The Demography of Kinship (3)

Kinship microsimulations

Diego Alburez Gutiérrez[†]

[†]Kinship Inequalities Research Group,
Max Planck Institute for Demographic Research

European Doctoral School of Demography 2022-23
26 April 2023

MAX PLANCK INSTITUTE FOR DEMOGRAPHIC RESEARCH
MAX-PLANCK-INSTITUT FÜR DEMOGRAFISCHE FORSCHUNG

EUROPEAN DOCTORAL SCHOOL OF DEMOGRAPHY

Structure of this talk

1. Principles of demographic microsimulation

2. Implementations

What is a simulation?

In groups, discuss:

1. What is a simulation?
2. What is a microsimulation?
3. What is a demographic microsimulation?
4. What is the difference between a demographic microsimulation and an agent-based model?

Principles of demographic microsimulation

General logic of demographic microsimulations

1. Model individual-level demographic behaviour using set of rules
2. Manufacture individual-level data
3. Simple inputs
4. Different alternatives:
   - SOCSIM
   - CAMSIM
   - R/python
   - Agent-based modelling

Expected number of children[1]

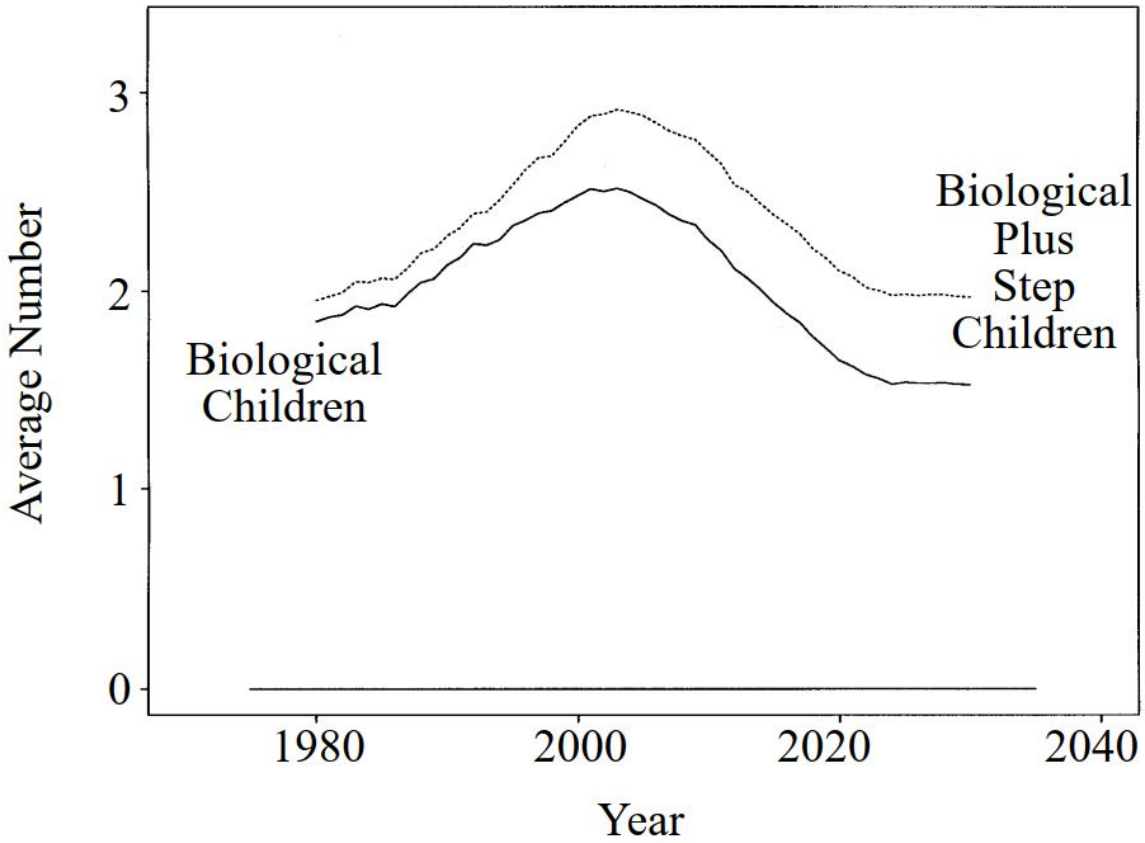


Figure 1. Living biological children and stepchildren, US whites aged 70–85. Outputs of Berkeley SOCSIM simulations, average of 40 replications.

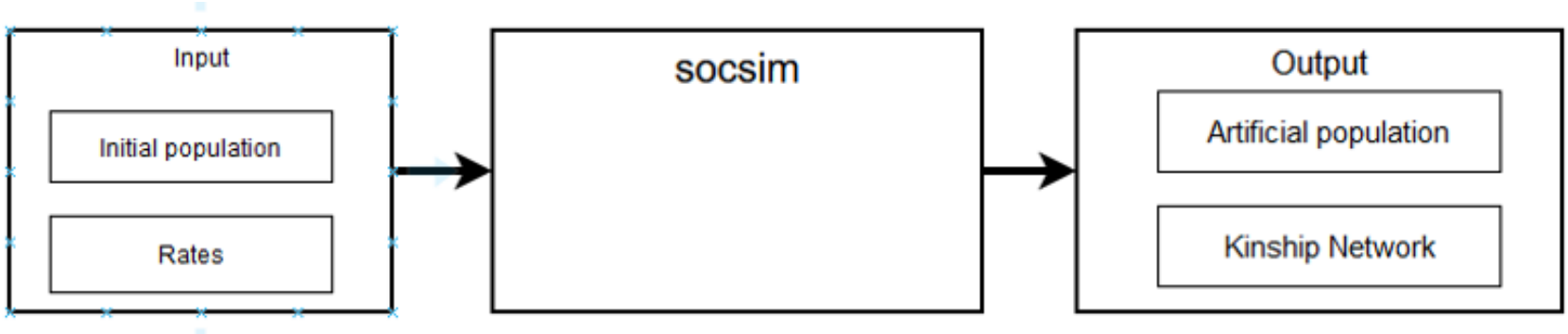
[1] Wachter, K. W. (1997). Kinship resources for the elderly. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *352*(1363), 1811–1817. https://doi.org/10.1098/rstb.1997.0166

Demographic microsimulations with SOCSIM

1. A stochastic microsimulation platform, 1970s at UC Berkeley
2. Starts with initial population
3. Each simulated individual experiences specific rates every month (e.g., mortality, fertility, marriage)
4. Keeps track of kinship ties to create a full genealogy
5. UC Berkeley SOCSIM User Manual[2]

[2] Mason, C. (2016). SOCSIM Oversimplified. UC Berkeley. https://lab.demog.berkeley.edu/socsim/CurrentDocs/socsimOversimplified.pdf

Are we living in a simulation?[3]



[3] This and next slides contributed by Tom Theile.

How does SOCSIM work? (1)

1. Every person is an individual object
2. The population is a long list of persons
3. When a birth happens, a new person is created and appended to that list
4. At the start of the simulation (or after an event), every person gets a "next event"

How does SOCSIM work? (2)

5. Events can be marriage, divorce, childbirth, death, etc.
6. Time increments in discrete timesteps
7. In every time step, all events scheduled for this time "happen"
8. A simulation can consist of 1 or more "segments", every segment can have different rates
9. At the end of the simulation, socsim writes the population into output files

Attributes of simulated individuals (1)

Every SOCSIM-person has the following parameters:

1. **status**: dead/alive
2. **sex** (female or male)
3. **marital status** (single, married, divorced, cohabiting, widowed)
4. **parity** (total number of children born to a woman)

Attributes of simulated individuals (2)

5. **group** (number between 0 and 63. Can be used to simulate different groups/countries/towns.... with transition rates between groups)
6. **age** (in months, changes with time steps, starts at 0. Max age is 200 years)
7. **next event**

Possible events

1. **birth/childbirth**: creation of a new person with age 0 and parameters that are random (sex) or derived (marital status=single at birth, . . .)
2. **death**: according to mortality rates. These are specific to parameters of the individual persons (age, gender, groups, parity, marital status)
3. **marriage**: the most complicated event, because it involves 2 persons. Three marriage market systems:
    1. Marriage rates for both males and females
    2. Marriage rates only for females - males just get picked
    3. No marriage rates; Marriage just before a childbirth to an unmarried mother
4. **divorce**

Running a SOCSIM simulation

```
##  init_new.opop|init_new.omar|init_new.opox|sim_results_socsim.sup_300_/result.pyr|sim_results
## fix pop pointers..
## Starting month is 601
## Initial size of pop 8000  (living: 8000)
## ------------aa3s------------aa32New events generated for all living persons
## ------------b1month:  700 PopLive:  9414 Brths:  16 Dths:   0 Mrgs: 11 Dvs:  0 Mq: 3728 Fq:0
## month:  800 PopLive: 10926 Brths:  12 Dths:   1 Mrgs:  6 Dvs:  0 Mq: 3890 Fq:0 ti1: 0.5 ti2:
## month:  900 PopLive: 12260 Brths:  14 Dths:   0 Mrgs:  4 Dvs:  0 Mq: 4031 Fq:0 ti1: 0.7 ti2:
## month: 1000 PopLive: 13397 Brths:   9 Dths:   2 Mrgs:  4 Dvs:  0 Mq: 4134 Fq:0 ti1: 0.8 ti2:
## month: 1100 PopLive: 14172 Brths:  16 Dths:   6 Mrgs:  6 Dvs:  0 Mq: 4135 Fq:0 ti1: 1.0 ti2:
## month: 1200 PopLive: 14518 Brths:  13 Dths:  11 Mrgs:  6 Dvs:  0 Mq: 4000 Fq:0 ti1: 1.2 ti2:
## month: 1300 PopLive: 14323 Brths:  14 Dths:  20 Mrgs:  4 Dvs:  0 Mq: 3891 Fq:0 ti1: 1.4 ti2:
## month: 1400 PopLive: 13816 Brths:  13 Dths:  15 Mrgs:  4 Dvs:  0 Mq: 3746 Fq:0 ti1: 1.6 ti2:
## month: 1500 PopLive: 13330 Brths:  11 Dths:  11 Mrgs:  5 Dvs:  0 Mq: 3679 Fq:0 ti1: 1.8 ti2:
## month: 1600 PopLive: 12944 Brths:  10 Dths:  15 Mrgs:  4 Dvs:  0 Mq: 3593 Fq:0 ti1: 2.0 ti2:
## month: 1700 PopLive: 12525 Brths:  10 Dths:  20 Mrgs:  5 Dvs:  0 Mq: 3436 Fq:0 ti1: 2.1 ti2:
## month: 1800 PopLive: 12009 Brths:  10 Dths:  16 Mrgs:  7 Dvs:  0 Mq: 3275 Fq:0 ti1: 2.3 ti2:
##
##
## Socsim Main Done
## Socsim Done.
```

At the end of a simulation

1. Virtual population written to output files
    1. **opop**: every person that ever lived and their attributes
    2. **omar**: every marriage and their attributes
2. text files, can then be read and analysed in R

SOCSIM output objects

```
## Read the opop file using the read_opop function
opop <- rsocsim::read_opop(folder = getwd(), supfile = "socsim_USA.sup",
                           seed = "120423", suffix = "",  fn = NULL)
```

```
## [1] "read population file: /Users/mallikasnyder/Desktop/rsocsim_workshop_paa/sim_results_socsim_USA
```

```
## Read the omar file using the read_opop function
omar <- rsocsim::read_omar(folder = getwd(), supfile = "socsim_USA.sup",
                           seed = "120423", suffix = "",  fn = NULL)
```

```
## [1] "read marriage file: /Users/mallikasnyder/Desktop/rsocsim_workshop_paa/sim_results_socsim_USA.s
```

SOCSIM output objects

```
head(opop)
```

```
##   pid fem group nev  dob mom pop nesibm nesibp lborn marid mstat  dod    fmult
## 1   1   1     1  65 1009   0   0      0      0 26579  1005     4 1593 1.767931
## 2   2   0     1  65 1179   0   0      0      0     0     0     1 2062 0.000000
## 3   3   1     1  65  956   0   0      0      0 26205   625     4 1795 0.706973
## 4   4   0     1  65  641   0   0      0      0     0     0     1 1349 0.000000
## 5   5   0     1  65 1015   0   0      0      0 25201  2810     3 1938 0.000000
## 6   6   1     1  65  797   0   0      0      0 22412   526     3 1555 1.581888
```

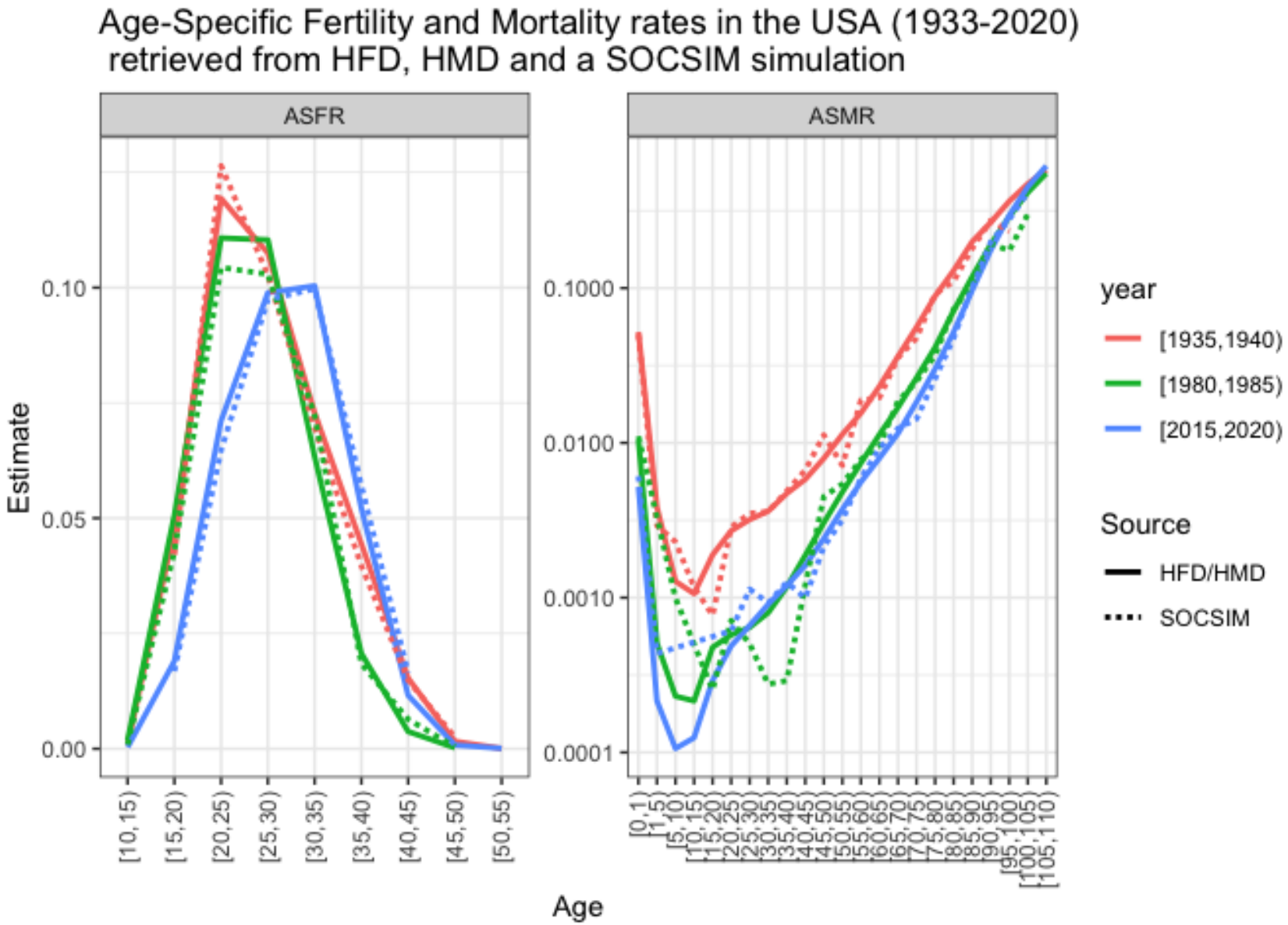
```
head(omar)
```

```
##   mid  wpid  hpid dstart dend rend wprior hprior
## 1   1  4473 11649   1201 1810    3      0      0
## 2   2  3418  1865   1201 1358    3      0      0
## 3   3  3569 11595   1201 1660    3      0      0
## 4   4 17771  3043   1201 1451    3      0      0
## 5   5  5388 17305   1201 1765    3      0      0
## 6   6 11717    33   1201 1402    3      0      0
```

Input data for SOCSIM microsimulation

1. User-provided
    1. Initial population
    2. Age-specific fertility rates
    3. Age-specific mortality rates
2. Optional or default parameters available
    1. Marriage transition rates
    2. Model for marriage market
    3. Other transition rates
    4. Other parameters (inheritance of fertility, etc.)

Discuss: comparison input-output rates[4]



[4] Liliana P. Calderón-Bernal, https://github.com/alburezg/rsocsim_workshop_paa

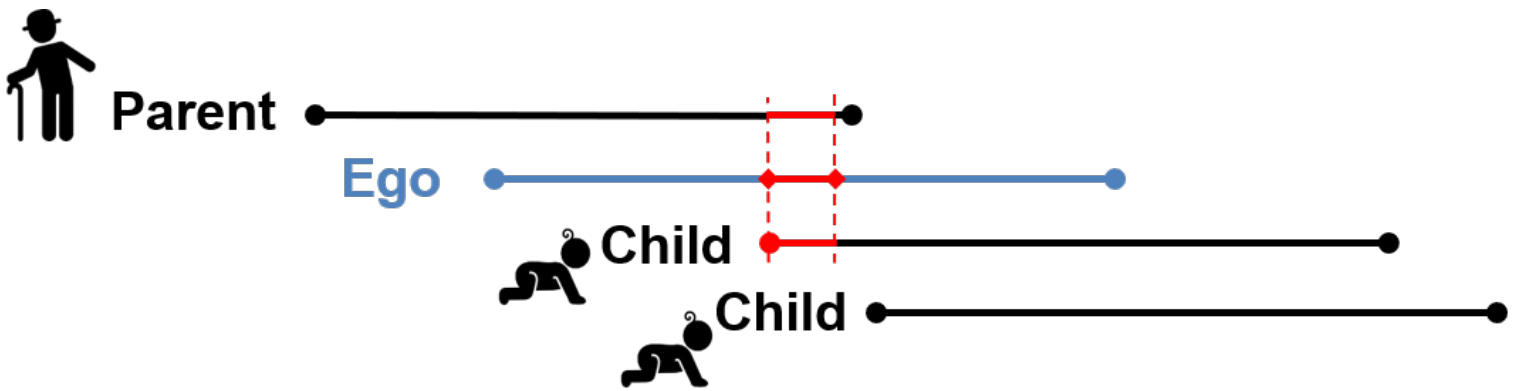
Break

Implementations

Demographic sandwiching

We define a person as 'sandwiched' if they simultaneously have:

1. one *child* younger than 15 years old, and
2. one older *parent or parent in-law* within 5 years from death



Measuring the 'sandwich generation' around the world[5]

Research question

Is the 'sandwich generation' more prevalent among the ageing populations of the Global North?

1. Methods: [redacted]
2. Data: [redacted]
3. Results: unequal geographic distribution of demographic sandwiching

[5] Alburez-Gutierrez, D., Mason, C., & Zagheni, E. (2021). The "Sandwich Generation" Revisited: Global Demographic Drivers of Care Time Demands. *Population and Development Review*, *47*(4), 997–1023

Discuss

1. How would you go about answering this question?
2. Which methods/data would you use?
3. What problems/limitations do you foresee?

Measuring the 'sandwich generation' around the world[6]

Research question

Is the 'sandwich generation' more prevalent among the ageing populations of the Global North?

1. Methods: demographic microsimulation in SOCSIM
2. Data: calibrated against 2019 UNWPP estimates and projections
3. Results: unequal geographic distribution of demographic sandwiching

[6] Alburez-Gutierrez, D., Mason, C., & Zagheni, E. (2021). The "Sandwich Generation" Revisited: Global Demographic Drivers of Care Time Demands. *Population and Development Review*, *47*(4), 997–1023

Anatomy of a demographic sandwich

$S(a)$ is the probability that a woman aged a has a child younger than 15 and a parent within 5 years of death:

$$S(a) = \underbrace{\left(1 - \prod_{x=1}^{15}[1 - m(a-x)]\right)}_{\substack{\text{fertility risk in the} \\ \text{15 years preceding age a}}} \times \underbrace{M_1(a)}_{\substack{\text{Prob. ego's mother} \\ \text{alive}}} \times \underbrace{\left(1 - \frac{M_1(a+5)}{M_1(a)}\right)}_{\substack{\text{Prob. ego's mother} \\ \text{dies in 5 years}}}$$

where the probability of having a living mother in a stable population is:

$$M_1(a) \approx \frac{l(\mu + a)}{l(\mu)}.$$

(1) Longer and healthier life

$S(a)$ is the probability that a woman aged a has a child younger than 15 and a parent within 5 years of death:

$$S(a) = \underbrace{\left(1 - \prod_{x=1}^{15}[1 - m(a-x)]\right)}_{\substack{\text{fertility risk in the} \\ \text{15 years preceding age a}}} \times \underbrace{M_1(a)}_{\substack{\text{Prob. ego's mother} \\ \text{alive}}} \times \underbrace{\left(1 - \frac{M_1(a+5)}{M_1(a)}\right)}_{\substack{\text{Prob. ego's mother} \\ \text{dies in 5 years}}}$$

where the probability of having a living mother in a stable population is:

$$M_1(a) \approx \frac{l(\mu + a)}{l(\mu)}.$$

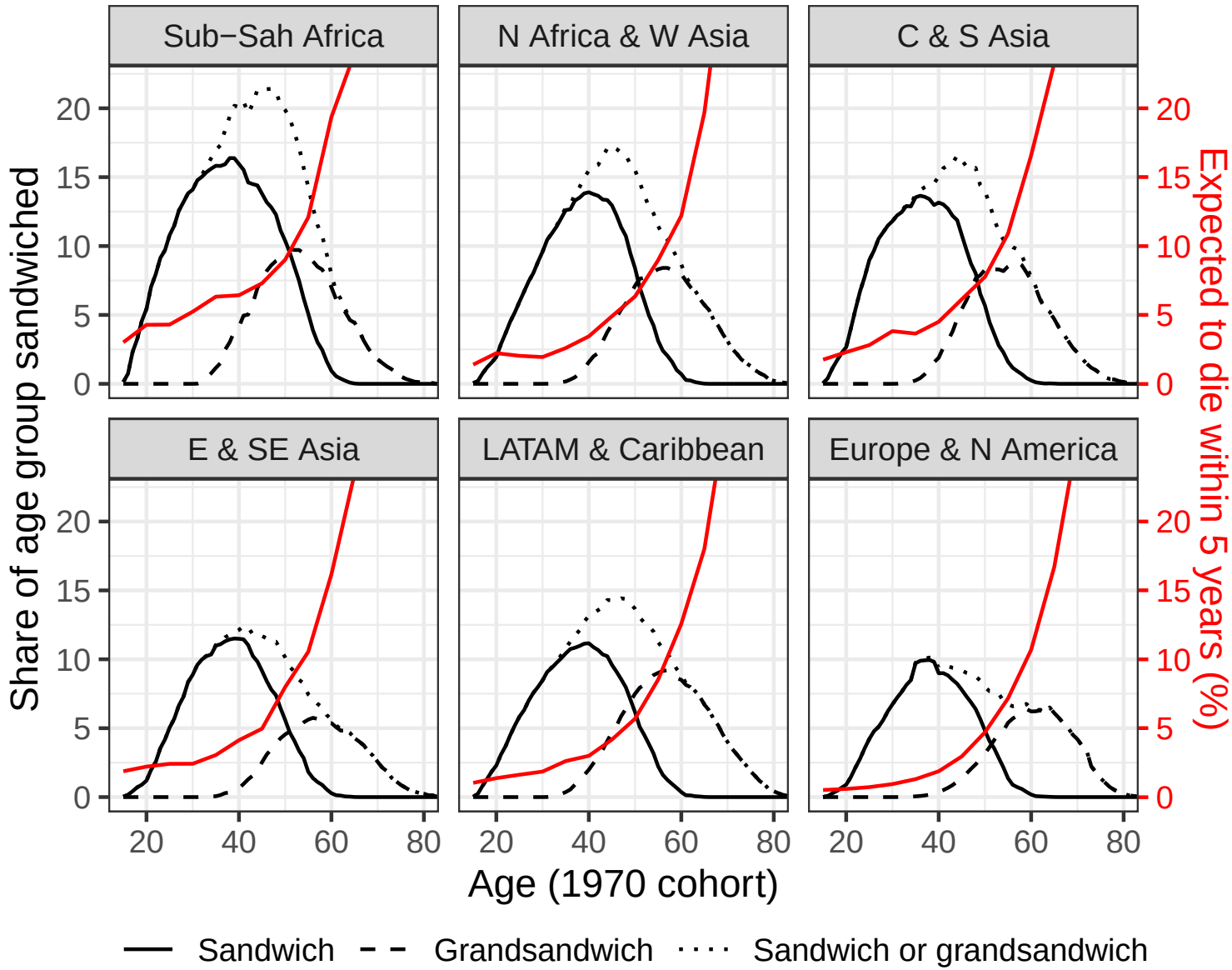
(2-3) Reduced fertility and later childbearing

$S(a)$ is the probability that a woman aged a has a child younger than 15 and a parent within 5 years of death:

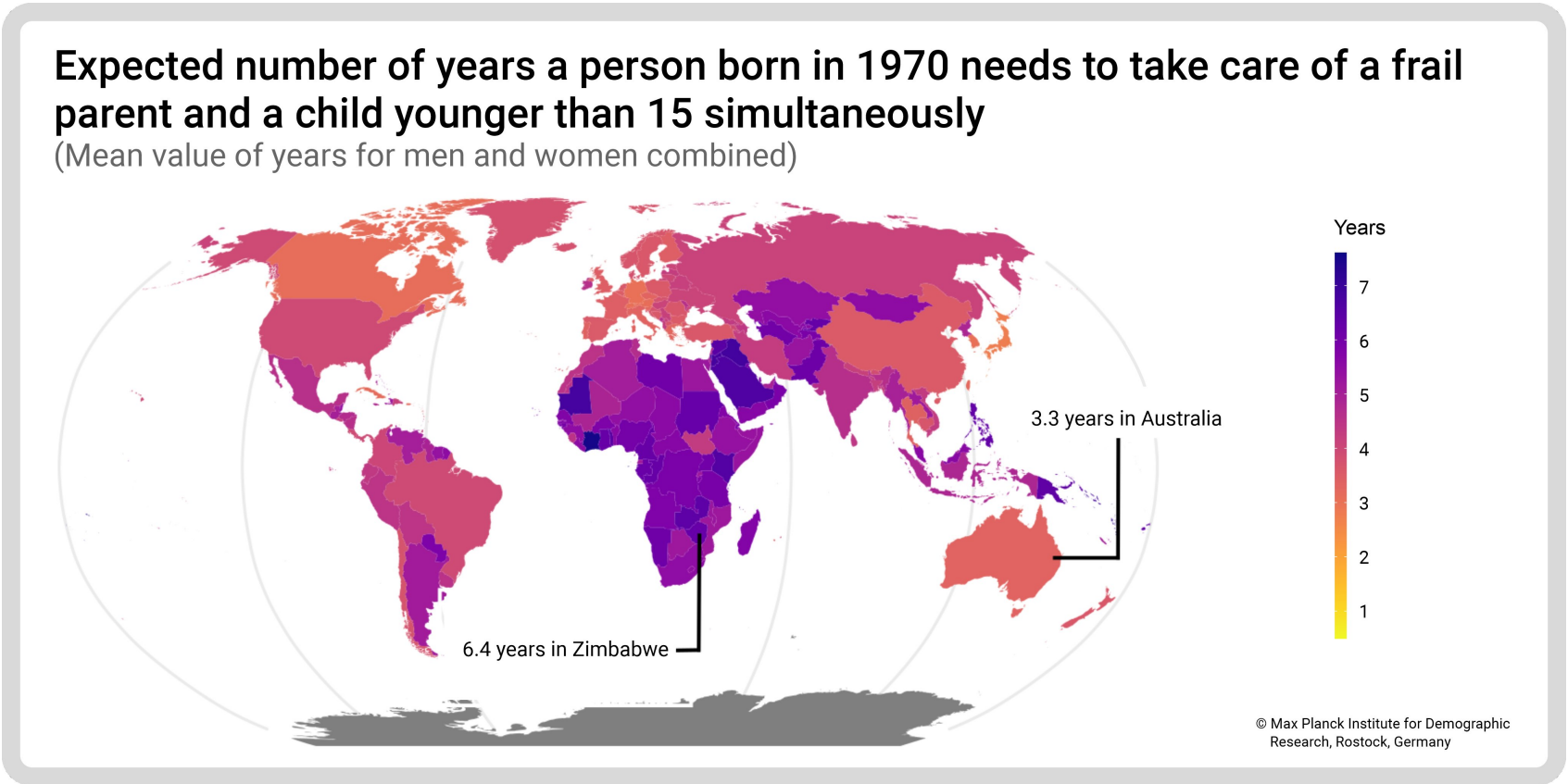
$$S(a) = \underbrace{\left(1 - \prod_{x=1}^{15}[1 - m(a-x)]\right)}_{\substack{\text{fertility risk in the} \\ \text{15 years preceding age a}}} \times \underbrace{M_1(a)}_{\substack{\text{Prob. ego's mother} \\ \text{alive}}} \times \underbrace{\left(1 - \frac{M_1(a+5)}{M_1(a)}\right)}_{\substack{\text{Prob. ego's mother} \\ \text{dies in 5 years}}}$$

where the probability of having a living mother in a stable population is:

$$M_1(a) \approx \frac{l(\mu + a)}{l(\mu)}.$$


Sub−Sah Africa
N Africa & W Asia
C & S Asia
E & SE Asia
LATAM & Caribbean
Europe & N America
Share of age group sandwiched
Expected to die within 5 years (%)
Age (1970 cohort)
0
5
10
15
20
20
40
60
80
Sandwich
Grandsandwich
Sandwich or grandsandwich

Sandwichness around the planet



`rsocsim`: demographic microsimulation in R

- Efficient R-language implementation
- Runs in any OS
- Built-in functions
- https://github.com/MPIDR/rsocsim
- More in the lab session...

Tom Theile

Liliana P. Calderón-Bernal

Mallika Snyder

Emilio Zagheni

Models and simulations

Discuss:

1. Compare and contrast model- and simulation-based approaches to study kinship
2. What are the relative advantages or each approach?
3. When would you chose one over the other?

Pros and cons of SOCSIM microsimulation

Strengths

1. Keep track of kinship ties
2. Full genealogies
3. Low data requirements
4. Flexible and adaptable states

Limitations

1. Not real populations
2. Correlated input rates
3. Computing power

When to use demographic microsimulations?

1. Complex inter-generational processes
2. Trace ancestry or relatedness
3. Improve the interval validity of simulations
   - Calibration
   - Comparing simulations to ground-truth
   - Methodological triangulation