+
- separately by sex
- Restricted to 24 highly developed countries from Europe, North America & Japan



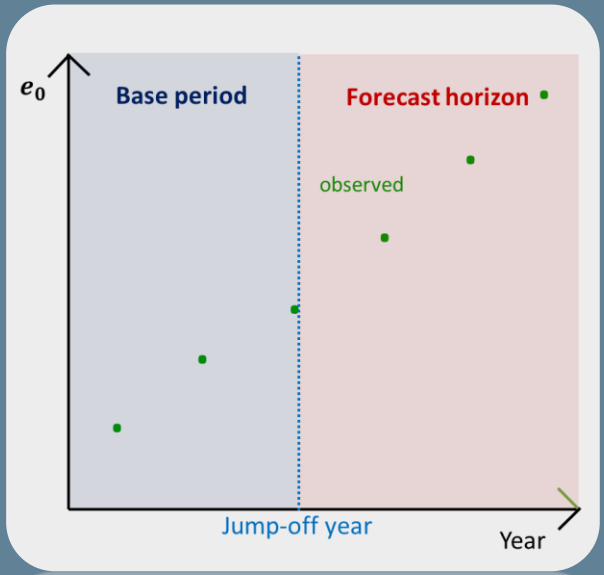
Mortality Development HMD Data

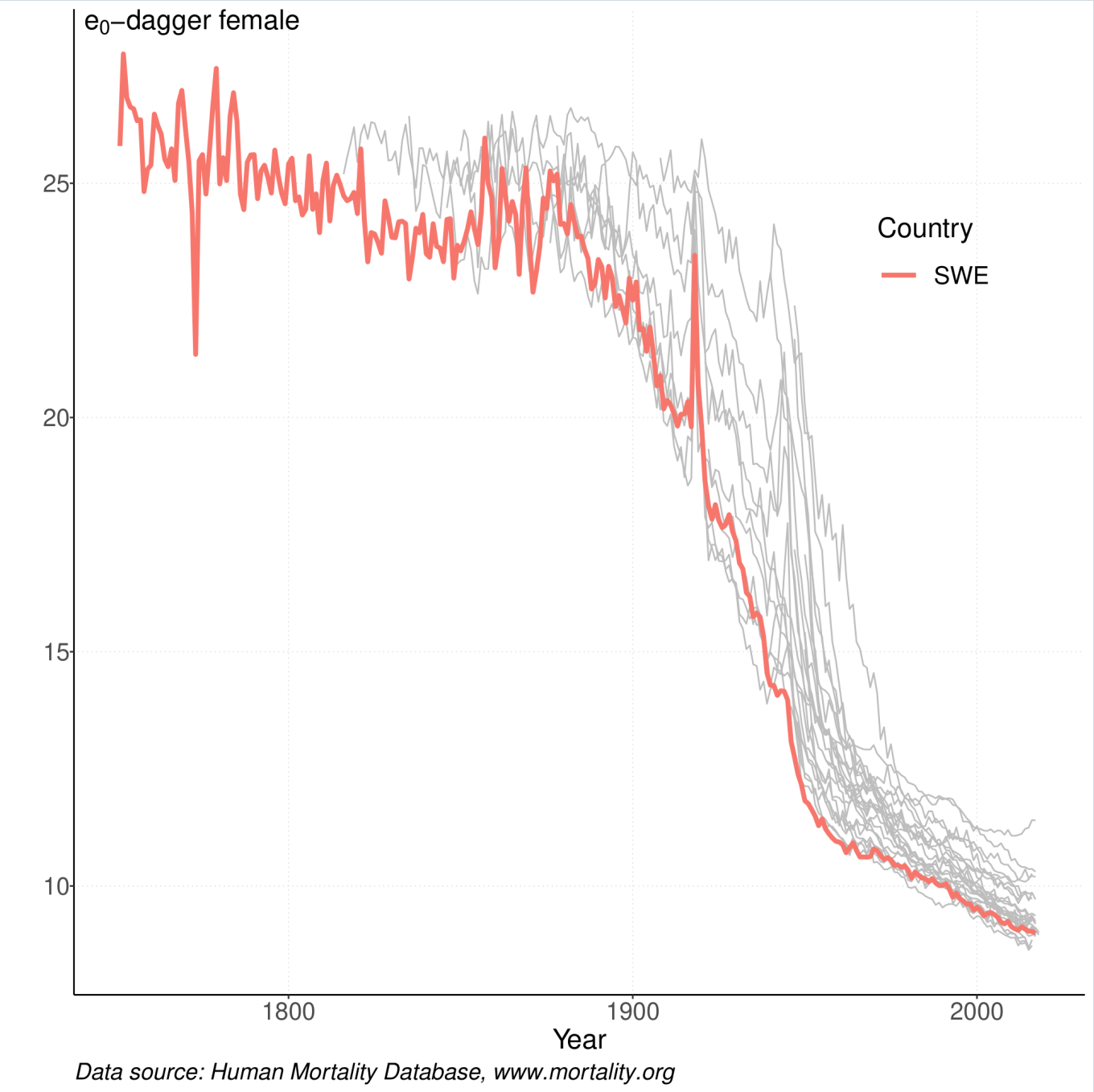




Question for you!

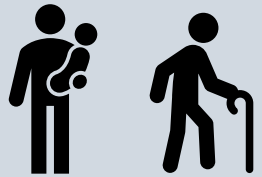
When is it especially difficult or easy for the Lee-Carter model to produce accurate and unbiased forecasts?

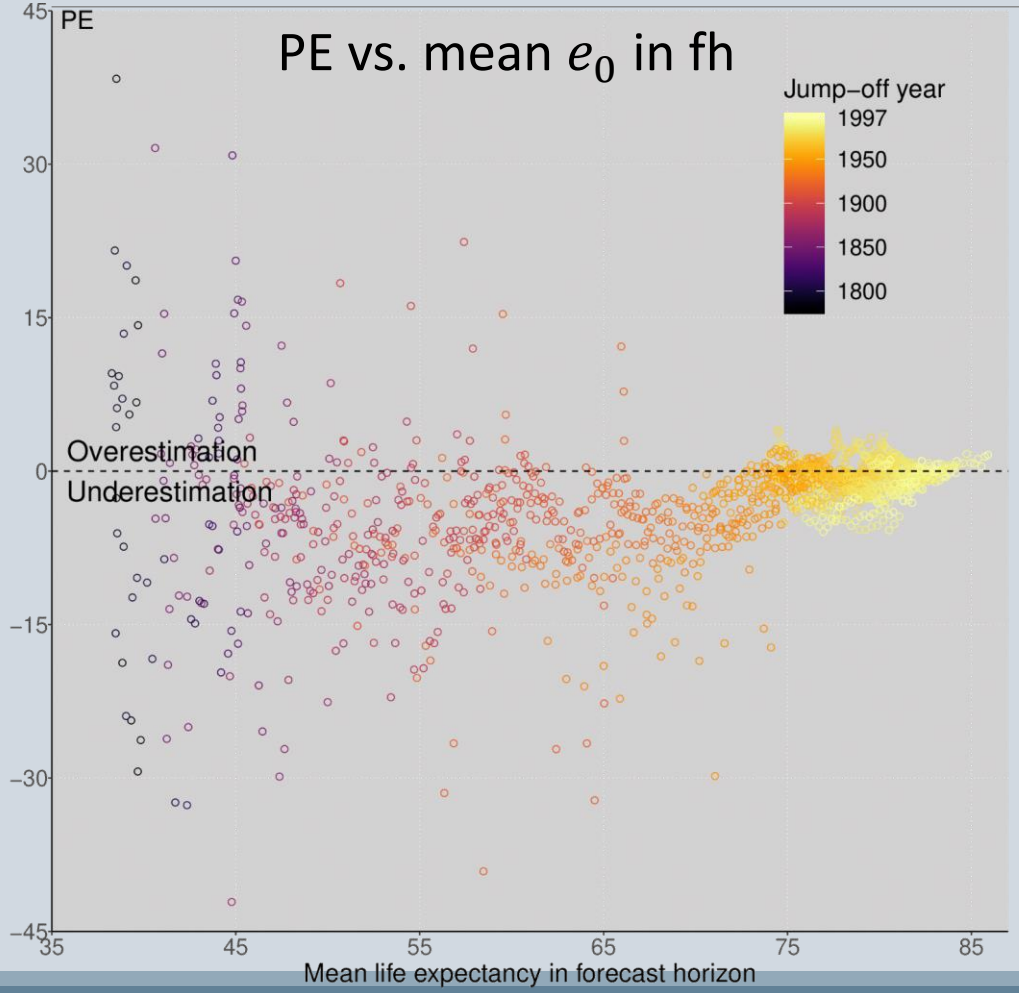




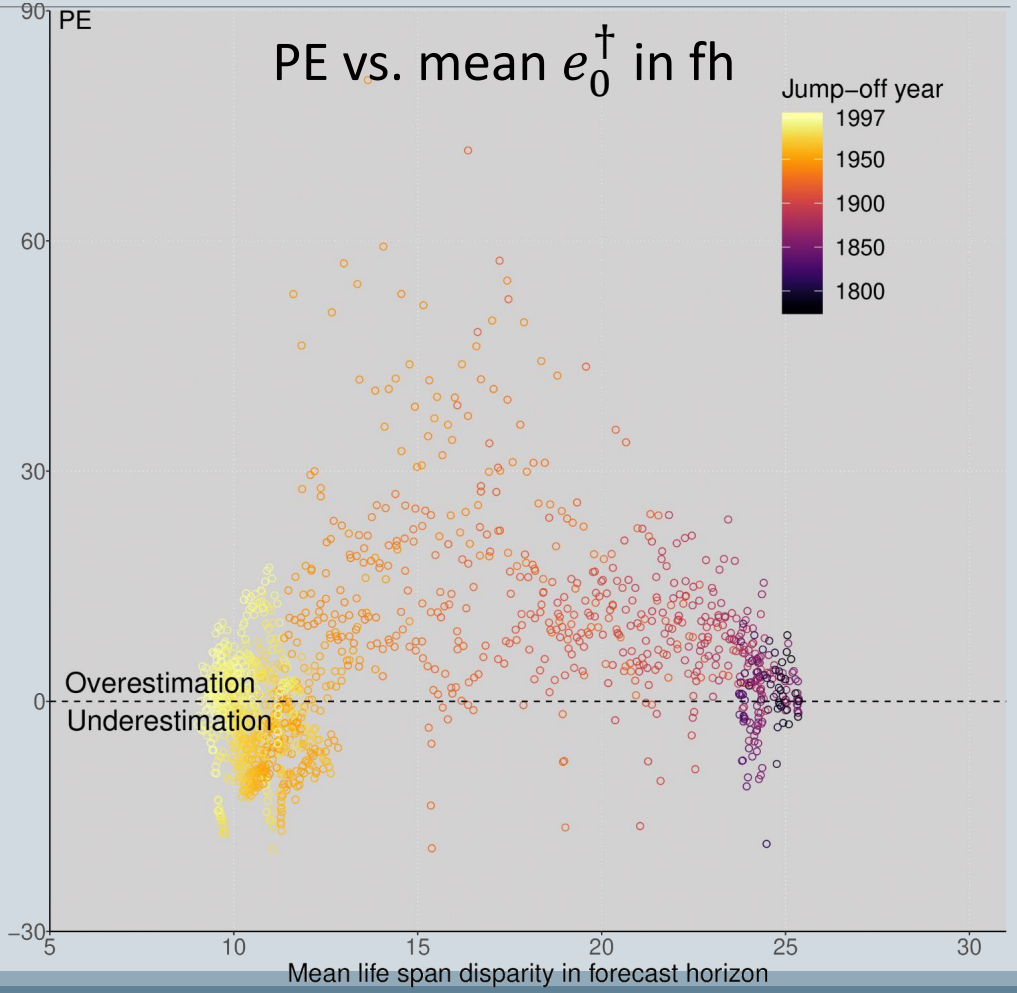
Data source: Human Mortality Database, www.mortality.org

Our Validation Framework:
Results for 1st Analytical Setting





Data source: Human Mortality Database, www.mortality.org

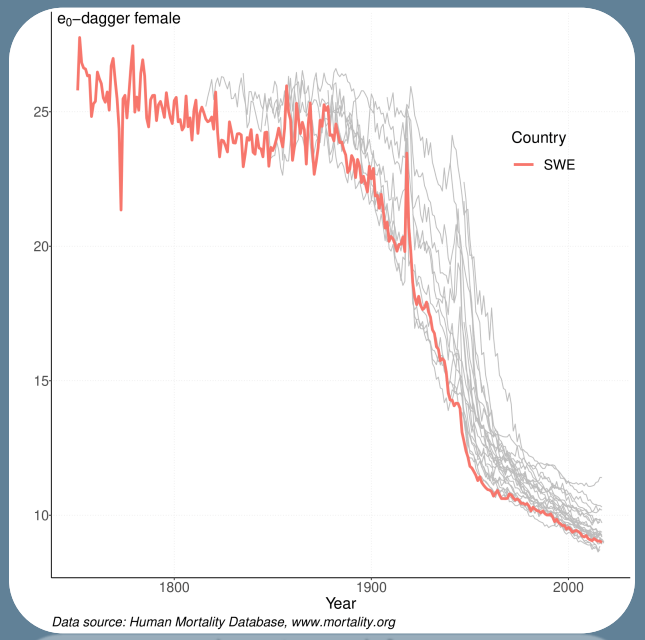


Data source: Human Mortality Database, www.mortality.org

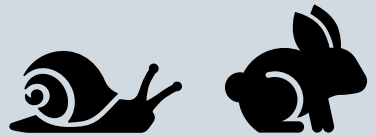
20 years Lee-Carter forecasts, female

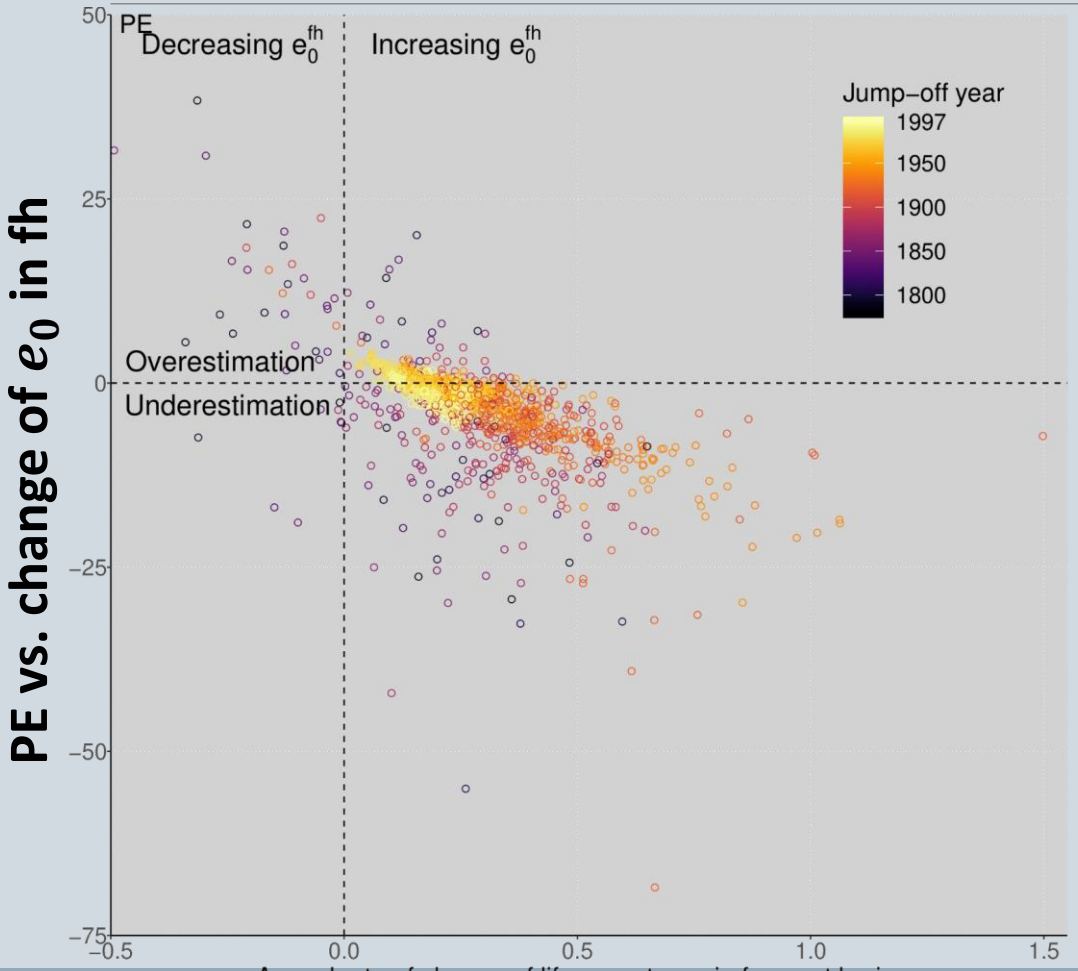
Question for you!

How would you interpret this plot?

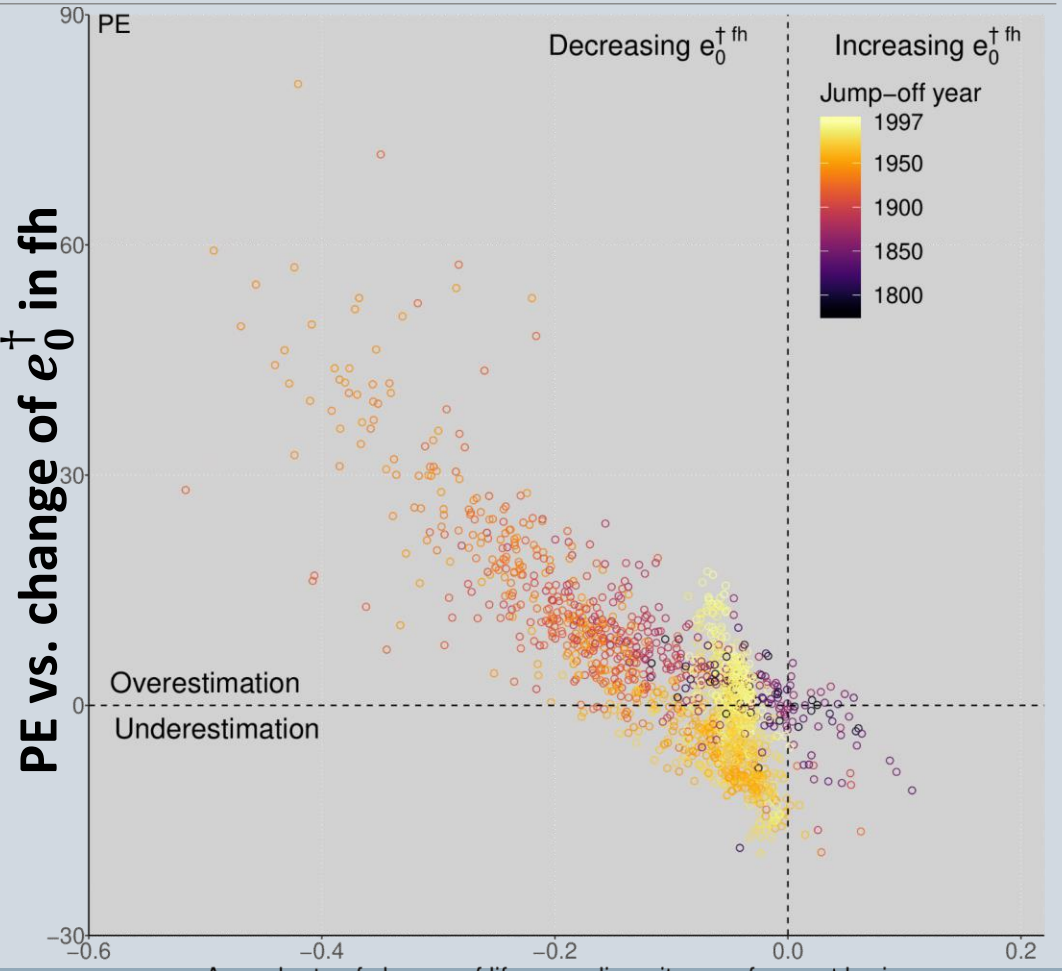


Our Validation Framework:
Results for 2nd Analytical Setting





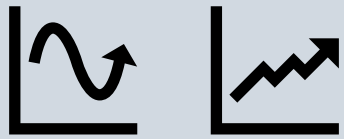
Data source: Human Mortality Database, www.mortality.org

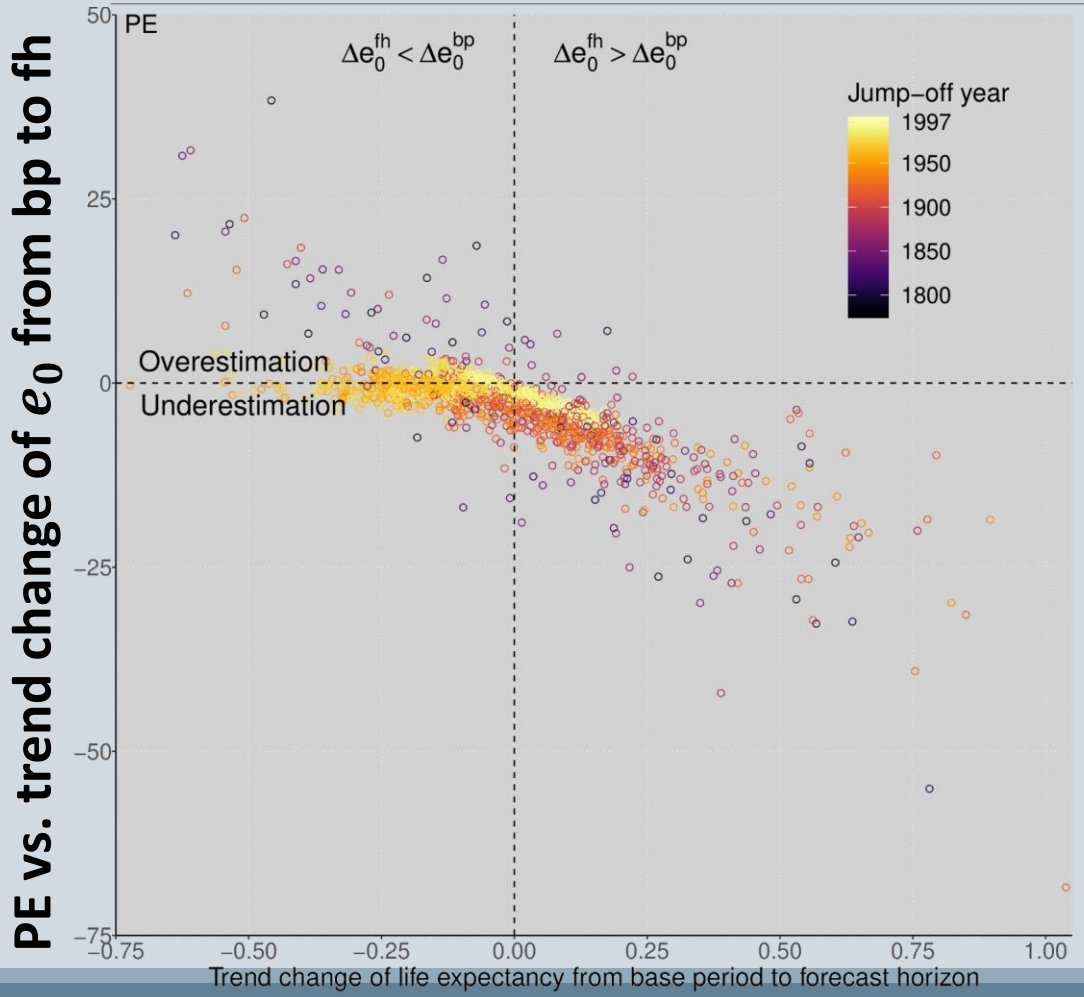


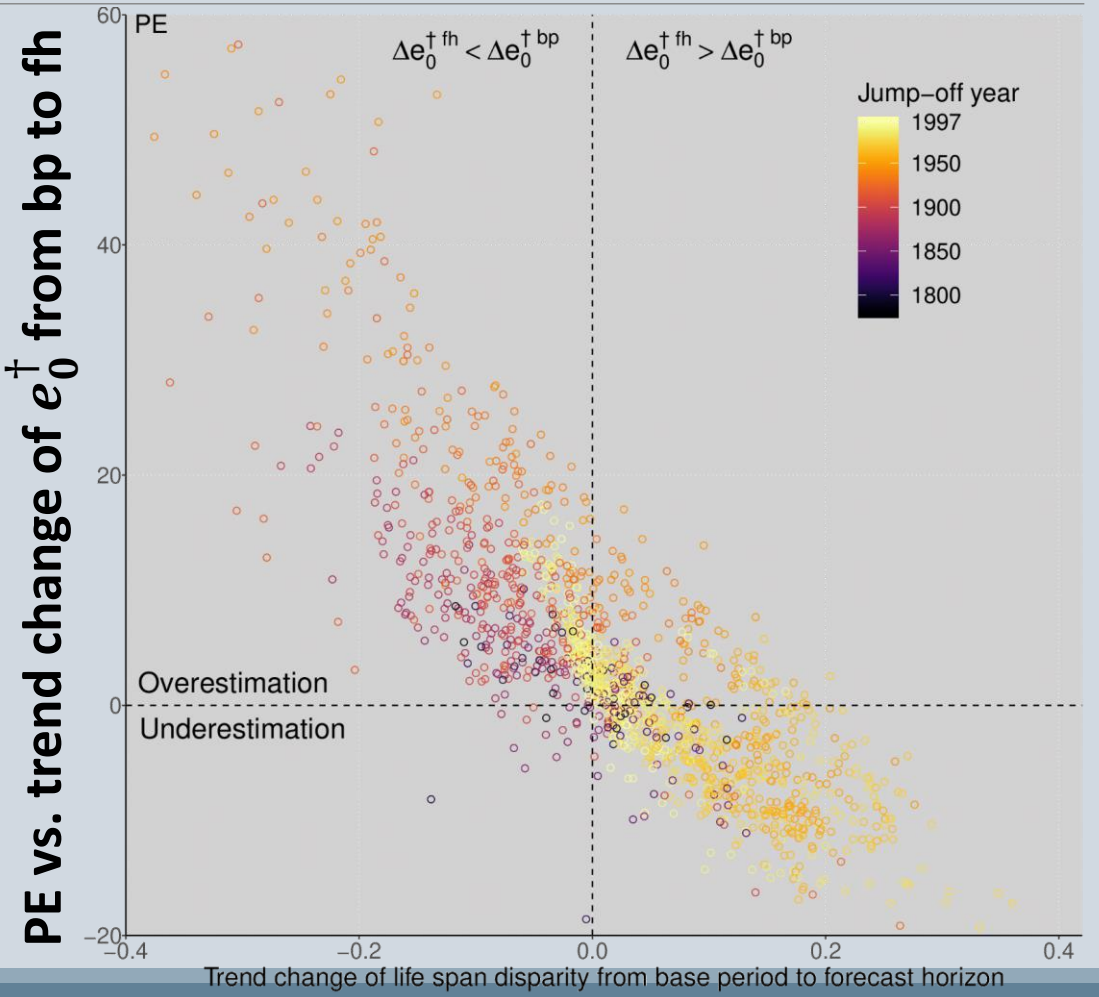
Data source: Human Mortality Database, www.mortality.org

20 years Lee-Carter forecasts, female

Our Validation Framework:
Results for 3rd Analytical Setting







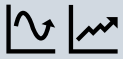
20 years Lee-Carter forecasts, female

Our Validation Design:
Conclusions

Regarding Lee-Carter forecasts:

- mean level of e0 appears has smaller effect on the PE of Lee-Carter forecasts than the mean level of e0†

- strong effect of the annual rate of change of e0 and e0† on accuracy and bias of the Lee-Carter forecasts



- trend changes have a strong impact on the PE of e0 and e0†

- similar patterns for women and men

→ LC suitable to forecast mortality of highly developed countries **in the most recent years**

Regarding our research question:

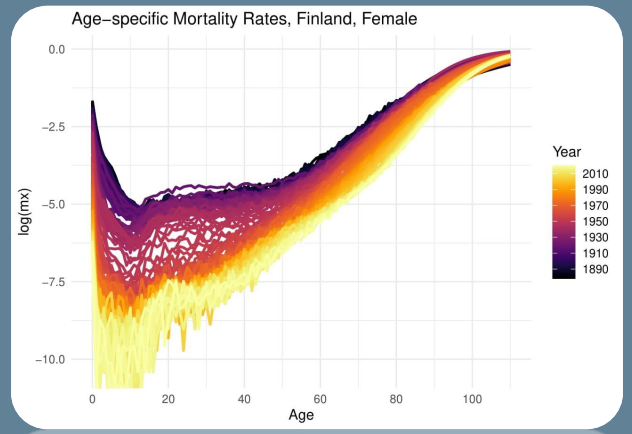
Validation serves as meaningful first test to decide whether a mortality forecast method is appropriate in a country of interest

What You Have Learned Today

- Measures of human mortality: age-specific mortality, life expectancy at birth, and life span disparity
- Development of Finnish mortality
- Data sources for human mortality
- The Lee-Carter model as an example of mortality forecasting
- The principle of ex-post forecast validation
- Example for the application of mortality forecast validation

Wednesday's Lab Session

Hands-on exercises on mortality forecast validation



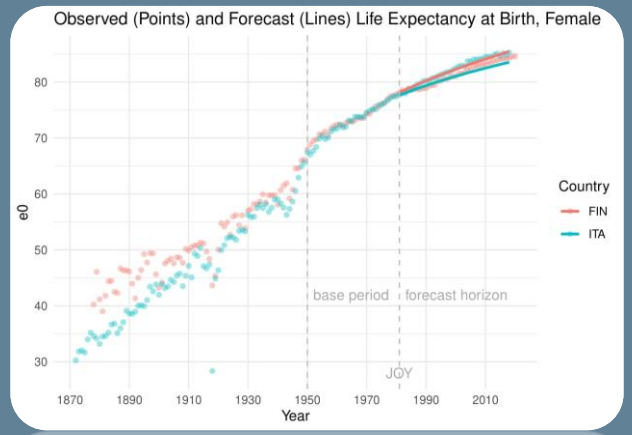
!! PREPARATIONS FOR WEDNESDAY!!

1. Please set up a free account for the Human Mortality Data Base in advance!
   - Go to www.mortality.org
   - Click on „New User“ on the left side
   - Create a login id with your e-mail address
   - Wait for the e-mail with your password

2. Please install the necessary R packages in advance!
   - Open an empty R script in R Studio
   - Copy paste the code on the following slide
   - Run the code
   - Watch the console! If R asks you to enter a number to update some packages, enter „3“ and hit „Enter“ in the console.
   - If you have any problems installing the packages, please contact me via e-mail: duerst@demogr.mpg.de

Wednesday‘s Lab Session

Hands-on exercises on mortality forecast validation



!! PREPARATIONS FOR WEDNESDAY!!

```
# install packages
install.packages(c("fda", "HMDHFDplus", "ggplot2", "viridis", "dplyr", "devtools",
"tidyr", "tibble", "MortalityLaws"), repos = "http://cran.us.r-project.org")

# load libraries
library(ggplot2)
library(HMDHFDplus)
library(fda)
library(dplyr)
library(viridis)
library(devtools)
library(tidyr)
library(tibble)
library(MortalityLaws)
devtools::install_github("mpascariu/MortalityForecast")
library(MortalityForecast)
```

You can‘t get it to work?
Let‘s figure it out together!
Tuesday 3:00pm Helsinki time:
https://zoom.us/j/98537126723?pwd=VjRXeElRVVgyNHlxUGtjZDBRUGYwUT09

Useful Resources

Data

Human Mortality Data Base: https://www.mortality.org/

UN World Population Prospects: https://population.un.org/wpp/

Learning material

- Demography and human mortality:

Preston, S, Heuveline, P & Guillot, M (2000). Demography: Measuring and Modeling Population Processes. MA: Blackwell Publishers.

- Human mortality & forecasting:

Oeppen, J & Vaupel, JW (2002). Broken limits to life expectancy. *Science*, 296:5570. DOI: 10.1126/science.1069675

Booth, H., & Tickle, L. (2008). Mortality modelling and forecasting: A review of methods. *Annals of actuarial science*, *3*(1-2), 3-43. https://www.actuaries.org/CTTEES_TFM/Documents/Booth_Tickle.pdf

- Lee-Carter model:

Lee, R & Carter, L (1992). Modeling and Forecasting U.S. Mortality. *Journal of the American Statistical Association, 87:419*. https://doi.org/10.1080/01621459.1992.10475265

- Our validation design:

Duerst, R & Bohk-Ewald, C. Validation and Mortality Forecasting: Establishing Method Validation Prior to Forecasting. *Annual Meeting of the Population Association of America*, May 5-8 2021, online. https://submissions2.mirasmart.com/PAA2021/ViewSubmissionFile.aspx?sbmID=2720&mode=html&validate=false

- Error measures:

Shcherbakov, M, Brebels, A, Shcherbakova, N, Tyukov, A, Janovsky, T & Kamaev, V (2013). A survey of forecast error measures. *World Applied Sciences Journal*, *24*(24). https://www.researchgate.net/profile/Adriaan-Brebels/publication/281718517_A_survey_of_forecast_error_measures/links/56f43b2408ae81582bf0a1a9/A-survey-of-forecast-error-measures.pdf