

Tobias Göllner
Statistics Austria

Den Haag
14 September 2018

Estimating Differential Mortality from EU-SILC UDB Longitudinal Data

- The FACTAGE project
- Differential Mortality
- EU-SILC
- R pipeline

- Fairer Active Ageing for Europe factage.eu
- First JPI-MYBL Call – BE,DE,ES,UK,AT
- one work package dedicated to Differential Mortality Estimations
- Literature overview by our FACTAGE colleagues Mosquera et al. (2018)

- Harmonized target variables
- All EU countries + (e.g.) Norway, Switzerland,...
- Available to researchers as User Database (UDB), with some restrictions to ensure anonymity
- Four filetypes (D, H, P and R)
- Released every year
- Cross-sectional and Longitudinal (4 y)
- Deterministic linkage (via ID, year, country)

2008 : D

Household A – 2008
Household B – 2006
Household B – 2007
Household B – 2008

2008 : H

Household A – 2008
Household B – 2006
Household B – 2007
Household B – 2008

2008 : R

Person A – 2008
Person B – 2006
Person B – 2007
Person B – 2008

2008 : P

Person A – 2008
Person B – 2006
Person B – 2007
Person B – 2008

2009 : D

Household A – 2008
Household A – 2009
Household B – 2006
Household B – 2007
Household B – 2008
Household B – 2009

2009 : H

Household A – 2008
Household A – 2009
Household B – 2006
Household B – 2007
Household B – 2008
Household B – 2009

2009 : R

Person A – 2008
Person A – 2009
Person B – 2006
Person B – 2007
Person B – 2008
Person B – 2009

2009 : P

Person A – 2008
Person A – 2009
Person B – 2006
Person B – 2007
Person B – 2008
Person B – 2009

2010 : D

Household A – 2008
Household A – 2009
Household A – 2010

2010 : H

Household A – 2008
Household A – 2009
Household A – 2010

2010 : R

Person A – 2008
Person A – 2009
Person A – 2010

2010 : P

Person A – 2008
Person A – 2009
Person A – 2010

ID	Country	Year	Obs start	Obs end	Died	Var of H/P	Verweildauer
Person A	AT	2008	15/05/2008	15/02/2011	0	<i>income</i>	3.75

- Create one line for each respondent, from multiple file types and versions
- Calculate a time at risk variable (“Verweildauer”)
- Know the vital status of the person
- Have a multitude of analytical variables (SES) for differential mortality estimations

- SAS code is in “user-friendly” macros
- User base limited to SAS license holders
- Translation is tricky... Things that are easy in SAS are hard in R and vice versa
- Increase user base and (potentially) performance

```
create.SIILCdata <-  
function(path, year_from, year_to,  
vars, countries=NULL) {...}
```

Raw code to only use the most recent observation, fixing some issues, create the date variables

```
first.last.merge <- function  
(first.vars=NULL, last.vars=NULL)  
{...}
```

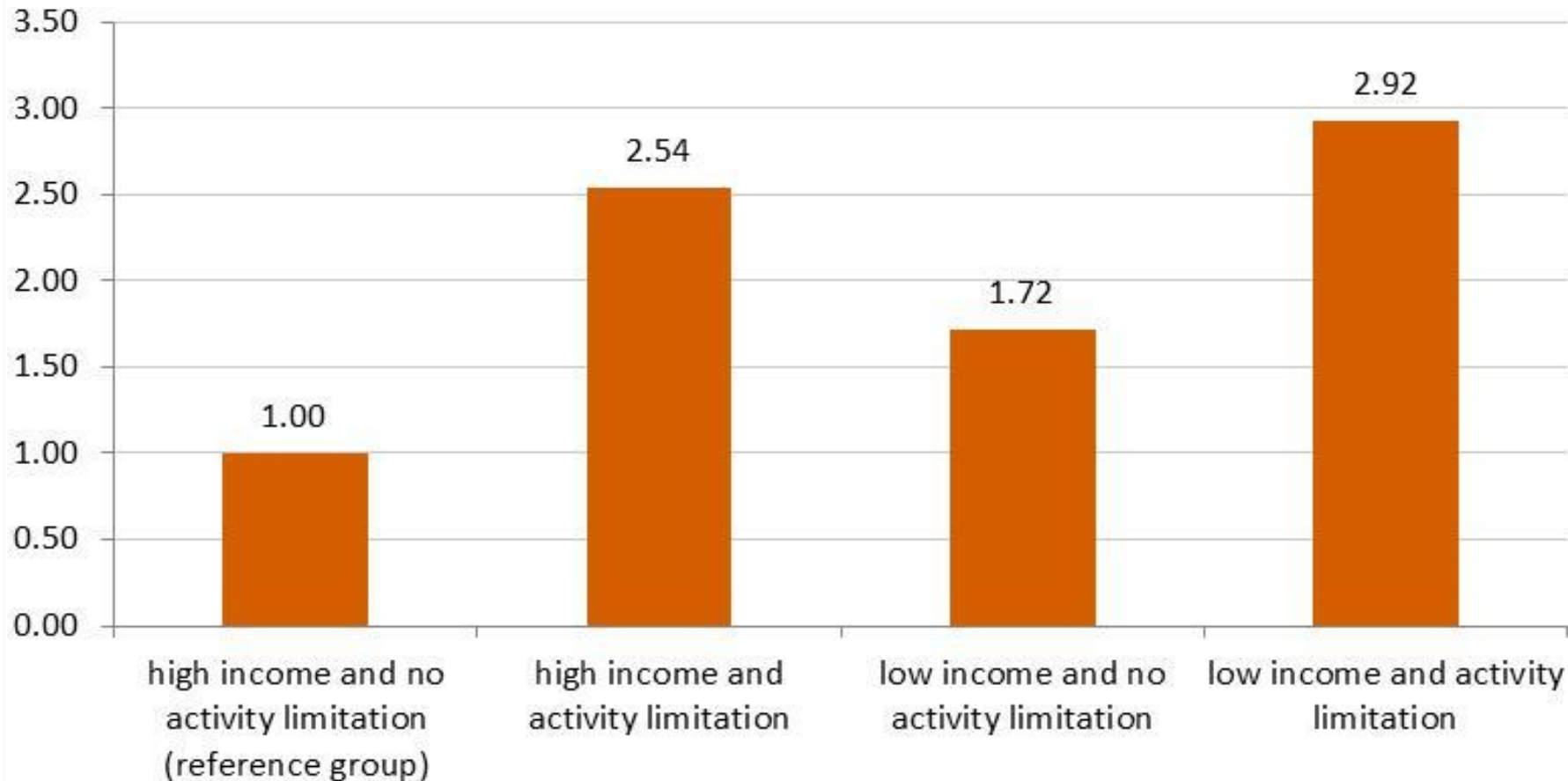

Data cleaning to only use eligible cases (16-79 yrs old, country specific stuff)

Write a csv file for differential mortality estimations

Create a model estimation using Cox Proportional Hazards Regression

Exemplary finding of FACTAGE

Age-adjusted mortality risk of Polish males



- Few packages (tidyr, dplyr, data.table, survival)
- Big potential for improvement
- Right now probably 4x slower than SAS

[https://github.com/TobiasGold/
FACTAGE-method Mortality/
0-and-1 Load-and-Apply v0-7-0.R](https://github.com/TobiasGold/FACTAGE-method_Mortality/0-and-1_Load-and-Apply_v0-7-0.R)

- Improve R code
- Use the code
- Spread the code

Potentially one-day course at GESIS
Mannheim @ 6th European User Conference
Microdata from Eurostat

Tobias Göllner
Johannes Klotz
Statistics Austria

tobias.goellner@statistik.gv.at

Thank You!