Racial disparities in infant outcomes in the United States

Insights from, and for, formal demography

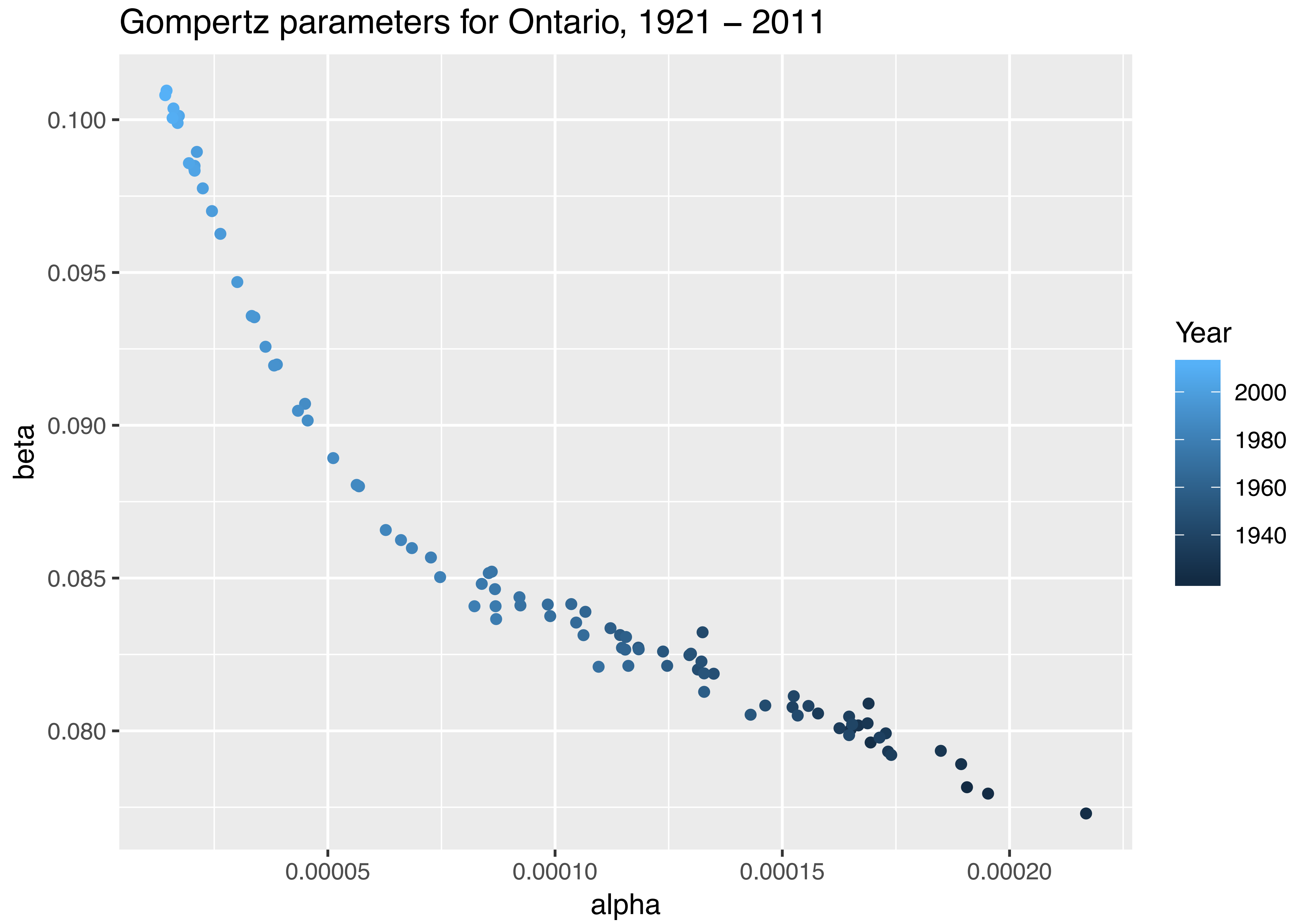
MONICA ALEXANDER, UNIVERSITY OF TORONTO

SOCIAL DEMOGRAPHY SEMINAR, 3 NOVEMBER 2022

Hello from sunny Toronto

Background

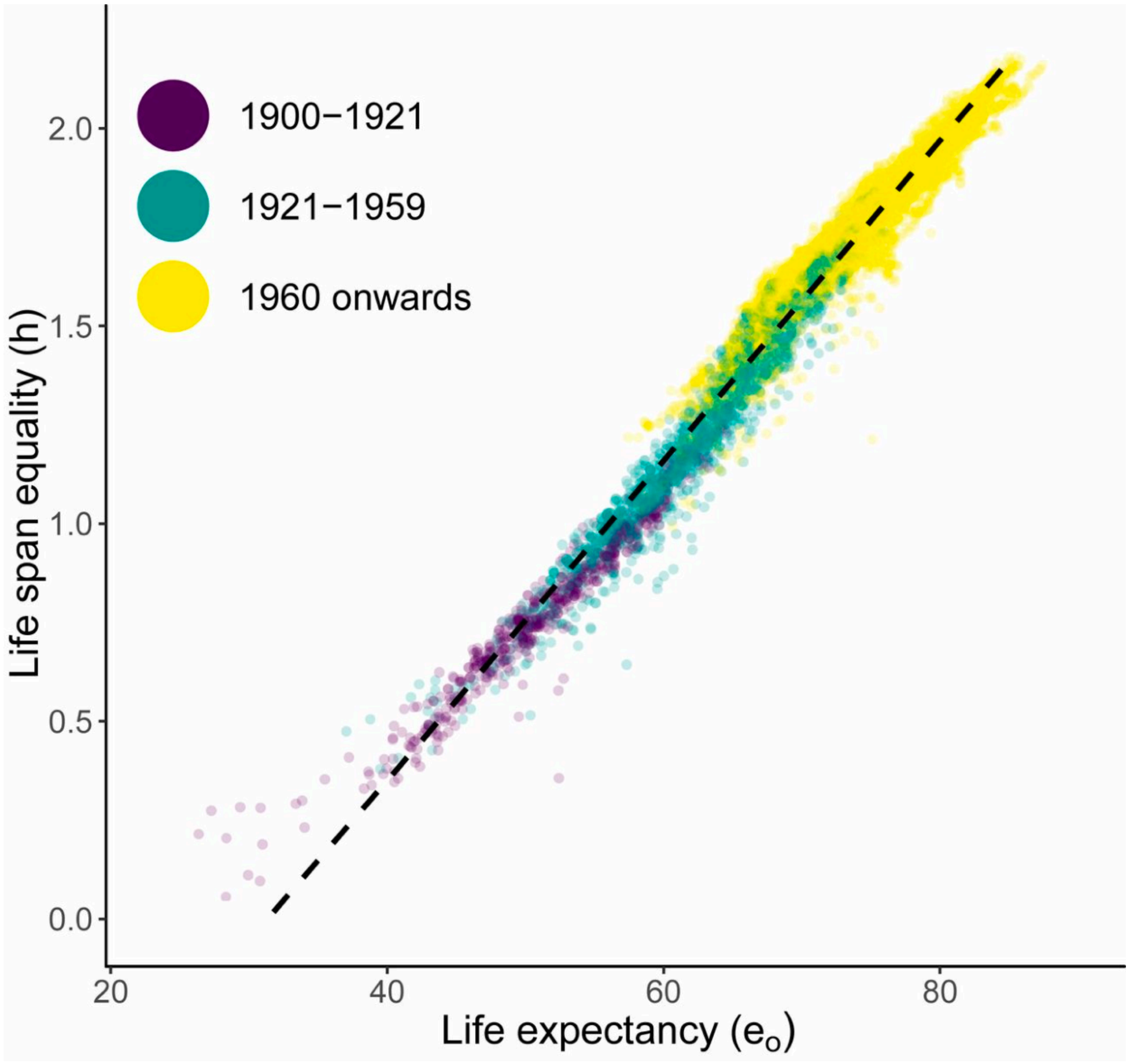
- Formal demography involves study of population processes using **mathematical** and **empirical** models
- Formal demographic models are useful to describe and compare populations through time and space with a reduced set of parameters



Source: My calculations using data from the Canadian Human Mortality Database

Background

- But these models are evolving over time, as we gain new insights from population processes
- Mathematical models: different priorities and viewpoints lead to different formulations and insights (e.g. lifespan disparity)
- Empirical models: we have new experiences, new questions, and new data



Source: Aburto, J. M., Villavicencio, F., Basellini, U., Kjærgaard, S., & Vaupel, J. W. (2020). Dynamics of life expectancy and life span equality. Proceedings of the National Academy of Sciences, 117(10), 5250-5259

This talk

- Demonstrate how formal demographic methods help to study racial disparities in infant outcomes, **and vice versa.**

- Touching on two projects:

1. How racial differences in infant outcomes help to improve demographic models

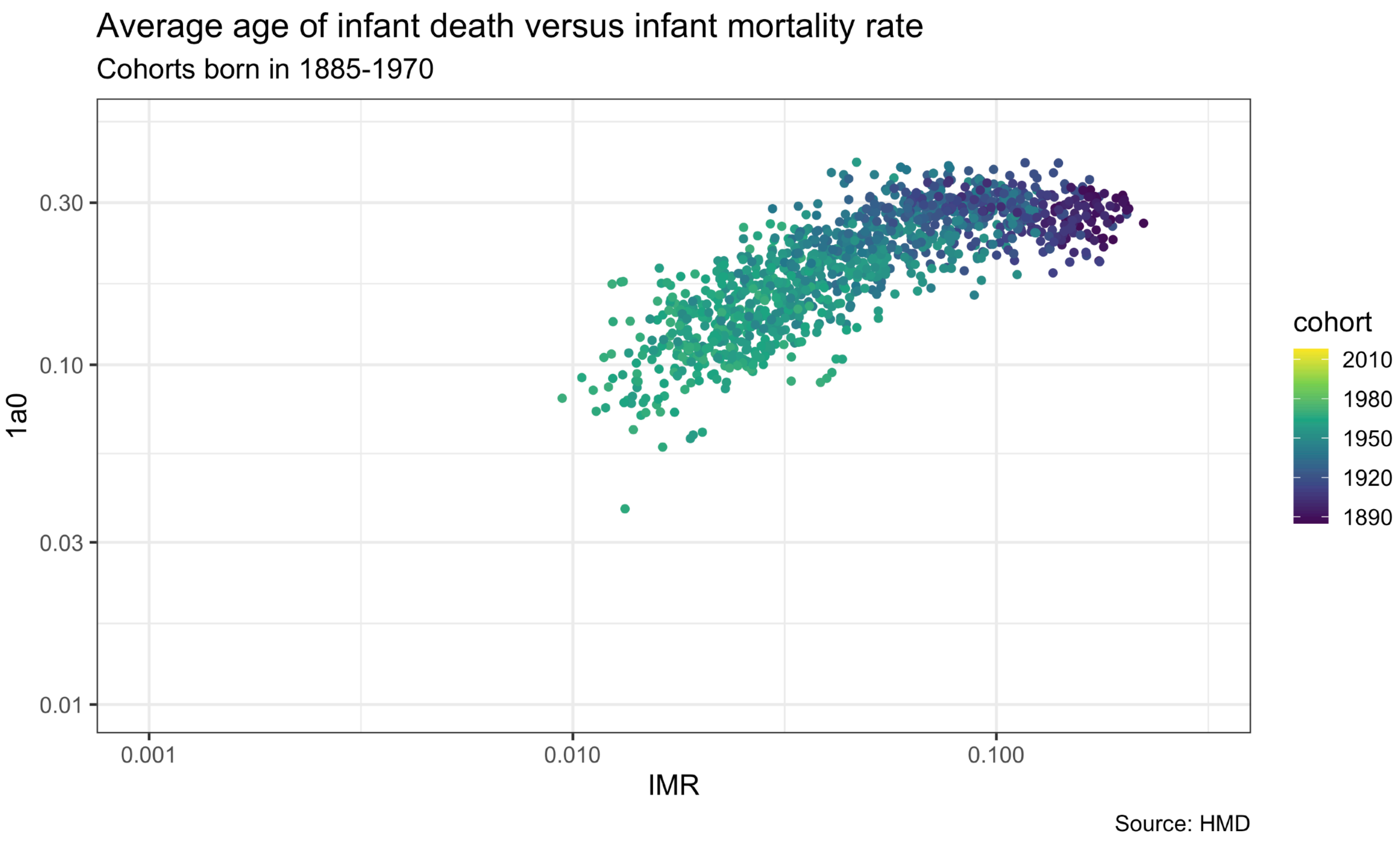
2. How demographic methods help to quantify the extent of disparities in fetal and infant outcomes across gestational age

Competing effects on the average age of infant death

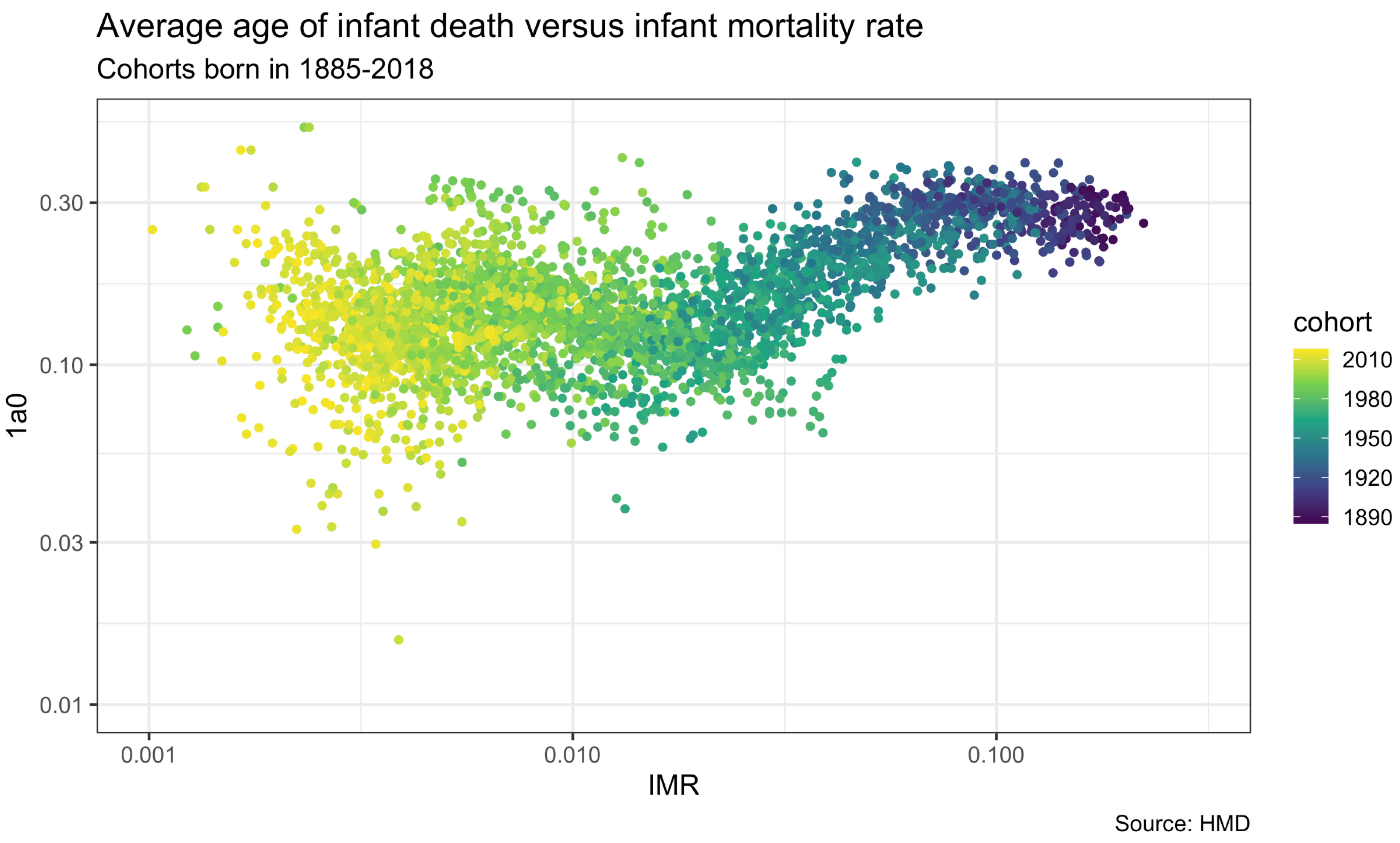
Background

- Life table quantity ${}_na_x$ (‘average time lived for those who died’)
- A particular quantity of interest is the average age of infant death, i.e. the average number of days lived of those who died in the first year of life: ${}_1a_0$
- To get accurate measure, need individual life lines (know everyone's date of birth and death)
- More realistically, we usually only have death and population counts aggregated by age, so need to approximate
- Existing approximations take advantage of the strong relationship between ${}_1a_0$ and the infant mortality rate

Empirical relationship



... oh no


Average age of infant death versus infant mortality rate
Cohorts born in 1885-2018
0.30
0.10
0.03
0.01
1a0
0.001
0.010
0.100
IMR
cohort
2010
1980
1950
1920
1890
Source: HMD

Existing approximations

- Coale-Demeny (1966)

Regional family of model life tables	For females	For males
infant mortality rate ${}_1q_0 \geq 0.1$		
"West," "North," "South"	0.35	0.33
"East"	0.31	0.29
infant mortality rate ${}_1q_0 < 0.1$		
"West," "North," "South"	$0.050+3.000 \cdot {}_1q_0$	$0.0425+2.875 \cdot {}_1q_0$
"East"	$0.010+3.000 \cdot {}_1q_0$	$0.0025+2.875 \cdot {}_1q_0$

- Keyfitz and Flieger (1968, 1970, 1971): ${}_1a_0 = 0.07 + 1.7\,{}_1M_0$

- Andreev and Kingkade (2015): updating using new data

Lower limits ${}_1q_0$	Upper limits ${}_1q_0$	Equation
Male		
0	0.0226	$0.1493 - 2.0367 \cdot {}_1q_0$
0.0226	0.0785	$0.0244 + 3.4994 \cdot {}_1q_0$
0.0785	+	0.2991
Female		
0	0.0170	$0.1490 - 2.0867 \cdot {}_1q_0$
0.0170	0.0658	$0.0438 + 4.1075 \cdot {}_1q_0$
0.0658	+	0.3141

Goals

- Understand and illustrate the relationships between ${}_1a_0$, infant mortality, and premature births
- Use these relationships to motivate a new approximation formula in low-mortality settings
  - Take a more 'explanatory' model building approach, rather than purely curve-fitting
  - Approximation formula depends on data/estimates that are widely available

Joint work with Leslie Root [*Demography* (2022) 59 (2): 587–605]

Infant mortality, 1a0, and prematurity

Data

- Data from U.S. Birth Cohort Linked Birth and Infant Death Data of the National Center for Health Statistics' National Vital Statistics System, years 2008–2012
  - Information on birth outcomes, e.g. gestational age, birthweight, characteristics of mother, birth complications…
  - Information on death outcomes (if applicable), e.g. age at death, cause of death

Mortality rates

Infant mortality rates (per 1,000 births) are higher for Non-Hispanic black babies in (basically) all subgroups

		By Age at Death			By Gestational Age at Birth			
	All	First Week	Neonatal	Post-neonatal	Extremely Preterm	Very Preterm	Later Preterm	Full-Term
Black	11.00	5.94	7.30	3.70	363.18	38.86	9.01	3.54
White	4.98	2.60	3.26	1.72	365.32	38.90	8.65	2.09
Ratio	2.21	2.29	2.24	2.15	0.99	1.00	1.04	1.70

What would you expect to see for race-specific values of ${}_1a_0$?

A Simpson's Paradox

${}_1a_0$ is lower for the Non-Hispanic black population overall, even though it is higher in every gestational age category

	Aggregate	Extremely Preterm	Very Preterm	Later Preterm	Full-Term
Mean Age at Death (days)					
Black	40.87	14.21	41.73	68.26	84.97
White	45.29	9.93	31.53	51.05	82.85
Ratio	0.90	1.43	1.32	1.34	1.03

Prematurity

Premature births are much more likely for non-Hispanic Black population

	Aggregate	Extremely Preterm	Very Preterm	Later Preterm	Full-Term
Birth Rates (per 1,000)					
Black	—	16.79	20.32	128.15	834.75
White	—	5.36	10.14	87.42	897.09
Ratio	—	3.13	2.00	1.47	0.93

Because values of ${}_1a_0$ are much lower for preterm births, the weighted average for non-Hispanic blacks is lower

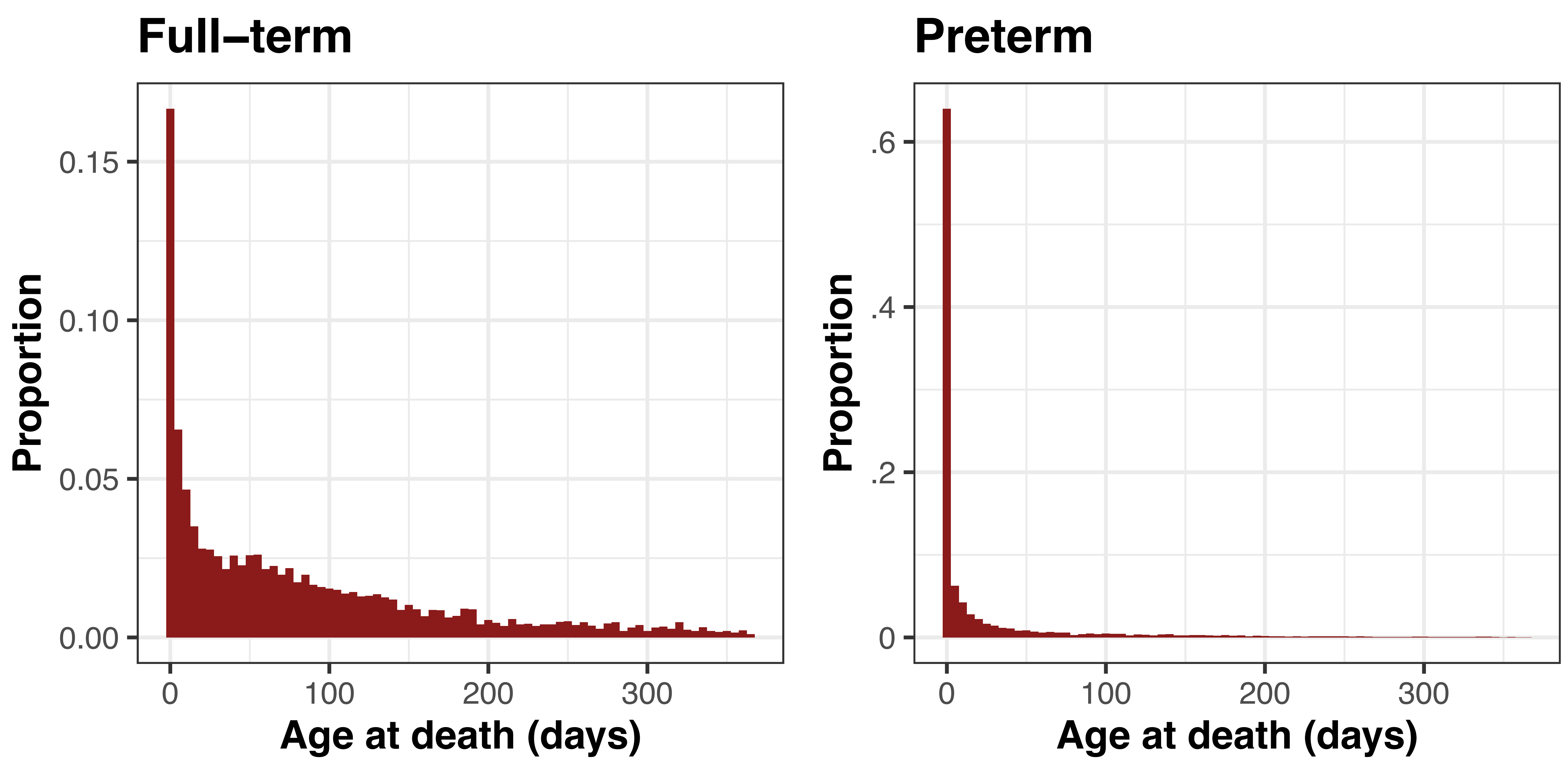
Relationships

1. The distribution of the timing of infant deaths is left skewed, with the majority of deaths in the first few days.

2. All other things being equal, the degree of this left skewness increases as mortality decreases.

3. The distribution of the timing of infant deaths conditional on births being premature is also heavily left skewed, with an even larger density of deaths in the first few days.

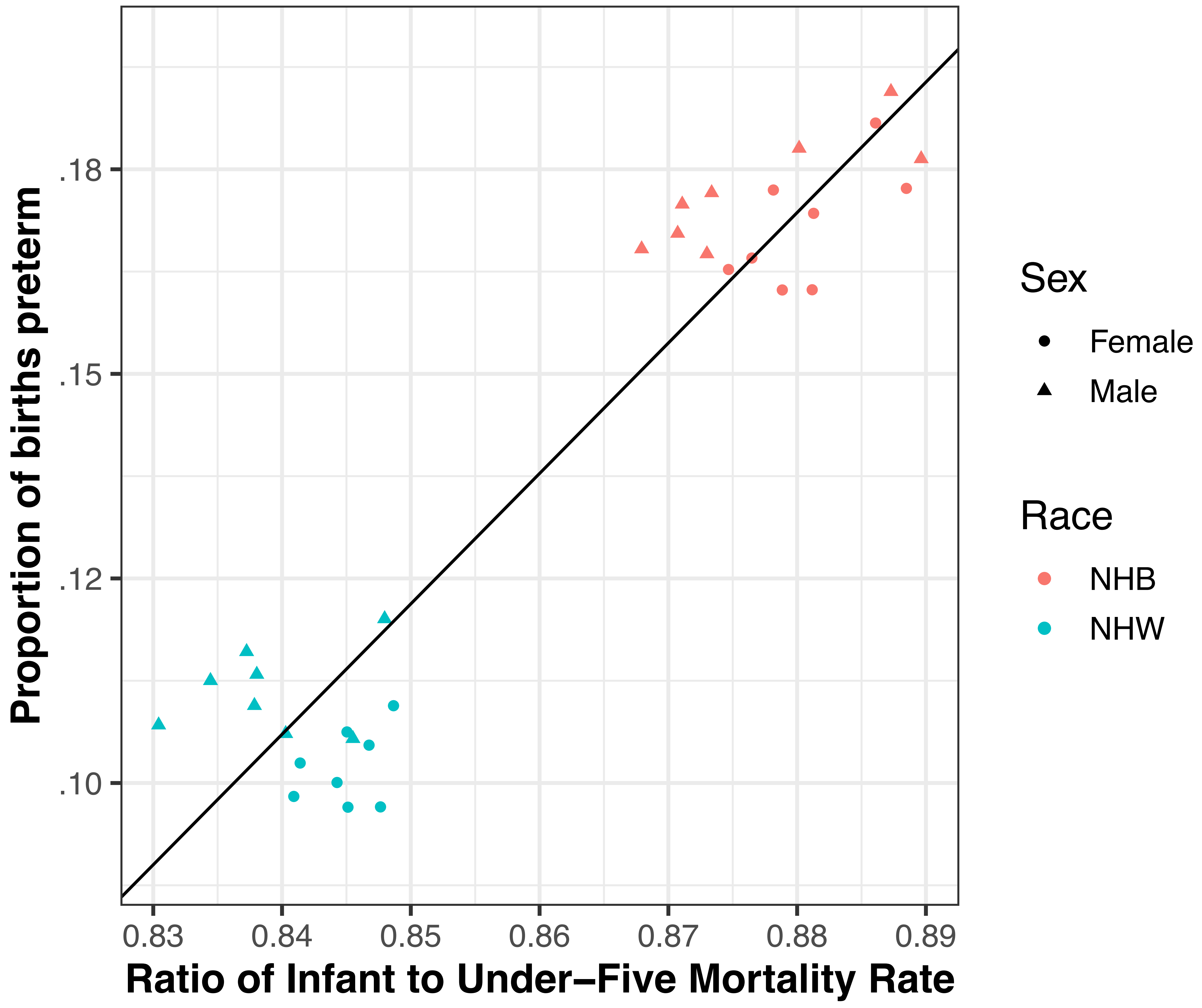
4. The share of births that are premature tends to decrease as mortality decreases



Statements 1 and 2 have the opposite effect from statements 3 and 4 on ${}_1a_0$

A new approximation in low-mortality settings

A proxy for prematurity



- The ratio of infant to total child mortality proxies prematurity
- As share of under-five mortality that is infant mortality increases, the proportion of births that are premature tends to increase
- Why do we use this? Because estimates of IMR and U5MR are readily available for all countries; while info on prematurity is not.

Model

'Ratio' model

$$ {}_1a_{0i} = \alpha + \beta_1 \left(\mathrm{IMR}_i\right) + \beta_2 \left(\frac{\mathrm{IMR_i}}{\mathrm{U5MR}_i} \right) + \varepsilon_i $$

Compare to piecewise linear model (e.g. from Andreev and Kingkade)

$$ {}_1a_{0i} = \alpha + \beta_1 \left(\mathrm{IMR_i}\right) + \beta_2 \left(\mathrm{IMR_i} - \theta\right) + \varepsilon_i $$

Data

- Data from the Human Mortality Database
- To evaluate model performance, we split data into a “training” data set, which comprises a random sample of 80% of all the available data, and a “test” data set, which comprises the remaining 20% of data.
- We also validated against the true race-specific values in the US population
- Code for model fitting and evaluation: https://github.com/MJAlexander/a0-competing-effects

Results

Model	Equation	Sex	$\widehat{\alpha}$	$\widehat{\beta}_1$	$\widehat{\beta}_2$	$\widehat{\theta}$
AK	2	F	0.1490	−2.0867	4.1075	0.0170
AK	2	M	0.1493	−2.0367	3.4994	0.0226
Piecewise (AK data)	2	F	0.149 [0.141, 0.155]	−1.984 [−2.820, −1.093]	4.856 [3.874, 5.695]	0.014 [0.013, 0.016]
Piecewise (AK data)	2	M	0.150 [0.144, 0.158]	−2.162 [−2.875, −1.588]	4.647 [4.027, 5.410]	0.019 [0.017, 0.021]
Piecewise	2	F	0.138 [0.134, 0.143]	−0.913 [−1.587, −0.328]	4.249 [3.619, 4.976]	0.017 [0.015, 0.019]
Piecewise	2	M	0.140 [0.136, 0.145]	−1.268 [−1.737, −0.883]	4.478 [3.987, 4.975]	0.023 [0.021, 0.025]
Piecewise	2	Both	0.139 [0.135, 0.142]	−1.004 [−1.381, −0.647]	4.297 [3.923, 4.708]	0.021 [0.019, 0.022]
Ratio	1	F	0.405 [0.375, 0.436]	1.975 [1.867, 2.084]	−0.359 [−0.398, −0.322]	—
Ratio	1	M	0.420 [0.386, 0.454]	1.551 [1.432, 1.669]	−0.382 [−0.425, −0.341]	—
Ratio	1	Both	0.426 [0.403, 0.447]	1.749 [1.666, 1.834]	−0.387 [−0.414, −0.359]	—

Model validation

HMD

Type	Sex	AK	Piecewise	Ratio
In-sample	F	0.038	0.038	0.037
In-sample (AK data)	F	0.035	0.038	0.036
Out-of-sample	F	0.038	0.038	0.036
In-sample	M	0.036	0.036	0.036
In-sample (AK data)	M	0.042	0.038	0.038
Out-of-sample	M	0.044	0.041	0.040
In-sample	Both	—	0.038	0.037
Out-of-sample	Both	—	0.041	0.038

US data

Race	Sex	Model	RMSE
NHB	F	Ratio	0.0090
NHB	F	Piecewise	0.0409
NHW	F	Ratio	0.0129
NHW	F	Piecewise	0.0689
NHB	M	Ratio	0.0080
NHB	M	Piecewise	0.0355
NHW	M	Ratio	0.0139
NHW	M	Piecewise	0.0710

Summary

- Patterns in infant outcomes by race in the US give insights into the dynamics of relationships between ${}_1a_0$, infant mortality, and prematurity
- Use these observations to motivate a new empirical model to approximate ${}_1a_0$ in low-mortality settings
- Model outperformed existing alternatives that were based on historically observed data, or purely curve fitting approaches

Comments

- We saw that the non-Hispanic Black population is more likely to experience preterm births AND the early death of a child
- These racial disparities extend to fetal deaths and stillbirths
- We tend to focus on studying these outcomes individually
- But there is a need to consider the cumulative and compounding risk of multiple potential outcomes in one comprehensive framework, and also to understand how risk evolves over gestational age

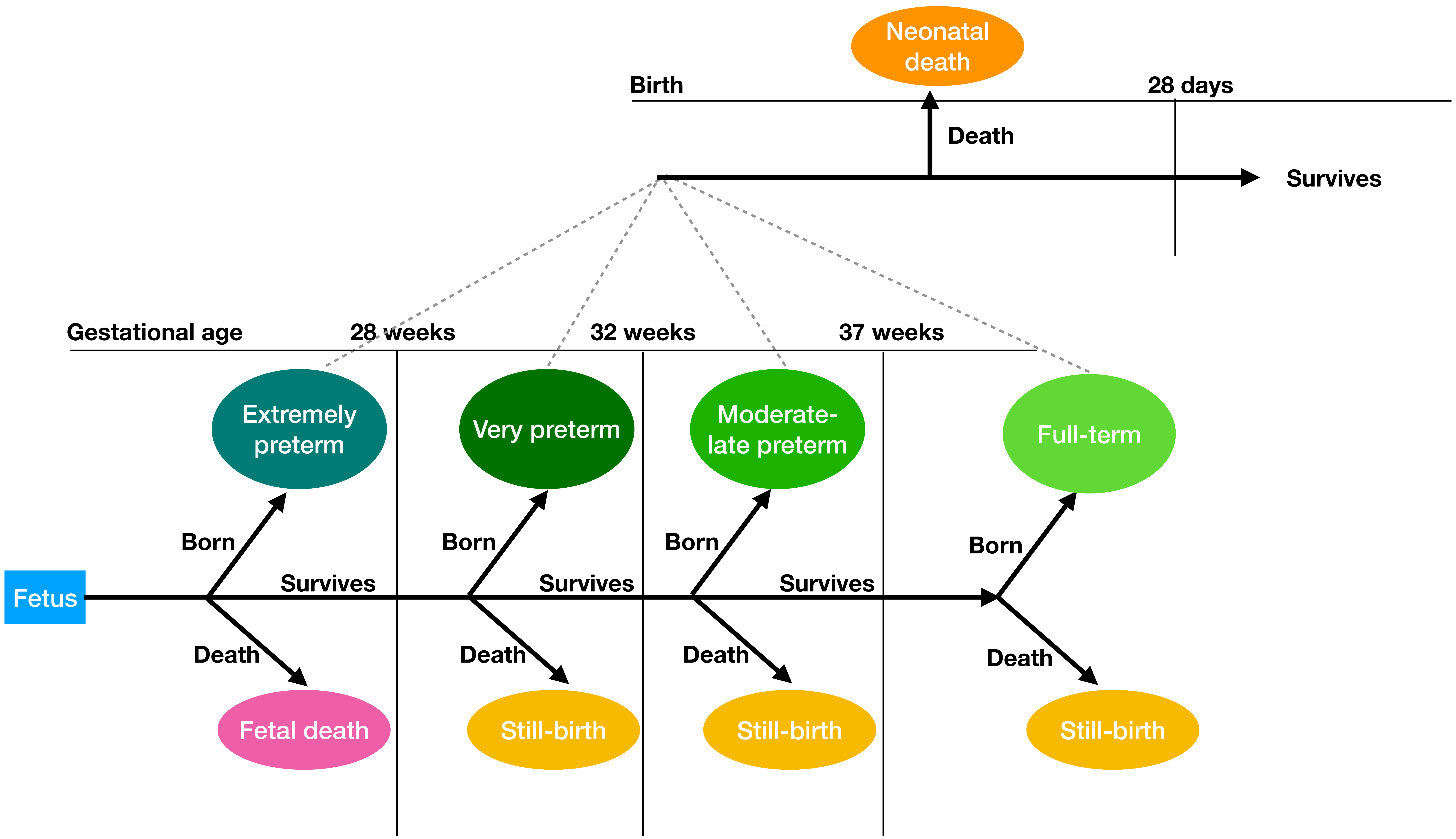
A multi-decrement lifetable approach to study outcomes by gestational age

Goals

- Propose a multi-decrement lifetable approach to study fetal and neonatal outcomes across different population groups
- Extend the life table framework to tabulate outcomes by gestational age (rather than age after birth)
- Tabulate decrements based on gestational outcome (death or birth, by prematurity) and also eventual neonatal outcome (death or survival)
- Allows us to encode the risk of all possible outcomes throughout the fetal-neonatal period, from 20 weeks of gestation

Joint work with Leslie Root (very much a work in progress!). Submission to PAA: https://www.monicaalexander.com/pdf/paa2023.pdf

Multiple potential outcomes across the fetal and neonatal period



Definitions

- Define x to be gestational age in weeks, $x = 20,\ldots,42+$
- We consider exists from gestation due to different causes i:
  - (extremely, very, moderate) premature live birth, survives to at least 28 days
  - full-term live birth, survives to at least 28 days
  - (extremely, very, moderate) premature live birth, neonatal death
  - full-term live birth, neonatal death
  - fetal death (<28 weeks gestation)
  - stillbirth (>= 28 weeks)

Definitions

We define the following life table columns:

- l_x is the number of fetuses still in gestation at gestational age x. l_{20} is referred to as the radix, which we set to be 100,000.

- d_x is the number of fetuses that have exited gestation (i.e. either through a live birth or death) at gestational age x

- d_x^i refers to the number of fetus that have exited gestation due to cause i at gestational age x.

- q_x is the probability that a fetus of gestational age x will exit through any cause

- q_x^i is the cause-specific probability that a fetus of gestational age x will exit

‘Lifetime’ risk of outcomes

- An additional measure that is interesting in studying fetal outcomes is one of ‘lifetime’ risk, that is, the probability that a fetus will eventually exit through a cause i, conditional on reaching gestational age x. This is calculated as

$$p_x^i = \frac{\sum_x^{\omega} d_x^i}{l_x}$$

- Where ω is the last gestational age group
- A nice property is that these are additive, so can be combined in various ways
- As before, data are from U.S. Birth Cohort Linked Birth and Infant Death Data of the National Center for Health Statistics' National Vital Statistics System, years 2008–2012

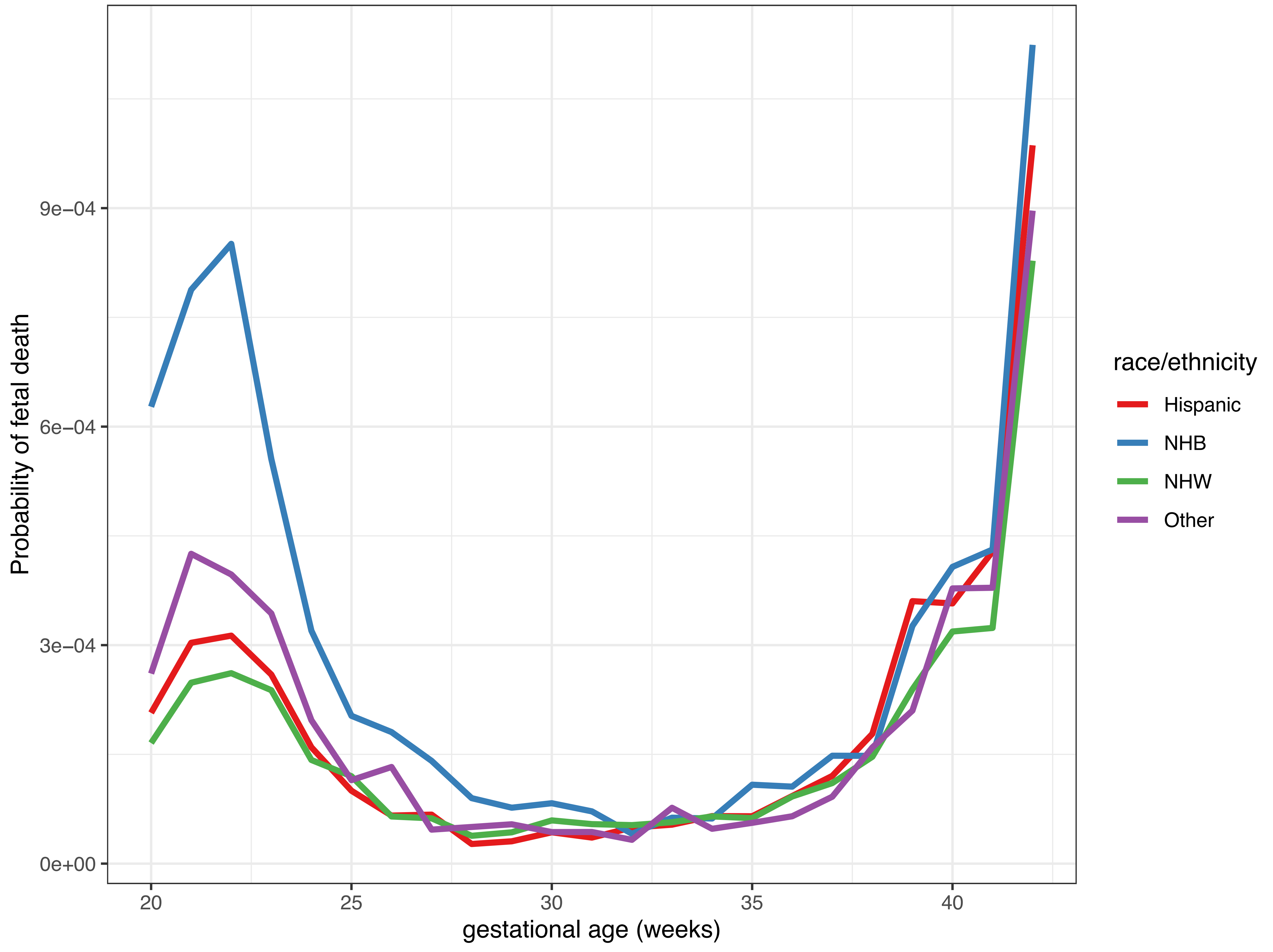
Some results

Life tables and age-specific probabilities

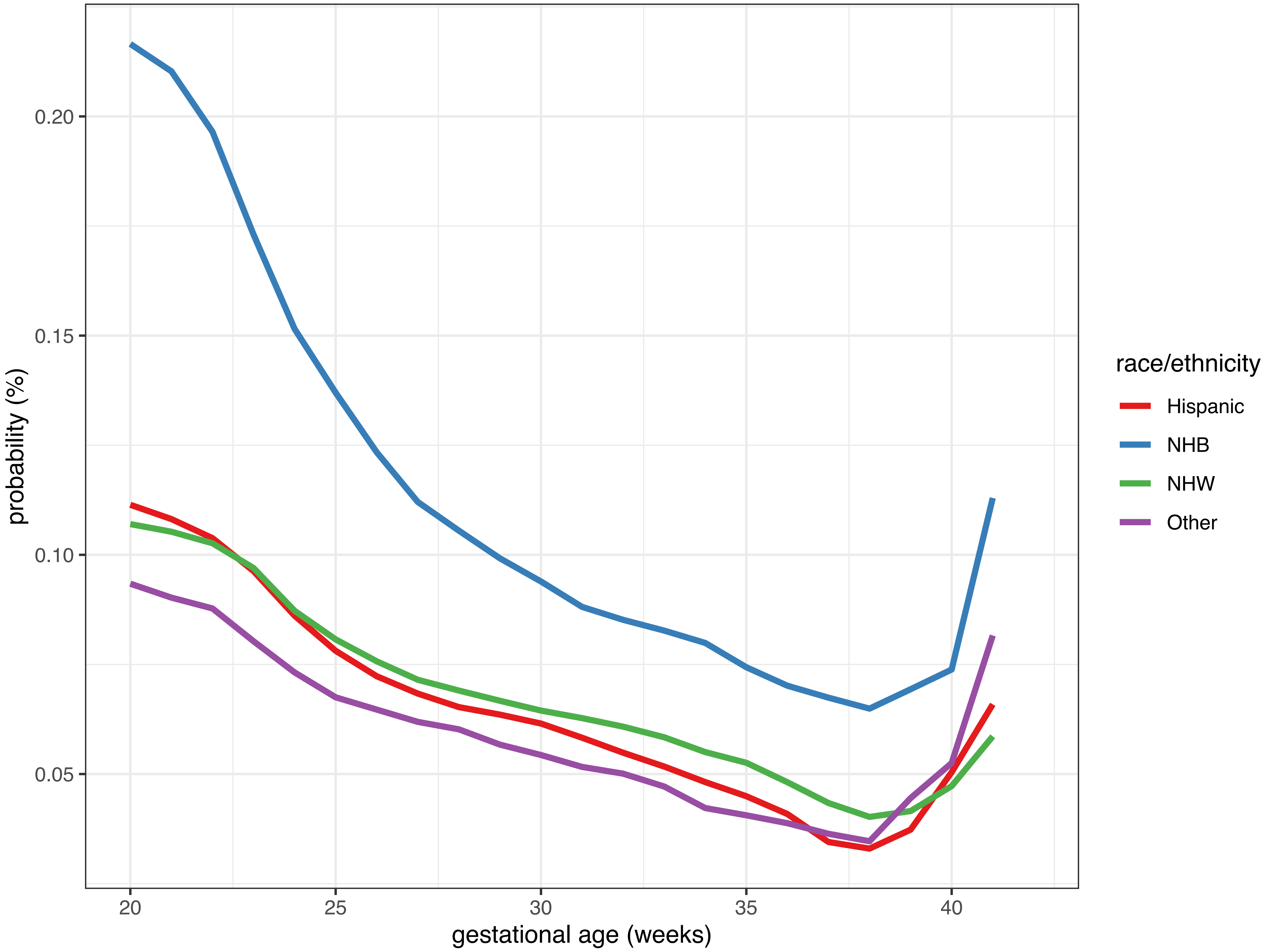
Fetal life table for non-Hispanic white population, with two possible outcomes: fetal death/stillbirth or live birth

x	lx	dx	dx_death	dx_livebirth	qx_death	qx_livebirth
20	100000.000	18.94715	16.552732	2.394420	0.0001655	0.0000239
21	99981.048	29.25357	24.829098	4.424472	0.0002483	0.0000443
22	99951.788	34.30267	26.130414	8.172261	0.0002614	0.0000818
23	99917.477	48.66920	23.788046	24.881151	0.0002381	0.0002490
24	99868.796	68.91766	14.210364	54.707300	0.0001423	0.0005478
25	99799.861	83.23213	11.972102	71.260032	0.0001200	0.0007140
26	99716.609	89.37434	6.454524	82.919819	0.0000647	0.0008316
27	99627.213	117.27455	6.194261	111.080284	0.0000622	0.0011150
28	99509.910	155.94964	3.799841	152.149799	0.0000382	0.0015290
29	99353.923	202.58878	4.268315	198.320470	0.0000430	0.0019961
30	99151.285	280.66771	5.881946	274.785764	0.0000593	0.0027714
31	98870.550	368.16816	5.361420	362.806740	0.0000542	0.0036695
32	98502.293	527.60532	5.205262	522.400063	0.0000528	0.0053034
33	97974.560	768.34868	5.569630	762.779048	0.0000568	0.0077855
34	97206.026	1328.95536	6.298367	1322.656995	0.0000648	0.0136067
35	95876.749	2105.42425	5.986051	2099.438196	0.0000624	0.0218973
36	93770.816	3825.86734	8.640734	3817.226605	0.0000921	0.0407080
37	89944.025	7959.73001	9.942050	7949.787964	0.0001105	0.0883859
38	81982.372	15189.16181	12.024154	15177.137658	0.0001467	0.1851268
39	66789.541	31003.05913	15.980153	30987.078979	0.0002393	0.4639511
40	35778.992	20649.01285	11.399523	20637.613325	0.0003186	0.5768081
41	15124.990	9336.78199	4.892946	9331.889047	0.0003235	0.6169848
42	5785.953	5784.55526	4.788841	5779.766419	0.0008277	0.9989308

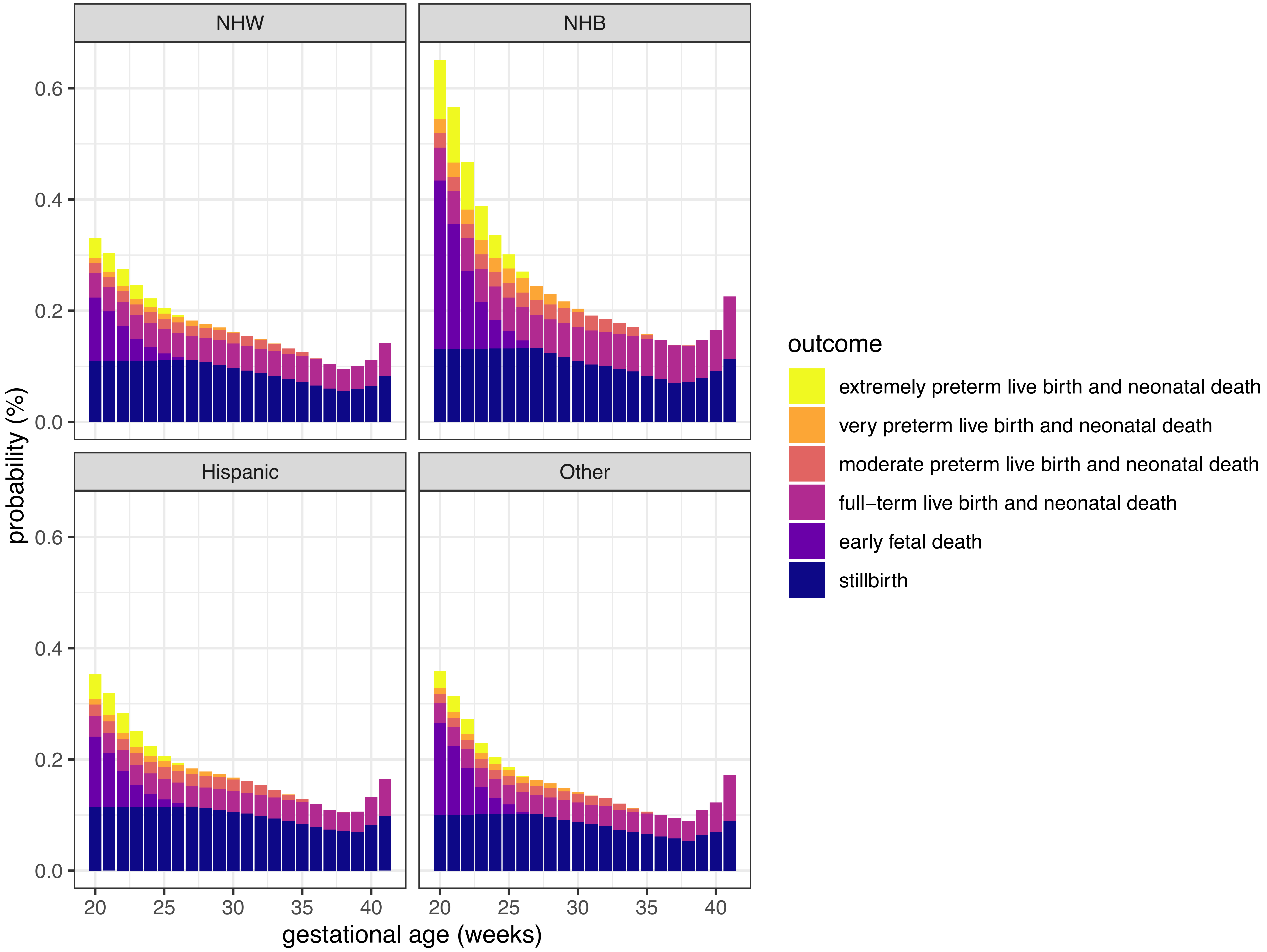
Age-specific probabilities of fetal death/stillbirth by race/ethnicity



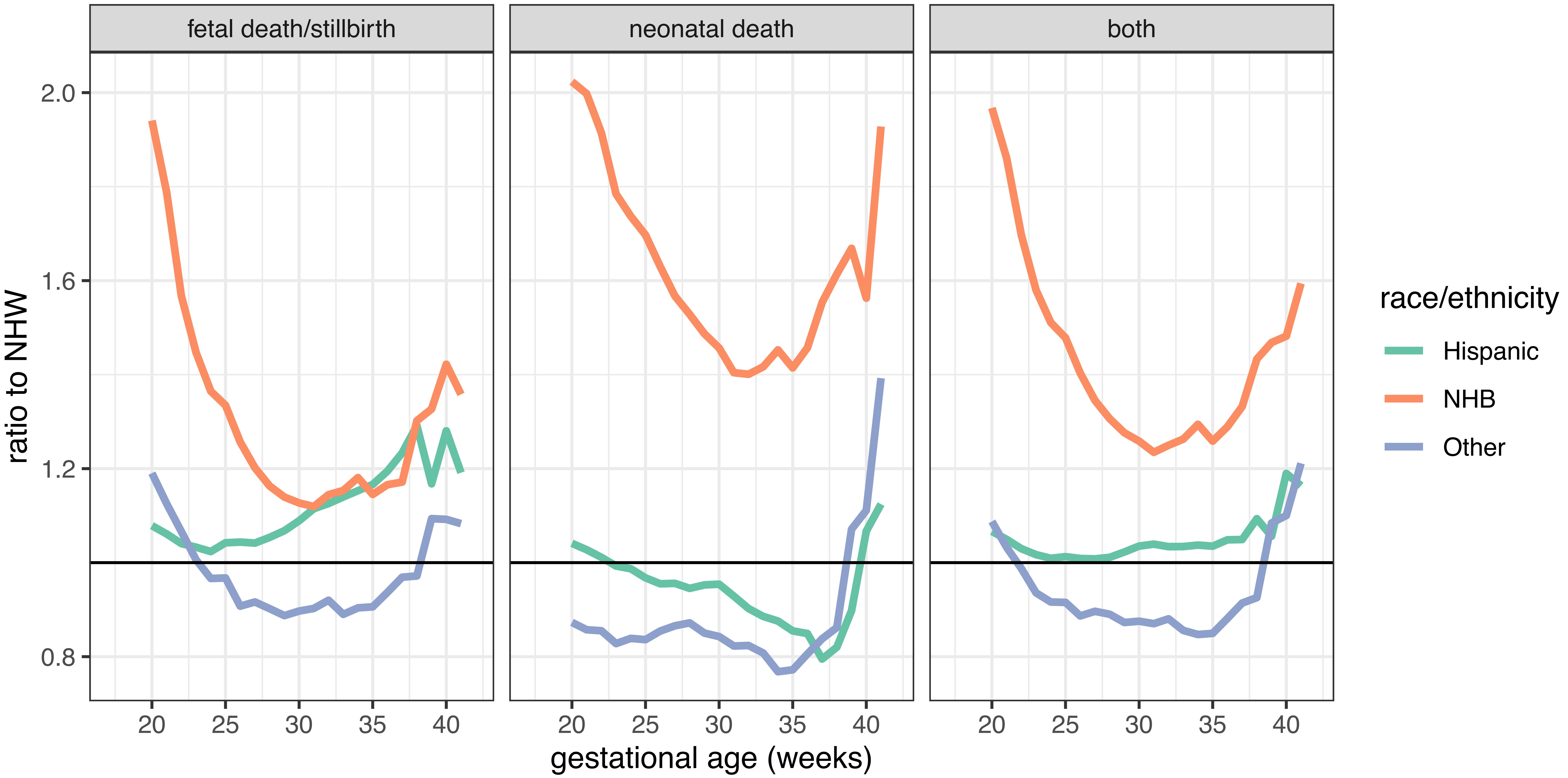
'Lifetime' risk of eventual neonatal death by gestational age



Combined outcomes



Risk ratios of lifetime risks



Summary

- Racial disparities are persistent across almost all dimensions of mortality in the United States, but are particularly stark around childbearing, childbirth, and the early years of life
- A need to fully quantify these risks and the experiences of these risks for the mother in concert
- Repurpose existing demographic methods to study gestational outcomes
- At 20 weeks gestation, NHB women have a ~0.8% chance of experiencing any adverse outcome. (contrast: ~0.3% for NHW women)
- Racial disparities generally persist through gestational ages, but are highest at early ages. Non-Hispanic blacks experience elevated risk of all outcomes across the full gestational period

Final thoughts

- The use of formal demographic models to study population outcomes is a dynamic process
- New questions, new data, traditional (but updated) methods
- Demography rediscovers its core
- Future work?
    - Extending ${}_1a_0$ work to LMIC/high-mortality contexts (challenging, because there are less available data and data that do exist may have measurement errors that mask true relationships)
    - Using multi-decrement approach to study and adjust for data with definitional issues

Thanks!

monica.alexander@utoronto.ca

monicaalexander.com

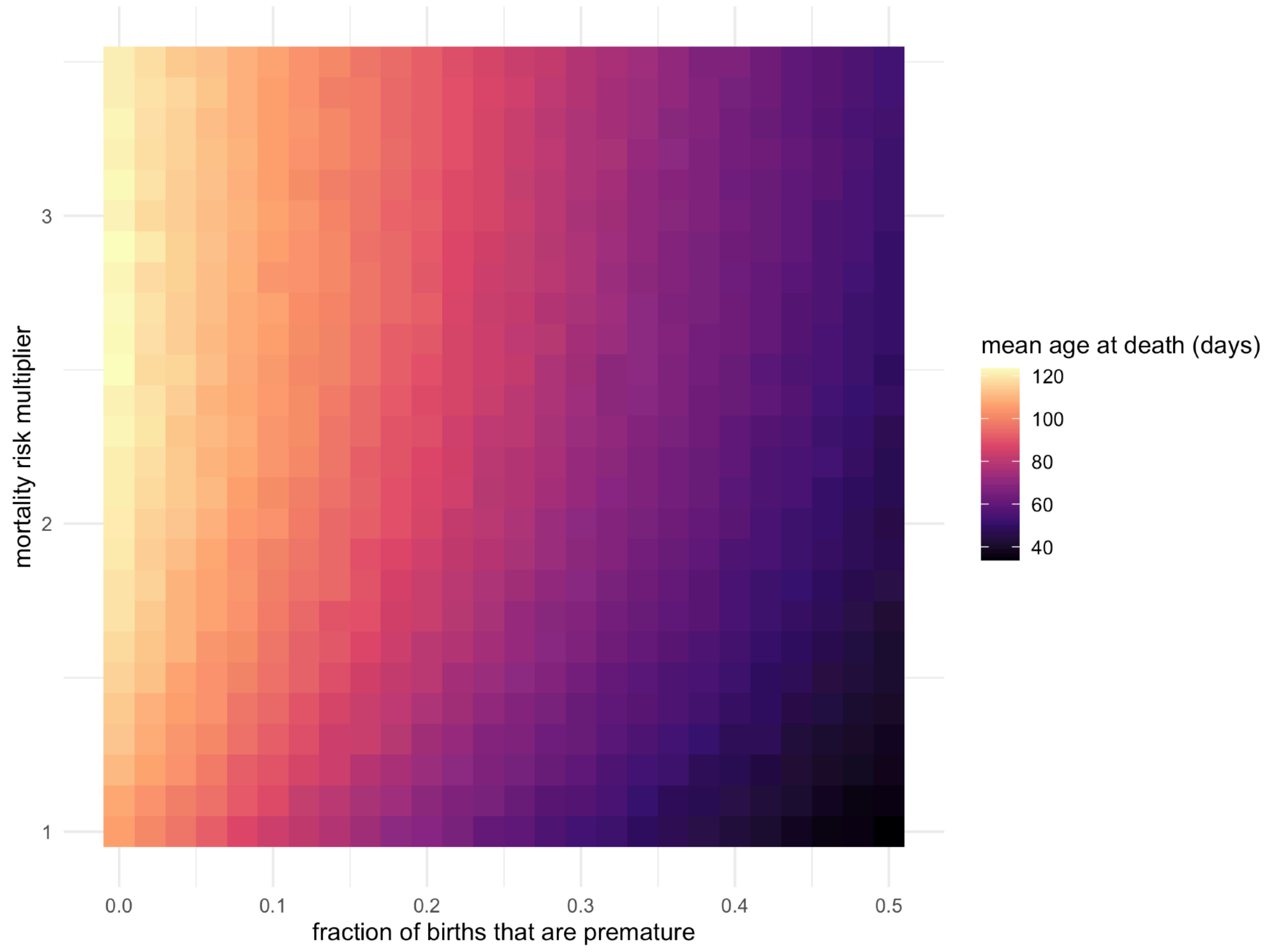
@monjalexander

MJAlexander

Extra

Simulation

- *Scenario 1*: Vary the risk of infant mortality of a population, holding the rate of prematurity constant.
- *Scenario 2*: Vary the share of premature births in a population, holding mortality risk constant.
- The trajectory of ${}_1a_0$ over time depends on the relative changes across the two dimensions



Fitting and evaluation

- Fit the ratio and piecewise models to training, test, and AK datasets
- All models fit in Stan (code: https://github.com/MJAlexander/a0-competing-effects)
- Models evaluated based on
  - Root mean squared error (RMSE), both in-sample and out-of-sample
  - Out-of-sample RSME on test dataset and US data
  - Approximate leave-one-out cross validation (LOO-CV), which gives an estimate of the expected log predictive density (ELPD)