4 - Computational approaches

Diego Alburez-Gutierrez
MPIDR
European Doctoral School of Demography 2019-20

02/04/2020

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

Agenda

1. Q&A
2. Demographic micro simulation
3. Example 1: Impact of the HIV/AIDS epidemic on kinship resources
4. Example 2: Projecting older adults without kin
5. Discussion

Q&A

- Previous days
- Questions about final assignment
- Other?

Demographic micro-simulation

- ▶ Model individual-level demographic behaviour applying set of rules
- ▶ Make up data where unavailable
- ▶ Science: compare to independent method
- ▶ Different alternatives:
  - ▶ SOCSIM
  - ▶ CAMSIM
  - ▶ NetLogo (Agent-based modelling)
  - ▶ R/python

Grow, A and Van Bavel, J. 2018. Agent-Based Modeling of Family Formation and Dissolution. In R. Schoen (Ed.), Analytical Family Demography (pp. 125-156). Springer Series on Demographic Methods and Population Analysis, (Vol. 47), Cham: Springer International Publishing.

Creating digital populations

- A stochastic micro-simulation platform, 1970s at UC Berkeley
- Starting with an initial population, applies age-specific demographic rates
- Creates kinship structure similar to a full genealogy
- Now maintained at the MPIDR!

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

A SOCSIM micro-simulation of Sweden (1603-2160)

```r
# Read sample Familinx data using data.table
read.csv("../../Assignment/Data/sweden_socsim.csv") %>%
  slice(1:4) %>%
  kable()
```

profileid	father	mother	birth_year	death_year
10000	2152	2390	1703	1705
10001	0	4343	1703	1707
10002	4593	5190	1703	1773
10003	0	3252	1703	1703

Zagheni, E. 2017. The Demographic Foundations of the Lived Experience of Kin Death. Working paper.

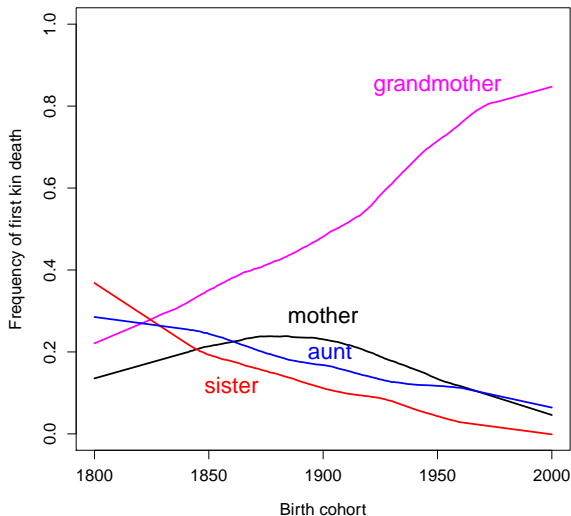# Historical and projected demographic processes

Figure 1: Frequency of different types of kin death, Sweden (SOCSIM)

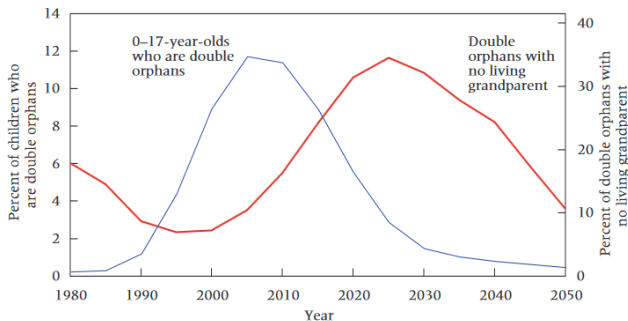Example 1: Impact of the HIV/AIDS epidemic
on kinship resources

Research at a glance

- ▶ RQ: estimate and project probabilities of orphanhood and evolution of kinship structure in Zimbabwe in context of HIV/AIDS epidemic (1980-2050)
- ▶ Data: SOCSIM, with rates from UN WPP, Demographic and Health Surveys, World Fertility and Marriage Database, UN HIV infection rates
- ▶ Findings:
  - ▶ increase in double orphans with no living grandparents
  - ▶ shift of responsibilities to aunts and uncles

{Zagheni, E. 2011. The impact of the HIV/AIDS epidemic on kinship resources for orphans in Zimbabwe, Population and Development Review 74(4), 761-783.}

Double-orphans and double-orphans without grandparents


FIGURE 2   Percent of children 0–17 years old who are double orphans (left scale), and percent of these double orphans who have no living grandparent (right scale)

{Zagheni, E. 2011. The impact of the HIV/AIDS epidemic on kinship resources for orphans in Zimbabwe, Population and Development Review 74(4), 761-783.}

Percent of double orphans

FIGURE 4 Estimates and projections of the percent of double orphans younger than 10 years who have at least one living sibling older than 15 years, Zimbabwe 1980–2050
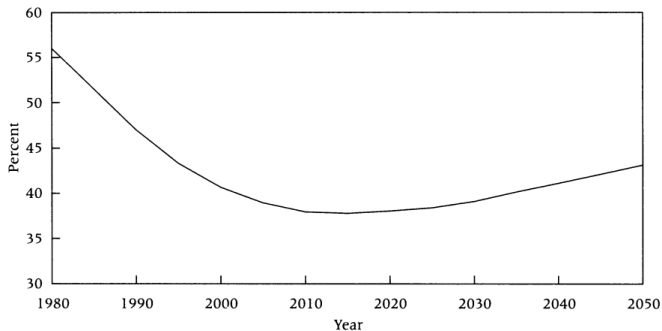
{Zagheni, E. 2011. The impact of the HIV/AIDS epidemic on kinship resources for orphans in Zimbabwe, Population and Development Review 74(4), 761-783.}

Availability of aunts and uncles

FIGURE 3 Estimates and projections of the average number of living uncles and aunts of double orphans 0–17 years old, Zimbabwe 1980–2050
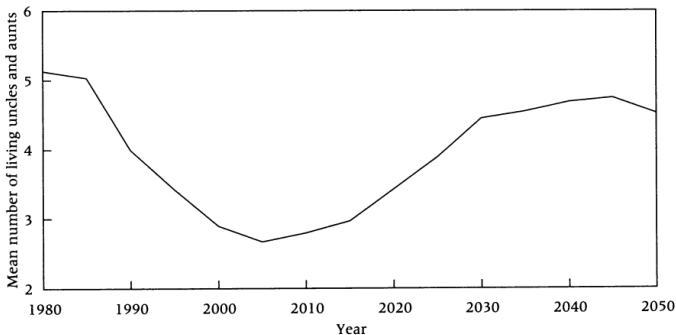
{Zagheni, E. 2011. The impact of the HIV/AIDS epidemic on kinship resources for orphans in Zimbabwe, Population and Development Review 74(4), 761-783.}

Example 2: Projecting older adults without kin

Research at a glance

- ▶ RQ: Examine the changing population of kinless individuals in American society over the coming decades
- ▶ Data: Rates from US census, Human Fertility Database, official statistics
- ▶ Findings:
  - ▶ impending increase of kinless older adults, especially amongst Black
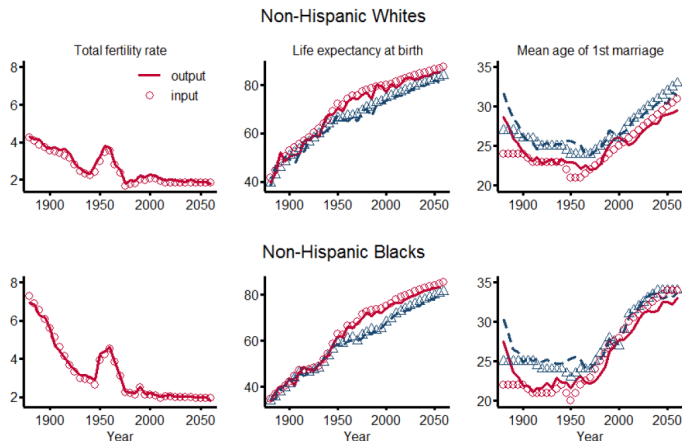 - ▶ Declines in marriage, one-child families, mortality

Verdery, A.M. and Margolis, R. (2017). Projections of white and black older adults without living kin in the United States, 2015 to 2060. Proceedings of the National Academy of Sciences 114(42):11109–11114.

Data sources

Demographic parameter	Time period	Source
SI.1. Sources of demographic parameters for the simulation models.		
Initial Populations		
	1880	(1)
Life expectancy at birth		
	1880-1949	(3)
	1950-2014	(5)
	2015-2060	(6)
Total fertility rate		
	1880-1939	(3)
	1940-1979	(3, 8)
	1980-2013	(7)
	2014-2060	(9)
Proportion male at birth		
	1880-2060	(12)
Marital status birth proportions		
	1901-2060	(1)
Parity status birth proportions		
	1901-2060	(10)
Marriage rates		
	1880-2010	(13)
	2011-2060	Extrapolation
Remarriage rates		
	1880-2060	(14, 15)
Divorce rates		
	1880-1967	(16, 17)
	1968-2014	(16, 18)
	2015-2060	Extrapolation
Partnership rates	1880-1997	Non-marital childbearing
	1998-2005	(19, 20, 22, 24)
	2006-2060	Extrapolation and assumptions
Aging and population size scaling		
	1880-2060	(1, 6)

Sanity checks: comparing to ground-truth

Key rates, historical and projected changes over time and simulated outcomes, 1880- 2060.

Verdery, A.M. and Margolis, R. (2017). SI Appendix. Projections of white and black older adults without living kin in the United States, 2015 to 2060. Proceedings of the National Academy of Sciences 114(42):11109–11114.
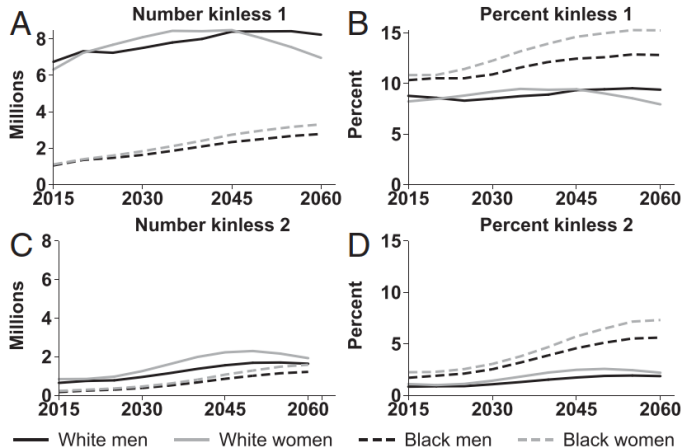
Kinlesness by gender and ethnicity in the US

Fig. 1. Projected numbers kinless 1 (*A*), percent kinless 1 (*B*), numbers kinless 2 (*C*), and percent kinless 2 (*D*), people age 50 and older, by year, sex, and race.

Verdery, A.M. and Margolis, R. (2017). Projections of white and black older adults without living kin in the United States, 2015 to 2060. Proceedings of the National Academy of Sciences 114(42):11109–11114.

Beyond description: looking at mechanisms
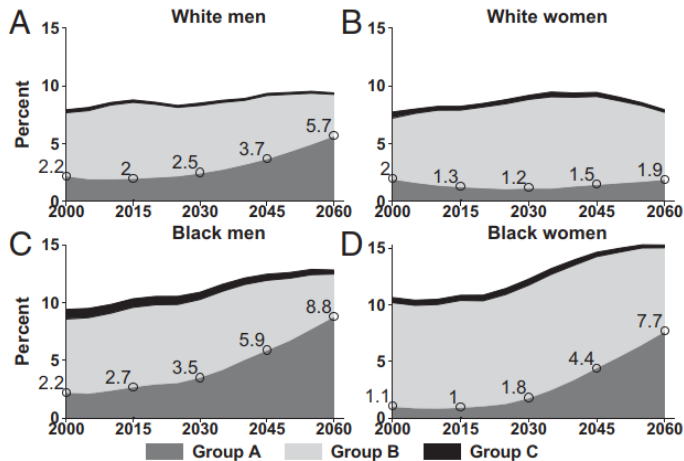
Verdery, A.M. and Margolis, R. (2017). Projections of white and black older adults without living kin in the United States, 2015 to 2060. Proceedings of the National Academy of Sciences 114(42):11109–11114.

Discussion

Strengths and weaknesses of simulated data

- ▶ Impact of the HIV/AIDS epidemic
  - ▶ Pro: No alternative data source
  - ▶ Pro: accounts for clustering of mortality
  - ▶ Con: comparison to ground-truth?
  - ▶ Con: high uncertainty of projected rates used as input in this context
- ▶ Projecting older adults without kin
  - ▶ Pro: Projection based on real rates
  - ▶ Pro: unpacks demographic dynamics leading to outcome
  - ▶ Con: model assumptions about marriage market
  - ▶ Con: comparison to ground-truth?
  - ▶ Con: does not account for future mortality shocks

When should we use real and simulated populations?

- ▶ Use real data whenever possible
- ▶ Improve the interval validity of simulations
    - ▶ Calibration
    - ▶ Methodological triangulation
    - ▶ Comparing simulations to ground-truth

A SOCSIM micro-simulation of Sweden (1603-2160)

```r
# Read sample Familinx data using data.table
read.csv("../../Assignment/Data/sweden_socsim.csv") %>%
  slice(1:4) %>%
  kable()
```

| profileid	father	mother	birth_year	death_year
10000	2152	2390	1703	1705
10001	0	4343	1703	1707
10002	4593	5190	1703	1773
10003	0	3252	1703	1703

Where did this SOCSIM simulation come from?

1. Initial population
2. Age-speficid fertility rates
3. Age-speficid mortality rates
4. Marriage transition rates
5. Model for marriage market
6. Other parameters (inheritance of fertility, etc.)

A quick example of such a comparison

- Cumulative number of child deaths for a woman surviving to a given age
  - Estimate from SOCISM-generated genealogy
  - Estimate formally:

$$\underbrace{CD_{(a,c,p)}}_{\text{Child deaths}} = \underbrace{\sum_{x=15}^{x=a} {_1F_{(x,c,p)}}}_{\text{Children born}} - \underbrace{\sum_{x=15}^{x=a} {_1F_{(x,c,p)}} l_{(a-x,c+x,p)}}_{\text{Children surviving or } CS_{(a,c,p)}} \quad (1)$$

Methodological triangulation: model and formal estimates
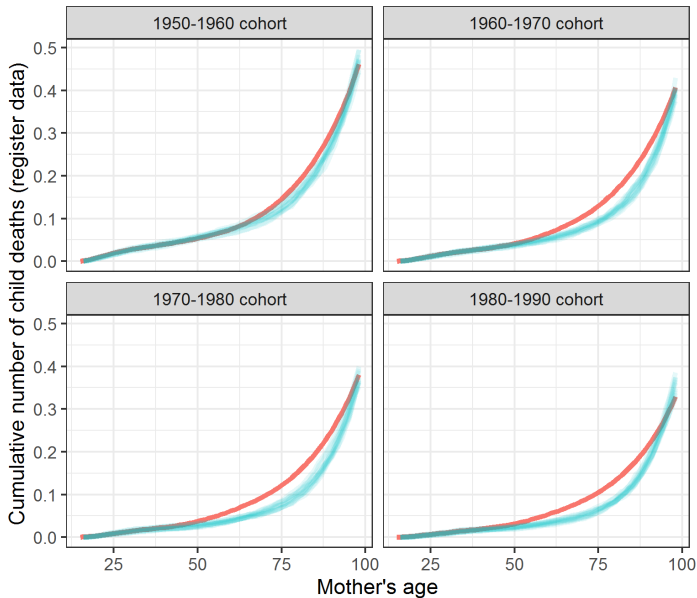
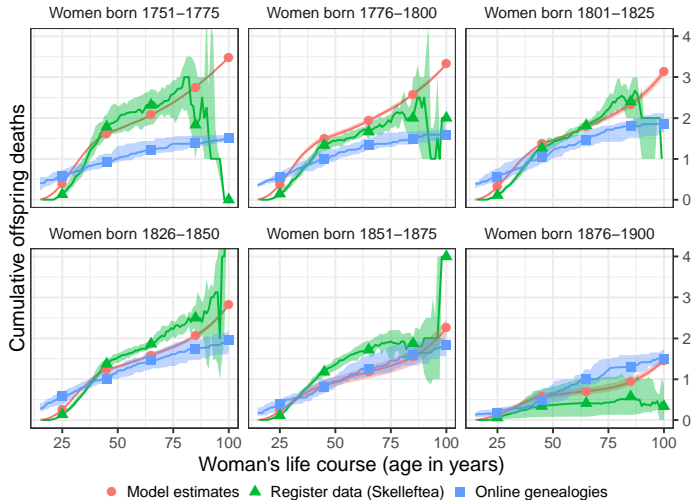# More validation: Compare to 'gold-standard' data

Figure 3: Expected number of child deaths

Another example: sandwichness

- 'Sandwiched' between having a young child and a parent close to death
- Double care responsibility
- Change over time

$$S(a, c) = \underbrace{(1 - \prod_{x=1}^{5} 1 - F_{a-x,c})}_{\substack{\text{Prob. of having given} \\ \text{birth in 5 preceding years}}} \times \underbrace{M_{a,c}}_{\substack{\text{P. mother} \\ \text{is alive}}} \times \underbrace{(1 - \frac{M_{a+5,c}}{M_{a,c}})}_{\substack{\text{Prob. that mother} \\ \text{dies within 5 years}}}$$

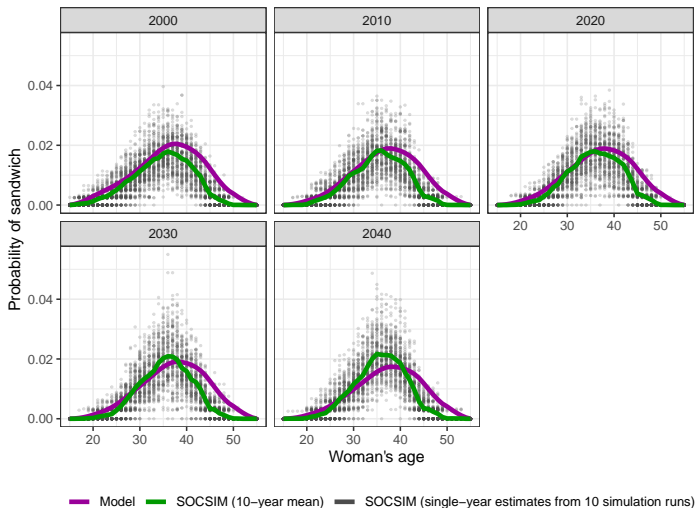# Comparing model and simulated estimates

Figure 4: Expected number of child deaths in the USA

Wrapping up

- ▶ How does this relate to your interests?
- ▶ Final thoughts?
- ▶ Brainstorming on project ideas?