



MAX-PLANCK-INSTITUT
FÜR DEMOGRAFISCHE
FORSCHUNG

MAX PLANCK INSTITUTE
FOR DEMOGRAPHIC
RESEARCH

Working with Demographic Life Table Data in Stata

Daniel C. Schneider

Max Planck Institute for Demographic Research

Stata Conference, July 27, 2017, Baltimore



Contents

Demographic Life Tables

The Human Mortality Database

hmddata

text to .dta conversion

data usage and examples

life expectancy: Oeppen / Vaupel (2002)

mortality rates

lifetable



The Life Table

- displays death-related statistics of a cohort/population
- columns: age and age-related functions pertaining to mortality
- cohort life table vs. period life table: "synthetic cohort"
- based on triangles from a Lexis diagram
- calculation of life expectancy
- related: `ltable` of official Stata



Life Table of the US, 2014

```
. hmddata use lifetables bothsexes, clear grid(5x1) popfilter(usa)
. list age mx-ex if year==2014, noobs sep(0)
```

age	mx	qx	ax	lx	dx	Lx	Tx	ex
0	0.0060	0.0059	0.06	100,000	592	99,447	7,897,283	79.0
1	0.0002	0.0010	1.64	99,408	96	397,407	7,797,837	78.4
5	0.0001	0.0006	2.41	99,312	57	496,414	7,400,429	74.5
10	0.0001	0.0007	2.82	99,255	69	496,125	6,904,015	69.6
15	0.0004	0.0023	2.98	99,186	224	495,476	6,407,890	64.6
20	0.0008	0.0042	2.60	98,962	415	493,810	5,912,414	59.7

[...]

65	0.0147	0.0710	2.62	84,222	5,983	406,867	1,644,162	19.5
70	0.0228	0.1080	2.62	78,239	8,450	371,069	1,237,295	15.8
75	0.0361	0.1662	2.62	69,789	11,596	321,393	866,226	12.4
80	0.0600	0.2621	2.59	58,193	15,252	254,268	544,833	9.4
85	0.1023	0.4080	2.52	42,941	17,522	171,309	290,565	6.8
90	0.1785	0.6061	2.35	25,420	15,408	86,325	119,256	4.7
95	0.2801	0.7737	2.11	10,011	7,746	27,655	32,931	3.3
100	0.4170	0.8948	1.81	2,266	2,027	4,861	5,276	2.3
105	0.5695	0.9554	1.52	238	228	400	415	1.7
110	0.6923	1.0000	1.44	11	11	15	15	1.4



Life Expectancy

- life expectancy (LE, e_x):
 - e_x : average years ahead of a population member aged X
 - e_0 (LE at birth):
 - average years lived
= mean age at death
- e_0 of period life table: average years lived *under current (period) mortality conditions*
- All statements are made with respect to members of a *hypothetical cohort*.



The Human Mortality Database (HMD)

- compiled by: UC Berkeley, Max Planck Institute for Demographic Research
- high-quality data
- variables: see next slide
- geographic coverage:
 - currently 39 countries / populations
 - many European countries, plus: US, Canada; Japan, Taiwan; Australia; Chile; Israel; Russia
- time coverage: Sweden 1750-, France 1816-, 10 other countries start before 1900
- www.mortality.org
- companion / similar databases: HLD, HFD, HFC



HMD: Data Contents

. hmddata info concepts

HMD full concepts:

<u>births</u>	<u>period</u> <u>raw</u>
<u>deaths</u>	<u>period</u> <u>raw</u>
<u>deathsbylexistriangles</u>	<u>period</u>
<u>populationsize</u>	<u>period</u> <u>raw</u>
<u>exposuretorisk</u>	<u>period</u> <u>cohort</u>
<u>deathrates</u>	<u>period</u> <u>cohort</u>
<u>lifetables</u>	<u>period</u> <u>cohort</u>
<u>lifeexpectancyatbirth</u>	<u>period</u> <u>cohort</u>



HMD: Data Acquisition

- consent to user agreement and registration required, but free of charge
- Data are distributed in text files.
- zipped text files
(http://www.mortality.org/cgi-bin/hmd/hmd_download.php)
 - available
 - by statistic / concept
 - by country
 - all data
 - **hmddata** can process any one and one or more of the zipped text files.



- development goal was a data exploration tool for researchers:
 - easy data access
 - quick generation of working-quality tables and graphs
- ```
net install hmddata ,
 from(http://user.demogr.mpg.de/schneider/stata)
```



# hmddata : text to .dta conversion

## Syntax

Set and query **hmddata** user settings

```
hmddata settings [parameter] , [value(valstring)]
```

Convert source data text files to hmddata files

```
hmddata convert fullconceptsspec , sourcedir(dirstring) [grid(gridlist) replace]
```

Load HMD data

```
hmddata use fullconceptsspec , [popfilter(poplistsspec) grid(gridsspec) long clear]
```



# hmddata : data handling

Generate age and year interval variables

```
hmddata intervals [intvalvars] , [noorder]
```

Filter data set according to a subset of populations

```
hmddata popfilter poplist , [iso noerror
droplist dummy(varname)]
```

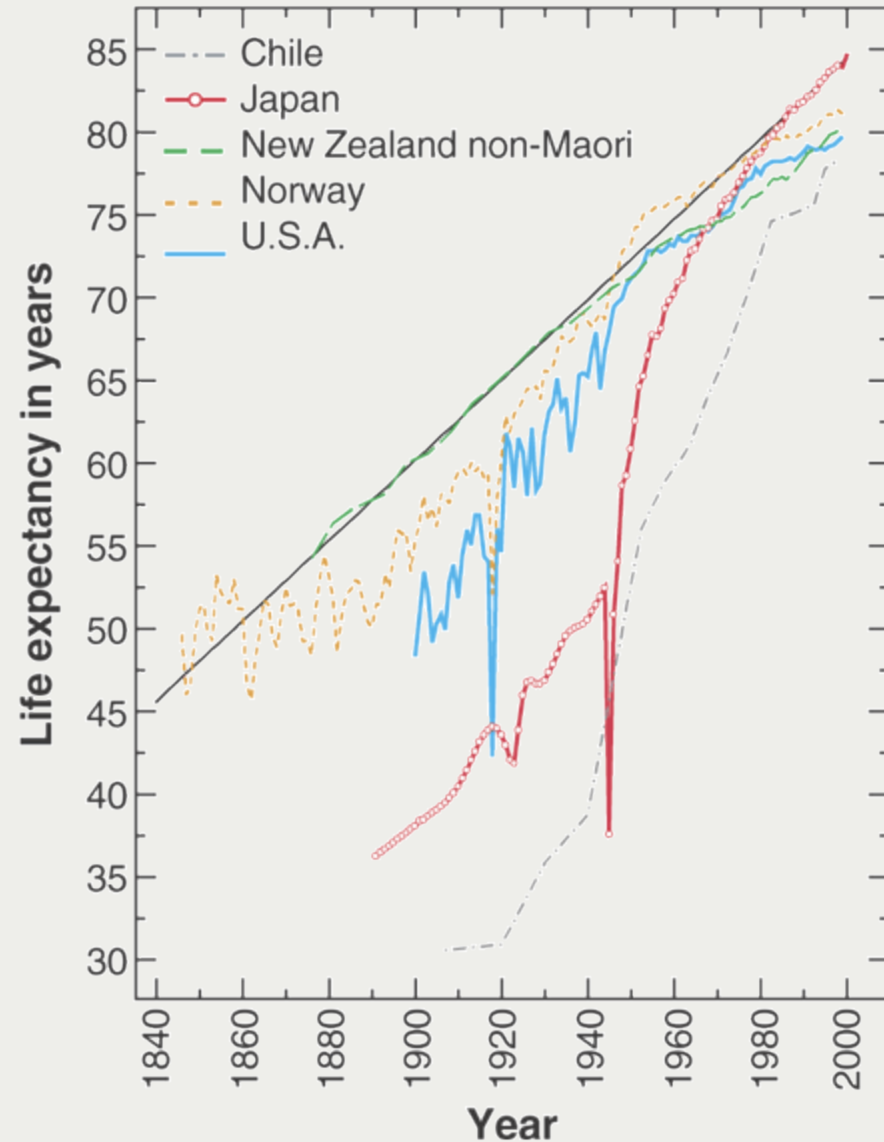
Generate graphs based on hmd data sets

```
hmddata graph plottype plotvars xvar [if] [in]
 , [at1(atspec) at2(atspec) by(varlist [,
byopts]) plotopts(cline_options)
twoway_options]
```



# Oeppen / Vaupel (2002)

Original graph from paper: Female life expectancy (LE) for selected countries and trend in record LE.

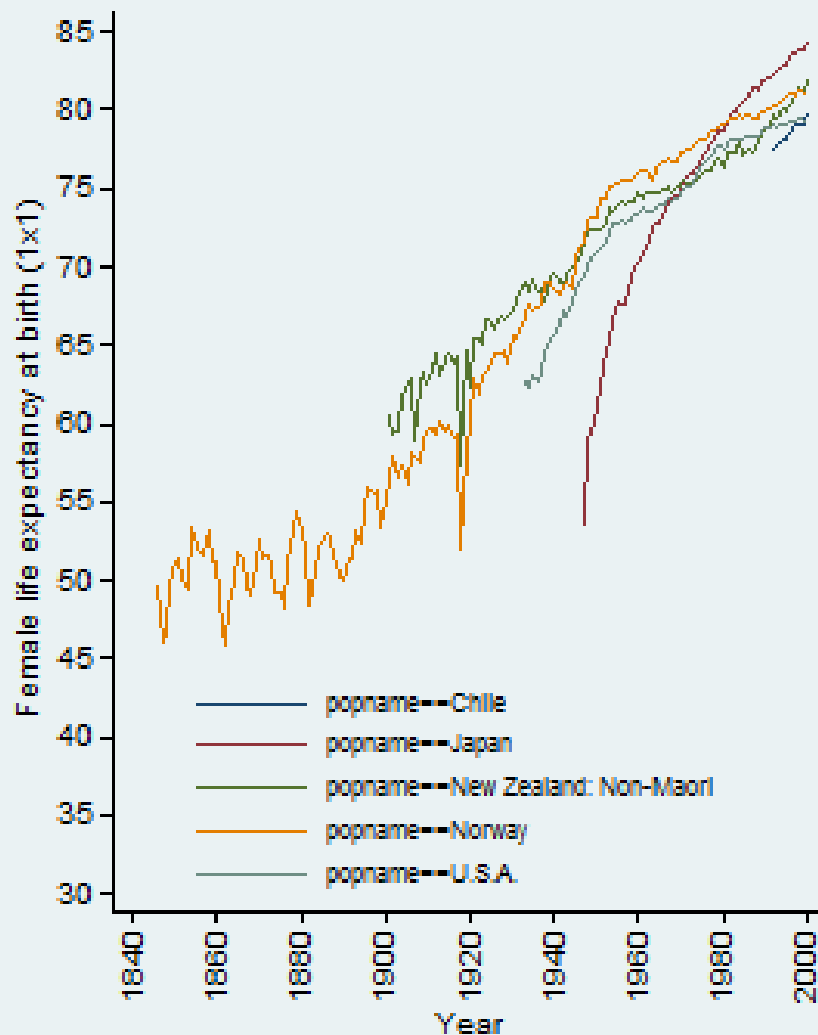




# Oeppen / Vaupel (2002)

Graph replication using `hmddata`:

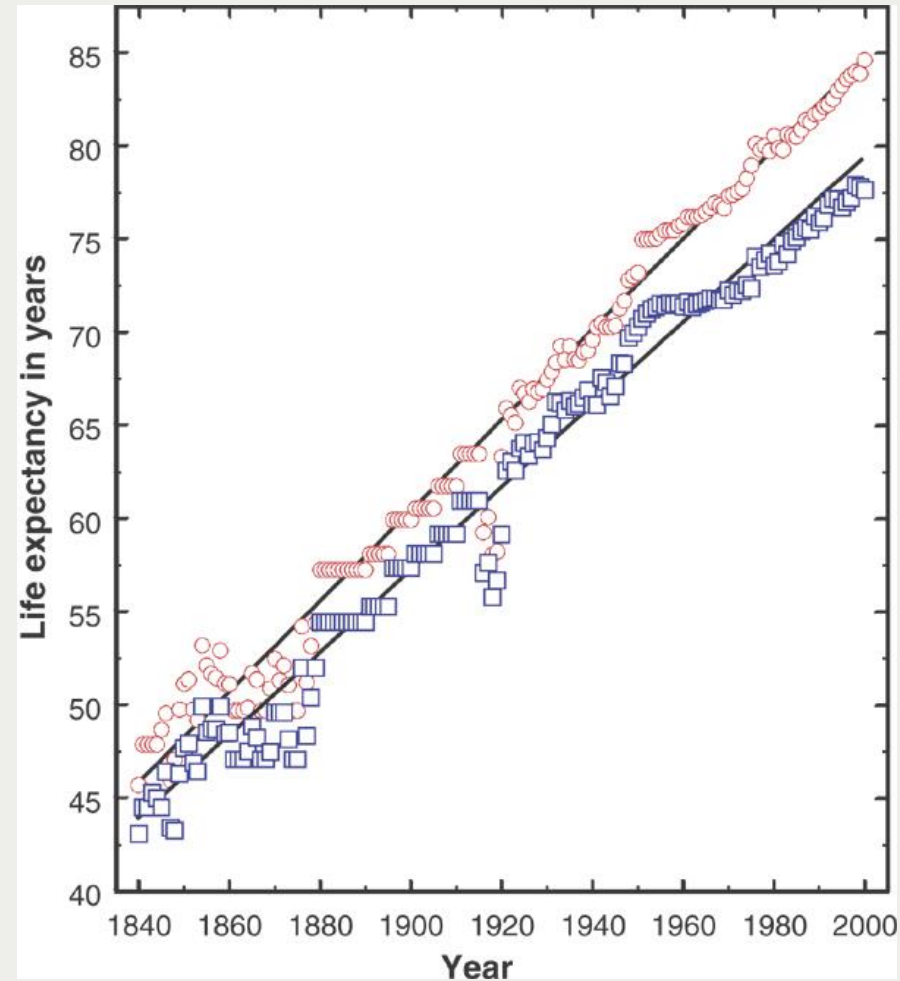
```
. hmddata use lifeexp , clear
. hmddata graph line female year
 if inrange(year, 1840, 2000),
 at1(popname chile japan
 newzealandnon nor usa) [...]
```





# Oeppen / Vaupel (2002)

Original graph from paper: Male (blue) and female (red) LE in the record-holding country.



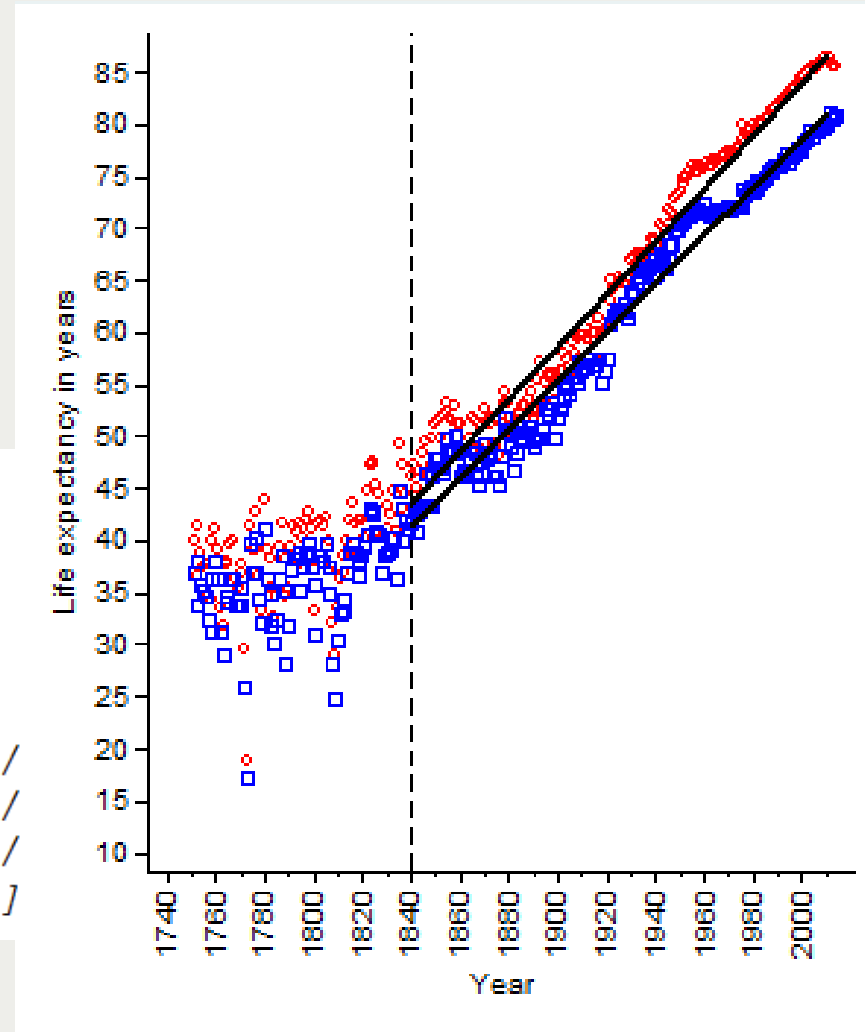


# Oeppen / Vaupel (2002)

- replication of graph plus additional history
- More complicated graphs: use `graph twoway` instead of `hmddata graph`.

```
hmddata use lifeexp , clear
hmddata popfilter francecivil [...]
[...] // generate LE rank variables
 // for males/females

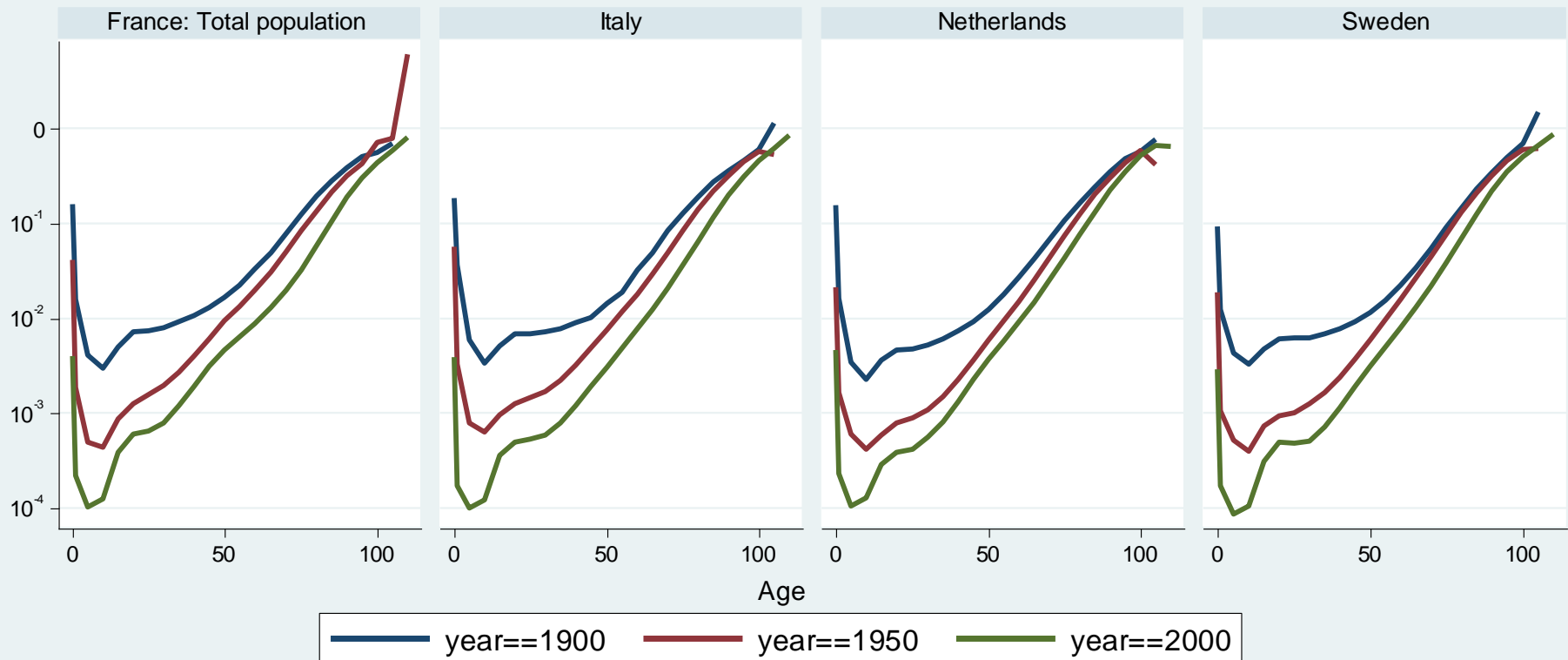
twoway ///
(scatter fem year if rank_f==1, [...]) || ///
(scatter mal year if rank_m==1, [...]) || ///
(lfit fem year if rank_f==1 [...]) || ///
(lfit mal year if rank_m==1 [...]), [...]
```





# Mortality Declines 1900-1949, 1950-2000

- . hmddata use deathrates , grid (5x10) clear
- . hmddata popfilter swe francetotal neth ita , dummy(d1)
- . replace total = log10(total)
- . hmddata graph line total age if d1 , at1(year 1900 1950 2000) by(popname) [...]



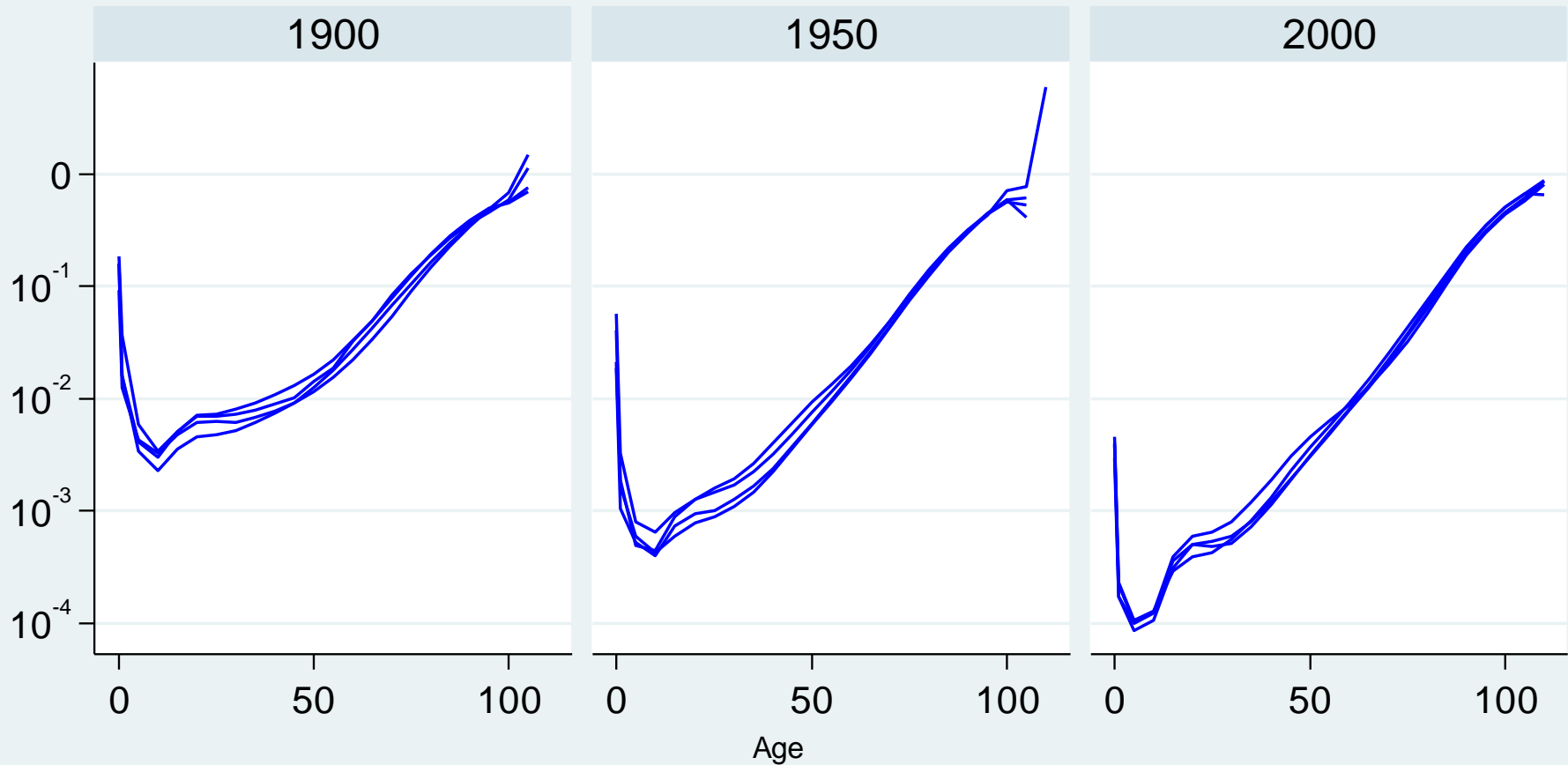
Graphs by Country / Population name





# Mortality Declines 1900-1949, 1950-2000

```
. gen smpl = inlist(year, 1900, 1950, 2000) & d1
. hmddata graph line total age if smpl , at1(popname)
 by(year, [...]) [...]
```



Graphs by Year



- not yet released
- development goal: versatile tool to generate and manipulate demographic life tables
- principles / features:
  - operation on multiple yet selected life tables at once
  - standardized/prescribed variable names
  - calculations using any valid minimum starting information
- to be added before release: CIs, methods for approximating  ${}_n a_x$ , ...



# The End

# Thank you!

Questions? Comments?

contact: [schneider@demogr.mpg.de](mailto:schneider@demogr.mpg.de)



# References

- Human Mortality Database. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at [www.mortality.org](http://www.mortality.org) and [www.humanmortality.de](http://www.humanmortality.de).
- J. Oeppen and James W. Vaupel (2002): Broken Limits to Life Expectancy. *Science*, 5570 (296), pp. 1029-1031.