Competing effects on the average age of infant death

Monica Alexander* Leslie Root†

Abstract

In recent decades, the relationship between the average length of life for those who die in the first year of life — the lifetable quantity ${}_1a_0$ — and the level of infant mortality, on which its calculation is often based, has broken down. The very low levels of infant mortality in the developed world correspond to a range of ${}_1a_0$ quantities. We illustrate the competing effect of falling mortality and reduction in preterm births on ${}_1a_0$, through two populations with very different levels of premature birth — infants born to non-Hispanic white mothers and to non-Hispanic black mothers in the United States. Through simulation, we further demonstrate that falling mortality reduces ${}_1a_0$, while a reduction in premature births increases it. We use these observations to motivate the formulation of a new approximation formula for ${}_1a_0$ in low-mortality contexts, which is a function of both the infant mortality rate and the ratio of infant to under-five mortality. Model results and validation show that this model outperforms existing alternatives.

*Departments of Statistical Sciences and Sociology, University of Toronto. monica.alexander@utoronto.ca.

†Department of Demography, University of California, Berkeley. leslie.root@berkeley.edu.

Both authors contributed equally to this work.

1 Introduction

The average length of life for those who die in the first year of life, ${}_1a_0$, is an important lifetable quantity, the first building block of the calculation of person-years lived that ultimately sums to the expectation of life at birth. However, the data required to calculate ${}_1a_0$ exactly are often not readily available, and as such, producers of lifetables usually rely on empirical relationships between the overall level of infant mortality and average age at death to calculate an approximate ${}_1a_0$. The most common of these relationships, the Coale-Demeny and Keyfitz-Flieger formulas, rely on the general rule that as infant mortality falls, deaths become increasingly concentrated early in the interval, so ${}_1a_0$ also falls (Coale, Demeny, and Vaughan 1983; Keyfitz and Flieger 1971). The Keyfitz-Flieger formula, for example, which is referenced in central demographic textbooks (e.g. Wachter 2014), is ${}_1a_0 = 0.07 + 1.7{}_1M_0$, where ${}_1M_0$ is the infant mortality rate.

This relationship can be explained with reference to Bourgeois-Pichat's theory (1951b, 1951a), which assumes all infant deaths can be categorized as arising from either endogenous or exogenous causes, with the latter tending to occur later in the first year. As overall mortality goes down, exogenous causes are increasingly eliminated, and the distribution of infant deaths shifts towards the endogenous causes, which are more likely to cause death relatively soon after birth.

However, at the very low levels of infant mortality currently observed in the developed world, the monotonic relationship between ${}_1a_0$ and infant mortality is no longer the case, and in fact, ${}_1a_0$ has been rising in many countries since the 1970s. Andreev and Kingkade (2015) attribute this shift to medical advances that reduce very early deaths due to congenital conditions and conditions of prematurity. Furthermore, they note that the relationship between level of mortality and ${}_1a_0$ becomes comparatively weak at these low levels of mortality. In this paper, we investigate the reason for this, and illustrate two competing effects on ${}_1a_0$: the overall level of infant mortality, and the share of births that are premature.

Recent work on age patterns in early-life mortality has not taken premature birth into consideration as a determinant of either the level or the shape of infant mortality (e.g. Mejía-Guevara and Tuljapurkar 2019; Mejía-Guevara et al. 2019; Galley and Woods 1999). This is no doubt partly because of the lack of availability of data on premature birth (Blencowe et al. 2012). However,

premature birth is clearly important in describing infant mortality patterns; despite advances, conditions of prematurity still play an outsize role in infant death (Callaghan et al. 2006). And because prematurity is difficult to predict and prevent, and is correlated with a number of individual-level factors including maternal age, health and behavior, as well as socioeconomic status and race, it may vary substantially both among and within populations with relatively low infant mortality (Tucker and McGuire 2004; Purisch and Gyamfi-Bannerman 2017). These two facts taken together indicate that an ideal model of infant mortality would include information on the preterm birth rate.

In this paper, we first illustrate the competing effect of mortality and prematurity rates on the average age of infant death, ${}_1a_0$, through an example using data from two U.S. populations with very different levels of premature birth – infants born to non-Hispanic white mothers and infants born to non-Hispanic black mothers. We then perform a simulation exercise, calculating ${}_1a_0$ at a wide range of mortality and prematurity levels, to further illustrate these competing effects. We then utilize these observations in developing a new formula for ${}_1a_0$, which aims to account for not only the level of infant mortality, but also the level of prematurity, using as a proxy the ratio of infant to under-five mortality.

2 The relationship between ${}_1a_0$, infant mortality, and prematurity

Although infant mortality relates to all deaths in the first year, the distribution of these deaths over time is far from uniform. Figure 1 shows the distribution of infant deaths in the United States in 2012. The distributions have been plotted separately for pre- and full-term births, where preterm births are defined as those occurring before a gestational age of 37 full weeks from the last menstrual period.

Irrespective of the prematurity of births, the largest share of infant deaths occurs in the first several days. However, as Figure 1 shows, the distribution is particularly skewed and concentrated for preterm births, with over 60% of infant deaths occurring within the first five hours.

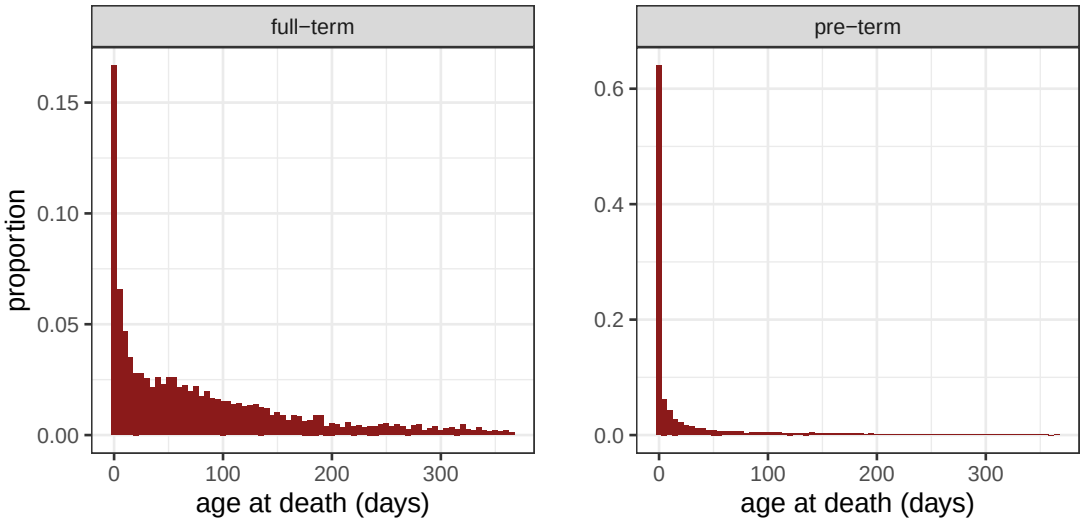

Figure 1: Proportion of infant deaths by five day intervals, United States, 2012 birth cohort. Data via National Center for Health Statistics National Vital Statistics System. Note differing y-axis scales.

In terms of the relationship between overall infant mortality, prematurity and the distribution of death times, the following general observations can be made:

1. The distribution of the timing of infant deaths is left skewed, with the majority of deaths in the first few days (Figure 1 above).
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 (Figure 1 above).
4. The share of births that are premature tends to decrease as mortality decreases (WHO 2018).

Why does the relationship between ${}_1a_0$ and infant mortality become unclear at lower mortality levels? Because statements 1 & 2 above have the opposite effect from statements 3 & 4 on ${}_1a_0$. As overall mortality conditions improve, we expect that exogenous causes of death that occur later in the first year of life decrease, and so ${}_1a_0$ will decrease. However, as overall mortality conditions improve, we also expect the share of births that are premature to decrease, and so ${}_1a_0$ will increase. These observations imply that trends in ${}_1a_0$ over time may go up or down; and that the degree

of similarity between populations' infant mortality rates does not necessarily imply any particular degree of similarity between their ${}_1a_0$ levels (and vice versa).

These competing effects are illustrated in this section; first, by examining infants deaths in the United States by race, and second, by simulating the effects on ${}_1a_0$ of changes in overall mortality risk and the prevalence of prematurity.

2.1 Example: U.S. infant deaths by race

Racial disparities in U.S. infant mortality are longstanding and well-known. Although mortality has fallen for all racial and ethnic groups, babies born to non-Hispanic black mothers have for over thirty years consistently died at more than twice the rate of those born to non-Hispanic white mothers (Ely and Driscoll 2019; Hummer et al. 1999; Mathews, MacDorman, and Menacker 2013). Given this substantial inequality, it could be assumed that the average age of infant death, ${}_1a_0$, would be noticeably higher for the black population than the white population. However, as illustrated in this section, this is not the case. The racial disparity in infant mortality is not simply one of magnitude; patterns of preterm birth, low birth weight, age at death and cause of death all differ substantially between these two groups (Bediako, BeLue, and Hillemeier 2015; Ely and Driscoll 2019; MacDorman 2011; Riddell, Harper, and Kaufman 2017), and the interaction of these factors causes surprising patterns in ${}_1a_0$.

2.1.1 Data

Data on live births and deaths in the first year of life for infants born to non-Hispanic black (NHB) and non-Hispanic white (NHW) mothers comes from the National Bureau of Economic Research collection of U.S. Birth Cohort Linked Birth and Infant Death Data of the National Center for Health Statistics' National Vital Statistics System, years 2008-2012 (NBER 2020). Cohort prematurity rates were calculated as the number of premature births divided by the total number of live births, and infant mortality rates were calculated as the number of deaths divided by the number of live births. Mortality was calculated according to race, birth cohort, premature status, and age at death. In line with WHO classification, premature status was grouped into 4 categories

based on the last menstrual period (LMP) measure of gestational age: extremely preterm (born at < 28 full weeks of gestation), very preterm (28 to <32 full weeks of gestation), later preterm (32 to <37 full weeks of gestation) and full-term (37 full weeks of gestation or more) (WHO 2018). Age at death was split into first-week (<7 days old at death), neonatal (<28 days; includes first-week deaths), and post-neonatal (28 days or older at death). An average of 3021 births (0.1% of the total) and 147 deaths (0.8% of the total) were excluded each year due to missing gestational age data.

2.1.2 Mortality, prematurity, and average age at death

Infant mortality rates for NHB and NHW populations in 2012, stratified by age at death and prematurity of birth, are shown below (Table 1).[1]

Large differences are observed in mortality rates by race that are not stratified by prematurity. Overall mortality is 2.21 times higher for infants born to NHB mothers than for those born to NHW mothers, at 11.00 per thousand versus 4.98 per thousand. When stratifying by age at death, a gradient is evident; racial ratios for mortality are highest in the first week (2.29), somewhat lower for neonatal mortality (2.24), and lowest, though still above 2, for post-neonatal mortality (2.15).

Among premature infants, mortality rates by gestational age are similar for those born to NHB and NHW mothers. For the extremely preterm, the ratio of the NHB mortality rate to the NHW mortality rate is 0.99. For the very preterm, the ratio is 1.00, and for later preterm infants, it is 1.04. For full-term births, however, NHB mortality is 1.70 times higher than NHW mortality.

		by age at death			by gestational age at birth			
	total	first week	neonatal	post-neonatal	extremely pre-term	very pre-term	later pre-term	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

Table 1: Infant mortality rates for U.S. children of non-Hispanic black and non-Hispanic white mothers and the ratio between them

This mismatch between racial patterns of mortality by gestational age and racial patterns by age at death is compositional; it is explained by a large difference in the distribution of births by

[1]Because there is no large time trend in racial difference – that is, infant mortality is declining in a similar way for both racial groups over this period – figures are given for the most recent birth cohort for which data are available (2012). Figures for four previous cohorts may be found in Appendix A.

gestational age, shown in Table 2. Fewer infants born to NHB mothers are born at full term, and among those born early, they are more likely than those born to NHW mothers to be extremely or very preterm. Because mortality risk drops rapidly with gestational age, this means that a larger share of infants born to NHB mothers are at high risk of dying.

	Births (rates per thousand)			
	extremely preterm	very preterm	later preterm	full-term
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

Table 2: Rates of preterm and full-term birth per 1000 U.S. live births to non-Hispanic black and non-Hispanic white mothers and the ratio between them

This compositional difference, in turn, leads to an instance of Simpson's paradox in the average age of infant death, wherein a trend observed in aggregate reverses when data are decomposed into subgroups. In aggregate, infants born to NHB mothers have a slightly lower average age at death, 40.87 days versus 45.29 days for infants of NHW mothers. This is somewhat surprising, given NHB infants' higher mortality: as described above, a shorter ${}_1a_0$ is generally associated with lower overall mortality and a lower share of preventable mortality, as deaths later in infancy are more likely to be caused by external factors and treatable diseases (Andreev and Kingkade 2015).

And indeed, within each subgroup by gestational age at birth, the average age at death is higher for NHB infants, indicating a greater proportion of deaths later in infancy. The difference is marked for all premature infants, regardless of gestational age; among full-term births, NHB infants' ${}_1a_0$ is only 3% longer than NHW infants'. (Table 3)

	mean age at death (days)				
	aggregate	extremely preterm	very preterm	later preterm	full term
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

Table 3: Mean age at death (days) by gestational age at birth for U.S. infants born to non-Hispanic black and non-Hispanic white mothers and the ratio between them

2.2 Illustrating the competing effects through simulation

In order to further investigate the competing effects on ${}_1a_0$, we performed a simulation exercise that allows us to change the overall level of mortality and prematurity independently, and then assess the consequent change in ${}_1a_0$. Specifically, we carry out two simulations:

- 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 coonstant.

In order to be able to simulate plausible times of infant deaths, we need a suitable expression for the distribution of infant deaths. As illustrated in Figure 1, the distribution of infant deaths at the low levels of mortality we are interested in is highly skewed, with a large proportion of deaths occurring in the first week. The shape of these distributions is not readily captured by any classic parametric distributions. However, we found that the shape of infant death distributions was well-captured by a piecewise constant hazard (PCH) model, with time intervals partitioned at days 1-7, 14, 28, 60, 90, 180, and 365. The PCH model assumes constant exponential hazards within each of these time intervals, which allows the model to be estimated using Poisson regression.

PCH models were fit separately to all full-term births/deaths and all preterm births/deaths using the 2012 U.S. births and deaths data. Once estimates of hazards were obtained, we simulated different sets of survival times based on varying: 1) the overall infant mortality risk factor and 2) the overall share of births that are premature. Details on the statistical model and simulation can be found in Appendix C.

2.2.1 Results

The results of these simulations are shown in Figures 2 and 3. The results clearly show that changes in overall infant mortality and changes in prematurity act in opposite directions. For increases in the risk profile of mortality, the average age of infant death steadily increases until a risk factor of around 2, when the estimated ${}_1a_0$ plateaus. This plateau is a consequence of the shape of the death

distribution, with increases in later infant mortality being balanced out by the increases in high hazards in the earlier-infant mortality. For increases in the share of premature deaths (Figure 3), the average age of infant death steadily declines. It is worth noting that the magnitude of changes in ${}_1a_0$ in response to changes in prematurity are larger than those based on increasing the overall risk of mortality. This is broadly consistent with the pattern observed in the U.S. data; faced with both higher rates of prematurity and a higher overall risk of mortality, NHB have a shorter ${}_1a_0$ than NHW infants, not a longer one.

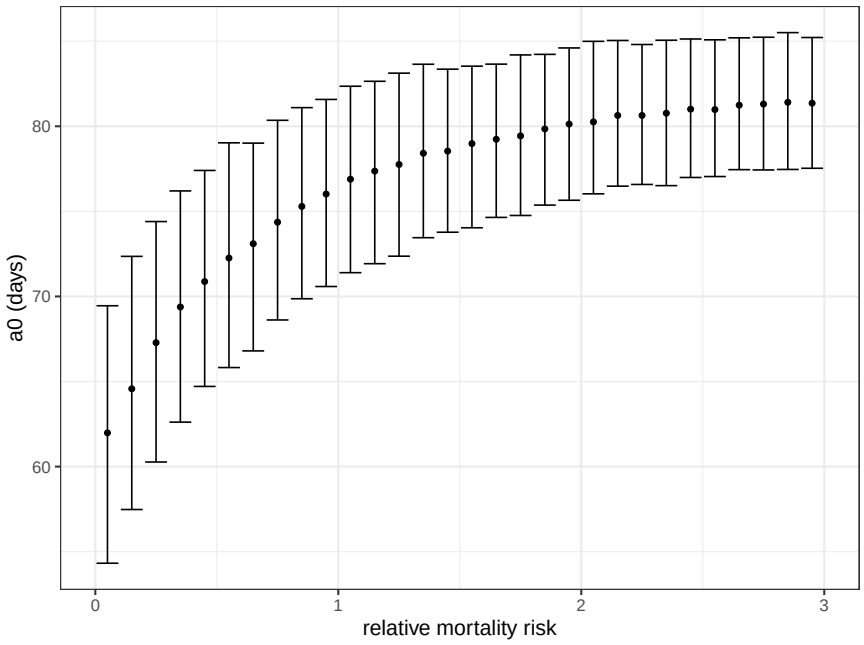

Figure 2: Estimated ${}_1a_0$ and 95% confidence intervals for different risk factors. Results from 1,000 simulations of 10,000 births with baseline hazards fitted to the U.S. NHB population in 2012.

Figure 4 combines variability across prematurity and mortality risk to illustrate changes in ${}_1a_0$ in two dimensions. The lighter the color, the higher the value of ${}_1a_0$. While ${}_1a_0$ increases monotonically with increased mortality risk, and decreases monotonically with increased prematurity, the trajectory of ${}_1a_0$ over time depends on the relative changes across the two dimensions.

3 A new approximation formula for ${}_1a_0$ in low-mortality conditions

We have illustrated that the rate of premature birth appears to play an important role in ${}_1a_0$ at low rates of infant mortality, and an ideal model for calculating ${}_1a_0$ would include information on both

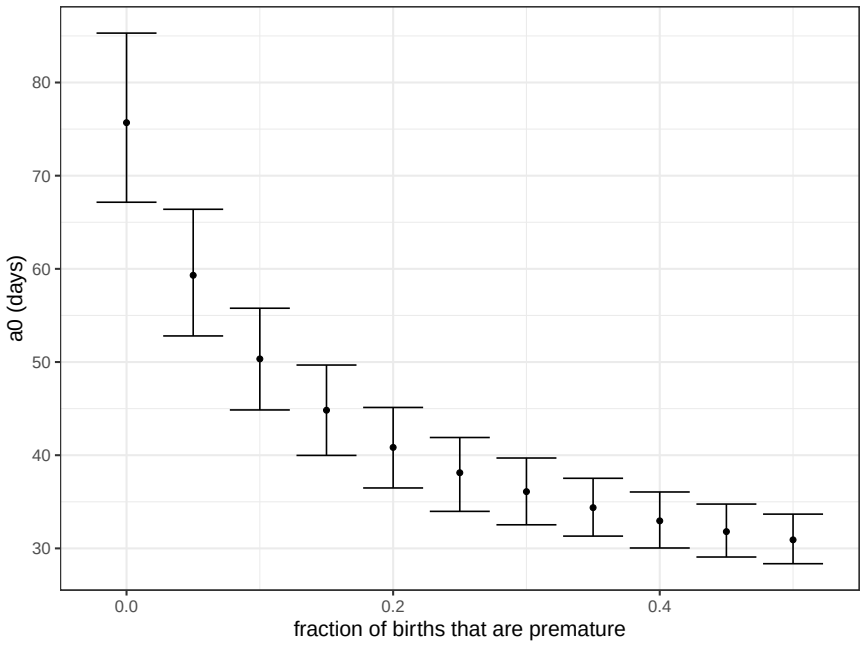

Figure 3: Estimated ${}_1a_0$ and 95% confidence intervals for different shares of premature births. Results from 1,000 simulations of 10,000 births with baseline hazards for pre- and full-term births fitted to the U.S. population in 2012.

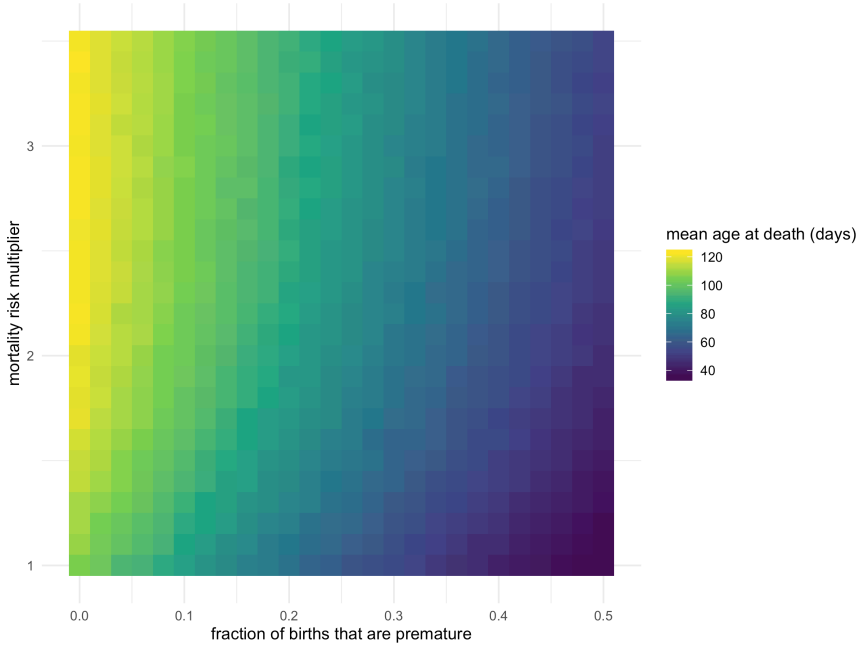

Figure 4: Competing effects on average age of infant death. Horizontal axis shows increasing prematurity, vertical axis shows increasing overall mortality risk

infant mortality and prematurity. Unfortunately, with the exception of U.S. microdata, detailed premature birth data is not widely available (Blencowe et al. 2012).

As an alternative measurement, we argue that, at these low levels of mortality, the ratio of early (i.e. infant) to later (i.e. under-five) child mortality proxies the early vs. late pattern within infant mortality. That is, at a given level of infant mortality, a higher level of overall mortality for those under age five signals that a relatively larger share of infant mortality is due to exogenous causes, while a lower level of overall under-five mortality signals the opposite – that infant mortality is more skewed toward endogenous causes arising from premature birth. This relationship appears to be borne out empirically, as shown in Figure 5, which illustrates a clear negative relationship between the average age of infant death and the ratio of infant to under-five mortality. Using the ratio as a explanatory variable within a model for ${}_1a_0$ is advantageous as it is easily calculable for all countries based on existing data and estimates. For instance, the United Nations Inter-agency Group on Mortality Estimation (IGME) produces infant mortality and under-5 mortality estimates for all member countries.[2]

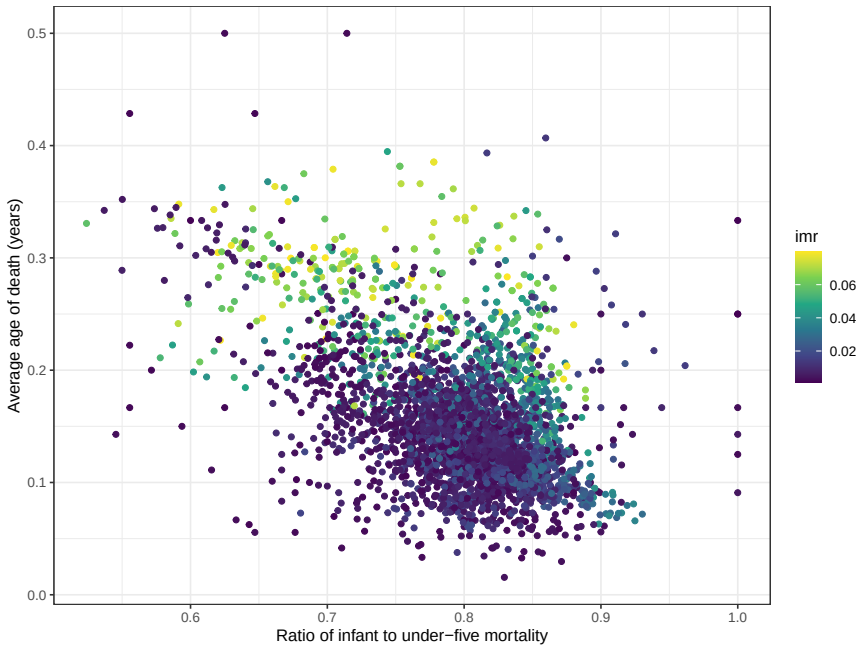

Figure 5: Relationship between the ratio of infant to under-5 mortality and ${}_1a_0$ in low-mortality conditions.

In this section, we outline a new proposed model to estimate ${}_1a_0$, the data we used for model

[2] Available at childmortality.org.

building and evaluation, and our evaluation and validation strategy. In brief, we compare the performance of our new proposed model to the existing best alternative, a piecewise linear model, as proposed by Andreev and Kingkade (2015), and show that it outperforms this model on several different model evaluation metrics.

3.1 Proposed model

We propose a linear model in which the ${}_1a_0$ is related to the infant mortality (${}_1q_0$) rate and the ratio of infant to under-5 mortality (U5MR):

$$ {}_1a_{0_i} = \alpha + \beta_1({}_1q_{0_i}) + \beta_2 \left(\frac{{}_1q_{0_i}}{\text{U5MR}_i} \right) + \varepsilon_i \tag{1} $$

For reference, existing models for ${}_1a_0$ have the form of a piecewise linear relationship with ${}_1q_0$ with one 'cutpoint' (θ) at which the slope of the linear relation with ${}_1q_0$ changes.[3]

$$ {}_1a_{0_i} = \alpha + \beta_1({}_1q_{0_i}) + \beta_2({}_1q_{0_i} - \theta) + \varepsilon_i \tag{2} $$

One of the main advantages of the proposed model over previous models is that it explicitly accounts for the two competing effects on ${}_1a_0$, by incorporating additional readily-available data. In contrast, existing methods model ${}_1a_0$ as a mathematical function of infant mortality rate using a set of arbitrary splines, without taking into consideration the demographic or biological reasons for why the relationship between ${}_1a_0$ and infant mortality may change with the level of mortality.

3.2 Data

This analysis draws on data compiled by the Human Mortality Database (HMD), a joint project of the University of California, Berkeley (USA), and the Max Planck Institute for Demographic Research (Germany) that collects, validates, cleans and adjusts detailed data on period and cohort mortality, available online at https://www.mortality.org. It currently includes 41 countries with

[3]Note that Andreev and Kingkade (2015) define two cutpoints, the second being the level of ${}_1q_0$ at which ${}_1a_0$ is estimated to be constant. However, as we are focusing on low levels of mortality, we are more interested in comparisons just in the range of the first cutpoint.

reliable mortality statistics. HMD data does not allow for direct observation of ${}_1a_0$, but it can be calculated using the Lexis triangle method detailed by Andreev and Kingkade (2015). Following Andreev and Kingkade, we use the initial raw numbers of deaths by Lexis triangle in these calculations, avoiding any adjustment which could bias the estimate of ${}_1a_0$. With some exclusions, this analysis makes use of raw Lexis triangle death data from all countries and cohorts currently available through HMD.[4] Excluded are: Bulgaria (before 2009), Estonia (before 1992), and Belarus, Russia, and Ukraine (entirely), for incomplete conformity to the WHO definition of 'live birth', which leads to an overcount of stillbirths and reduction in first-day infant mortality; Chile and Taiwan, for suspected problems with the registration of very early infant deaths; Iceland and Luxembourg, for very low numbers of infant deaths resulting in unusual ${}_1a_0$ or mortality rates; the Netherlands (before 1950), because of data adjustments made to account for differences in the definition of live birth; Switzerland (before 1880) for irregularities implying unusual migration patterns at age zero; and Belgium (1941-1945 and 1958-1960), Israel, and Poland (entirely), for implausible mortality rates or ${}_1a_0$ estimates. Populations whose inclusion would result in duplication of data were also excluded (e.g., because East Germany and West Germany are included separately, the aggregate data available for Germany as a whole is not included). In addition, an infant mortality cutoff of 80 per thousand for females and 95 per thousand for males was applied, as this analysis focuses on low-mortality contexts. This yielded 2852 ${}_1a_0$ observations from 27 countries or areas, for cohort years ranging from 1902 to 2018, broken down by sex (1430 female observations and 1422 male). A table listing all included populations can be found in Appendix A.

3.3 Modeling strategy

Our primary goal is to estimate parameters in Equation 1, which we will refer to as the 'ratio model', and compare the performance of this model to the existing best alternative proposed by Andreev and Kingkade. We estimated model parameters for both the ratio model and equation 2, i.e., a piecewise linear model, in a Bayesian framework using a Hamiltonian Monte Carlo algorithm, implemented using the Stan programming language in R (Carpenter et al. 2017).[5]

[4] Data downloaded May 25, 2020.

[5] We chose to estimate models using Stan because of two main reasons; firstly, uncertainty intervals (Bayesian credible intervals) are a natural by-product of the estimation process (unlike other optimization options that are usually required to estimate the position of the cutpoints), and secondly, the ability to calculate approximations to

To evaluate model performance, we split the available data into two separate datasets: the 'training' dataset, which comprises a random sample of 80% of all the available data, and a 'test' dataset, which comprises of the remaining 20% of data. The idea is that the estimates of the parameters in the ratio and piecewise models are obtained by fitting models on the training dataset, and then the predictive accuracy of each model is evaluated based on the test dataset. This train/test splitting technique, which is common in statistical learning methods, minimizes issues of over-fitting and better allows for the evaluation of out-of-sample fit (James et al. 2013).

Our modeling strategy introduces two comparison issues: differences in data, and differences in estimation methods. Firstly, we have at our disposal substantially more data from HMD than was available five or more years ago, which means parameter estimates for a piecewise linear, as was estimated by Andreev and Kingkade, may differ purely because of data reasons. Secondly, our choice to estimate models within a Bayesian framework using Stan also differs from Andreev and Kingkade's analysis. To account for these differences, in addition to estimating the parameters for thr piecewise linear model on the training dataset, we also obtain estimates on a dataset which corresponds to the data used by Andreev and Kingkade (2015) (termed the 'AK' dataset). In doing so, we show that our estimates for a piecewise linear model fit to AK data do not differ significantly from Andreev and Kingkade's estimates.

To summarize, we fit two models (the ratio model and the piecewise linear model, Equations 1 and 2). We estimated the ratio model and piecewise linear model by sex and also for both sexes combined. Both the ratio model and piecewise linear model were estimated on the training datasets (male, female and both sexes combined) and the AK dataset. All models, including the estimates presented by Andreev and Kingkade (2015), were then evaluated on several metrics, as discussed below, including their predictive performance based on the test datasets.

3.4 Evaluation metrics

In order to compare candidate models for ${}_1a_0$ and evaluate their performance, several different evaluation metrics were used. First, we calculated the mean squared error (RMSE), which is defined as

the leave-one-out cross validation using Stan allows us to easily compare and evaluate models.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} \left({}_1a_0(i) - {}_1\hat{a}_0(i)\right)^2}{N}}$$

where ${}_1a_0(i)$ is the ith observed ${}_1a_0$ value and ${}_1\hat{a}_0(i)$ is the corresponding fitted value for all N observations $i = 1, \ldots, N$.

The RMSE was calculated on three different data sets for each of the three models (the ratio, piecewise, and AK model) under consideration:

1. The **training** data set, i.e. the HMD data on which the model was fitted, which comprises all available data (as described above) with 20% excluded. This is referred to the **in-sample RMSE** and reflects the goodness of model fit within the dataset that was used for estimation.
2. The **test** data set, which is the 20% of HMD data that was excluded above. This is referred to the **out-of-sample RMSE** and reflects the model fit on a 'new' dataset.
3. The **Andreev and Kingkade (AK)** data set, i.e. the set of HMD data that was used by Andreev and Kingkade (2015).[6] This was included to ensure that the piecewise estimates in Andreev and Kingkade (2015) and our piecewise estimates were as comparable as possible.

In addition to evaluating models based on in- and out-of-sample RMSE, we also used approximate leave-one-out cross-validation (LOO-CV) to assess relative model performance of the piecewise linear and the ratio model. LOO-CV refers to the technique of leaving out one data observation (the ith data point), refitting the model to the new dataset with everything except point i, and evaluating the ability of the model to predict point i. This process can be repeated N times, leaving out each of the i points and then assessing predictive power each time. In practice, fitting each model N times is computationally inefficient, and so we used approximate LOO-CV using Pareto smoothed importance sampling (Vehtari, Gelman, and Gabry 2017). This was done using the `loo` package in R (Vehtari et al. 2020). Approximate LOO-CV gives an estimate for the expected log predictive density (ELPD). The higher the ELPD, the better the model fit.

[6] As discussed above, we used similar criteria for data inclusion as Andreev and Kingkade. However, a great deal of new data has been added to HMD since Andreev and Kingkade conducted their analysis, in early 2011, so their data set comprises only a subset of our data.

4 Results

Table 4 shows the estimated coefficients for Equations 1 and 2 for various sex and data combinations. The first two rows show the estimates from Andreev and Kingkade (2015) for reference. The third and fourth rows show the results of estimation of the piecewise model fit to a dataset that is the same as was used by Andreev and Kingkade (2015). These results suggest that our estimates are very similar, which is encouraging given our estimation approach is quite different. The next three rows show results of the same piecewise model that has been fitted to all available HMD data. Notice that we also estimate this model for both sexes combined; this was motivated by the fact that, based on looking at the 95% uncertainty bounds of the sex-specific estimates, there is generally no significant difference between estimates for males and females. The coefficient estimates are slightly different to the AK data case, which perhaps is not surprising given the addition of many more data points at increasingly lower mortality conditions. The comparison between the piecewise coefficients is also shown in Figure 6.

Model	Equation	Sex	$\hat{\alpha}$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\theta}$
AK	2	f	0.149	-2.0867	4.1075	0.017
AK	2	m	0.1493	-2.0367	3.4994	0.0226
piecewise (AK data)	2	f	0.149 [0.14,0.155]	-1.984 [-2.82,-1.093]	4.856 [3.87,5.695]	0.014 [0.01,0.016]
piecewise (AK data)	2	m	0.15 [0.14,0.158]	-2.162 [-2.87,-1.588]	4.647 [4.03,5.41]	0.019 [0.02,0.021]
piecewise	2	f	0.138 [0.13,0.143]	-0.913 [-1.59,-0.328]	4.249 [3.62,4.976]	0.017 [0.01,0.019]
piecewise	2	m	0.14 [0.14,0.145]	-1.268 [-1.74,-0.883]	4.478 [3.99,4.975]	0.023 [0.02,0.025]
piecewise	2	b	0.139 [0.14,0.142]	-1.004 [-1.38,-0.647]	4.297 [3.92,4.708]	0.021 [0.02,0.022]
ratio	1	f	0.405 [0.37,0.436]	1.975 [1.87,2.084]	-0.359 [-0.4,-0.322]	-
ratio	1	m	0.42 [0.39,0.454]	1.551 [1.43,1.669]	-0.382 [-0.42,-0.341]	-
ratio	1	b	0.426 [0.4,0.447]	1.749 [1.67,1.834]	-0.387 [-0.41,-0.359]	-

Table 4: Coefficient estimates (and 95% credible intervals) for all model combinations

The last three rows of Table 4 show the coefficient estimates for females, males and both sexes combined for the ratio model (shown in Equation 1). The sign of the coefficient estimates of the ratio model suggest that as infant mortality rises, so does ${}_1a_0$, but as the fraction of under-five mortality that is attributed to infant mortality rises, ${}_1a_0$ falls. The coefficient estimates on infant mortality ($\hat{\beta}_1$) are 1.5-1.9, which suggests that for every 1 per 100 increase in infant mortality, ${}_1a_0$ will increase by 0.015-0.02 years, or somewhere between 5 to 7 days, holding the ratio constant. The coefficient estimates on the ratio ($\hat{\beta}_2$) suggest that, as the ratio increases by 0.1, ${}_1a_0$ decreases by 0.03-0.04 years, around 11-15 days, holding infant mortality constant.

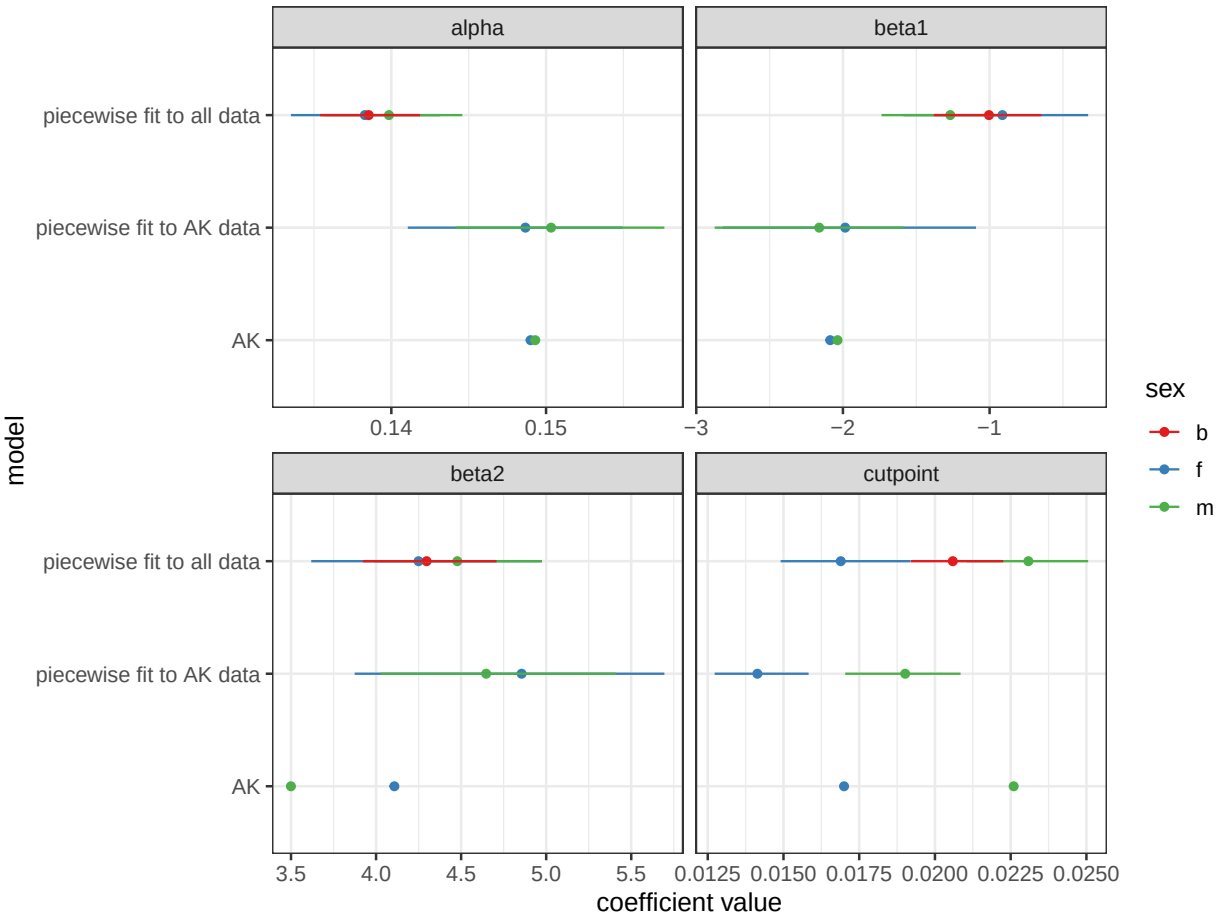

Figure 6: Coefficient estimates and 95% credible intervals for all piecewise model combinations.

4.1 Model evaluation

Table 5 shows in-sample and out-of-sample root mean squared error (RMSE) for each model, for each sex and for both sexes together. The smaller the RMSE, the better the model fit. Note that for the estimates from AK data, we can only calculated RMSE for either males or females, because estimates were not provided for both sexes combined. In general, the ratio model RMSEs are the smallest for all sexes, suggesting the fit of this model is superior to both the training (in-sample) and test (out of sample) data sets. The RMSE 'gain' going from the piecewise model to the ratio model is largest in the out-of-sample contexts, suggesting that the ratio model performs relatively well in predicting ${}_1a_0$ based on new data.

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
insample	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	b	-	0.038	0.037
out of sample	b	-	0.041	0.038

Table 5: In- and out-of-sample root mean squared error for different models

Finally, we can compare the expected log predictive density (ELPD) for the piecewise model and ratio model across all sex combinations. The higher the ELPD, the better the model's predictive ability. Table 6 shows the difference in ELPD (piecewise minus ratio) and the standard error for that difference for all sex combinations. The results suggest that in all cases, the ratio model has a higher ELPD (i.e. the difference is negative), although the difference is not significant in the male case.

Sex	ELPD difference	SE difference
f	-54.27	20.47
m	-10.96	21.67
b	-83.98	30.89

Table 6: ELPD difference (piecewise - ratio) and standard error of difference for all sex combinations

5 Discussion

This paper increases our understanding of how and why length of life varies with other forms of inequality. We illustrate that higher infant mortality tends to raise ${}_1a_0$, while a rising share of births that are premature tends to lower it. Thus, at a given level of infant mortality, a population with more premature birth will have a lower ${}_1a_0$. This paradox explains the observed similarity of ${}_1a_0$ among U.S.-born children of black and white non-Hispanic mothers; while children of black mothers have higher mortality overall, thus implying a longer ${}_1a_0$, they also have higher rates of prematurity, implying a shorter one. In the U.S. case, the prematurity disadvantage rate outweighs the mortality disadvantage, resulting in a shorter ${}_1a_0$ for infants born to black mothers. We believe this is an important descriptive finding; it illustrates that the classic association of shorter ${}_1a_0$ with better mortality conditions does not hold at low levels of mortality, where the role of prematurity is greater.

We used this descriptive finding to motivate the formulation of a new approximation for ${}_1a_0$ in low mortality conditions. As data on the prevalence of prematurity is not readily available for all countries, we use the ratio of infant to under-5 mortality as a proxy, and model ${}_1a_0$ as a function of overall infant mortality and the ratio. We showed that this model outperforms the existing best alternative, a piecewise model which is a linear function of infant mortality with a cutpoint of changing slope. As well as illustrating an improvement in a purely statistical sense, we believe this model is an improvement in because it is based on substantive reasoning rather than a mathematical function.

The data and model estimation presented in this paper considered the low mortality context (i.e. infant mortality rates of less than 80 per 1,000), as data from HMD mostly covers low-mortality populations. Future work could explore the relationship between ${}_1a_0$, infant mortality and the ratio of infant to under-five mortality in a range of different mortality contexts, using, for example, data from Demographic and Health Surveys as well as chld mortality estimates from UN-IGME.

Practically speaking, small improvements in the calculation of ${}_1a_0$ are of relatively little importance for the life table (e.g., for the calculation of e_0); especially at low levels of mortality, the contribution of person-years lived by those who die within the first year of life to total person-years lived is very

small, and the fact that a different ${}_1a_0$ formula adjusts their modeled average length of life by mere days makes the effect smaller still. For example, using the Centers for Disease Control's 2008 U.S. male period life table (Arias 2012), which gives an infant mortality rate of seven per thousand, changing the given ${}_1a_0$ of 47.45 days by a full week in either direction changes e_0 by a little over one ten-thousandth of a year, or about an hour and twelve minutes.

Nonetheless, accuracy is desirable, and perhaps more importantly, this framework allows us to explain previously unaccounted-for variation in ${}_1a_0$ at low levels of mortality. The effectiveness of using the ratio of infant to under-five mortality as a proxy for prematurity in this application suggests a possible avenue for exploring a phenomenon – the premature birth rate – that has large implications for population health but for which we have very little data. It also suggests that if we had more data on premature birth, ${}_1a_0$ could be modeled with even greater accuracy.

A Mortality and prematurity 2008-2012

The tables below show mortality rates and prematurity rates for the non-Hispanic white and black populations for the United States for the years 2008–2012.

	2008	2009	2010	2011	2012
	infant mortality rate				
black	12.34	12.03	11.32	11.07	11.00
white	5.40	5.19	5.06	4.98	4.98
ratio	2.29	2.32	2.24	2.22	2.21
	first-week mortality rate				
black	6.48	6.33	5.85	6.03	5.94
white	2.71	2.59	2.56	2.57	2.60
ratio	2.39	2.44	2.28	2.35	2.29
	neonatal mortality rate				
black	8.07	7.93	7.28	7.43	7.30
white	3.44	3.32	3.29	3.25	3.26
ratio	2.35	2.39	2.21	2.29	2.24
	post-neonatal mortality rate				
black	4.27	4.09	4.04	3.64	3.70
white	1.96	1.87	1.77	1.74	1.72
ratio	2.18	2.19	2.29	2.10	2.15

Table 7: Mortality rates by age at death, race, and year

	2008	2009	2010	2011	2012
	extremely preterm				
black	398.48	384.46	366.24	362.18	363.18
white	387.33	376.18	369.73	369.00	365.32
ratio	1.03	1.02	0.99	0.98	0.99
	very preterm				
black	46.62	45.77	44.10	40.36	38.86
white	40.26	41.40	37.39	41.58	38.90
ratio	1.16	1.11	1.18	0.97	1.00
	later preterm				
black	10.46	9.92	9.67	9.88	9.01
white	8.84	8.18	8.30	8.26	8.65
ratio	1.18	1.21	1.16	1.20	1.04
	full term				
black	3.72	3.64	3.53	3.36	3.54
white	2.29	2.22	2.13	2.11	2.09
ratio	1.63	1.64	1.66	1.60	1.70

Table 8: Mortality rates by premature status, race, and year

	2008	2009	2010	2011	2012
	extremely preterm				
black	17.19	17.43	16.89	16.99	16.79
white	5.41	5.36	5.44	5.24	5.36
ratio	3.18	3.25	3.10	3.25	3.13
	very preterm				
black	21.24	21.25	20.99	20.58	20.32
white	10.58	10.38	10.36	10.21	10.14
ratio	2.01	2.05	2.03	2.02	2.00
	later preterm				
black	137.00	135.98	133.30	130.17	128.15
white	95.39	93.43	91.89	89.56	87.42
ratio	1.44	1.46	1.45	1.45	1.47
	full term				
black	824.57	825.34	828.82	832.26	834.75
white	888.62	890.82	892.30	895.00	897.09
ratio	0.93	0.93	0.93	0.93	0.93

Table 9: Rates of preterm and full-term birth per 1000 live births by race and year

B HMD data used

The table below lists the countries and years included in the model estimation process.

Population	HMD abbreviation	Cohorts
Australia	AUS	1990-2015
Austria	AUT	1971-2016
Belgium	BEL	1956-1957, 1961-2016
Bulgaria	BGR	2009-2016
Canada	CAN	1950-2009
Croatia	HRV	2001-2017
Czech Republic	CZE	1950-2017
Denmark	DNK	1921-2018
England and Wales	GBRTENW	1914-1949
Estonia	EST	1992-2016
Finland	FIN	1917-2017
France (civilian population)	FRACNP	1907-2016
Germany (East)	DEUTE	1956-2016
Germany (West)	DEUTW	1956-2016
Hungary	HUN	1950-2016
Italy	ITA	1915-2016
Japan	JPN	1947-2017
Netherlands	NLD	1850-1998
New Zealand	NZL_NP	1980-2012
Norway	NOR	1980-2017
Portugal	PRT	1980-2017
Slovakia	SVK	1965-2016
Slovenia	SVN	1983-2016
Spain	ESP	1975-2017
Sweden	SWE	1901-2017
Switzerland	CHE	1880-2017
United States of America	USA	1959-2016

Table 10: Human Mortality Database populations and cohorts that meet criteria for inclusion in analysis. Note that some cohorts listed here are excluded based on the infant mortality rate cutoff of 80 per thousand (female) and 95 per thousand (male).

C Details on simulation

C.1 Goals

The goal of the simulation exercise is to see how ${}_1a_0$ varies with changes in 1) the overall mortality risk and 2) the share of births that are premature. In order to do this, we need to first obtain a set of hazards for full-term births a preterm births. We then need to obtain a set of simulated death times based on these hazards under the conditions of 1) and 2).

C.2 Estimating hazards

In order to simulate infant death times under varying mortality risk and prematurity conditions, an expression for the hazard of dying at a particular time t is required. However, the shape of the distribution of infant death times is non-parametric and thus not well represented by standard parametric distributions like the Exponential or Weibull.

As such, in order to obtain estimates for the hazards over time, we used piecewise constant hazard (PCH) models. We assume exponential hazards within each piece, and split the first year at days 1-7, 14, 28, 60, 90, 180, and 365. It can be shown that piece-wise exponential proportional hazards model is equivalent to a Poisson log-linear model where the death indicator is the response and the log of exposure time enters as an offset. We fit these models in `R` using the `glm` function separately for all full-term births and all preterm births in the 2012 U.S. dataset. The resulting regression estimates give the log-hazards in each of the separate intervals, which can then be used for simulation.

C.3 Piecewise exponential survival function

Define the set of time intervals to be $0 = \tau_0 < \tau_1 < ... \tau_m < \tau_{m+1}$. The piecewise constant hazard function is

$$h(t) = h_0 \sum_{l=0}^{m} g_l I_l(t)$$

where h_0 is the baseline hazard, g_l are the relative hazards and $I(t)$ is an indicator function which is equal to 1 if $\tau_l \leq t < \tau_{l+1}$ and 0 otherwise.

The survival function is

$$S(t) = \exp\left(-H(t)\right) = \exp\left(-h_0 \sum_{l=0}^{m} g_l \int_0^t I_l(s)ds\right)$$

The piecewise survival function for interval if $\tau_i \leq t < \tau_{i+1}$ is

$$S_i(t) = \exp\left(-h_0 \left(\sum_{l=0}^{i-1} g_l(\tau_{l+1} - \tau_l) + g_i(t - t_i)\right)\right)$$

Solving for t implies

$$t = \tau_i - \frac{\log(S_i(t))}{h_0 g_i} - \frac{1}{g_i} \sum_{l=0}^{i-1} g_l(\tau_{l+1} - \tau_l) \tag{3}$$

As $\tau_i \leq t < \tau_{i+1}$ this gives us two conditions:

$$\log(S_i(t)) \leq -h_0 \sum_{l=0}^{i-1} g_l(\tau_{l+1} - \tau_l) \tag{4}$$

$$\log(S_i(t)) > \sum_{l=0}^{i} g_l(\tau_{l+1} - \tau_l) \tag{5}$$

C.4 Simulation framework

We have an analytical form for the survival function S, which means there is also an analytical form for the cumulative distribution function $F = 1 - S$. We used inverse transform sampling and equations 1-3 above to generate samples from the death distributions specified by a set of piecewise hazards estimated using Poisson regression. To reiterate, the two simulation scenarios were:

- Scenario 1: Vary the risk of infant mortality of a population with similar infant mortality to the full-term birth population in the U.S. Risk multipliers simulated were between [0.1,3] in increments of 0.1.

- Scenario 2: Vary the share of premature births in a population which has similar hazard profiles of pre- and full-term births in the U.S. Shares considered were between [0, 0.5] in increments of 0.1.

The following simulation process was used:

1. Set the time intervals, the baseline hazard, and all relative hazards. For Scenario 1, the hazards come from the full-term birth estimation. For Scenario 2, both the hazards from full- and preterm births are used. For every iteration in each scenario (i.e. each risk multiplier and each share s)
2. Set the size of the sub-population. For Scenario 1 this was 10,000. For Scenario 2 this was $10,000 * s$ preterm births and $10,000 * (1 - s)$ full-term births.
3. Draw a uniformly (0,1) distributed random variable S = 1 - F. (For Scenario 2, two random variables are drawn.)
4. Determine the right interval using the conditions in Equations 2 and 3.
5. Compute the random time t using Equation 1.
6. Combine the computed random times to one file.
7. Repeat this process 1,000 times.

Once the set of 1,000 simulation times are obtained, the mean age of death ${}_1a_0$ was calculated for each simulation. The mean, 2.5th and 97.5th quantiles of ${}_1a_0$ were then calculated.

References

Andreev, Evgeny M., and W. Ward Kingkade. 2015. "Average Age at Death in Infancy and Infant Mortality Level: Reconsidering the Coale-Demeny Formulas at Current Levels of Low Mortality." *Demographic Research* 33 (13): 363–90. https://doi.org/10.4054/DemRes.2015.33.13.

Arias, Elizabeth. 2012. "United States Life Tables, 2008." Volume 61, Number 3. Hyattsville, Maryland: National Center for Health Statistics.

Bediako, Phylicia T., Rhonda BeLue, and Marianne M. Hillemeier. 2015. "A Comparison of Birth Outcomes Among Black, Hispanic, and Black Hispanic Women." *Journal of Racial and Ethnic Health Disparities* 2 (4): 573–82. https://doi.org/10.1007/s40615-015-0110-2.

Blencowe, Hannah, Simon Cousens, Mikkel Z Oestergaard, Doris Chou, Ann-Beth Moller, Rajesh Narwal, Alma Adler, et al. 2012. "National, Regional, and Worldwide Estimates of Preterm Birth Rates in the Year 2010 with Time Trends Since 1990 for Selected Countries: A Systematic Analysis and Implications." *The Lancet* 379 (9832): 2162–72. https://doi.org/10.1016/S0140-6736(12)60820-4.

Bourgeois-Pichat, Jean. 1951a. "La Mesure de La Mortalité Infantile: II. Les Causes de décès." *Population (French Edition)* 6 (3): 459–80. https://doi.org/10.2307/1523958.

———. 1951b. "La Mesure de La Mortalité Infantile. I. Principes et Méthodes." *Population (French Edition)* 6 (2): 233–48. https://doi.org/10.2307/1524151.

Callaghan, William M, Marian F MacDorman, Sonja A Rasmussen, Cheng Qin, and Eve M Lackritz. 2006. "The Contribution of Preterm Birth to Infant Mortality Rates in the United States." *Pediatrics* 118 (4): 1566–73.

Carpenter, Bob, Andrew Gelman, Matthew D Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. 2017. "Stan: A Probabilistic Programming Language." *Journal of Statistical Software* 76 (1).

Coale, Ansley J., Paul George Demeny, and Barbara Vaughan. 1983. *Regional Model Life Tables and Stable Populations*. 2nd ed. Studies in Population. New York: Academic Press.

Ely, Danielle M., and Anne K. Driscoll. 2019. "Infant Mortality in the United States, 2017: Data

from the Period Linked Birth/Infant Death File." 68: 10. U.S. Department of Health; Human Services, Centers for Disease Control; Prevention.

Galley, Chris, and Robert Woods. 1999. "On the Distribution of Deaths During the First Year of Life." *Population: An English Selection* 11: 35–59. http://www.jstor.org/stable/2998689.

Hummer, Robert A., Monique Biegler, Peter B. De Turk, Douglas Forbes, W. Parker Frisbie, Ying Hong, and Starling G. Pullum. 1999. "Race/Ethnicity, Nativity, and Infant Mortality in the United States." *Social Forces* 77 (3): 1083–1117. https://doi.org/10.2307/3005972.

James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani. 2013. *An Introduction to Statistical Learning*. Vol. 112. Springer.

Keyfitz, Nathan, and Wilhelm Flieger. 1971. *Population: Facts and Methods of Demography*. San Francisco: W. H. Freeman.

MacDorman, Marian F. 2011. "Understanding Racial and Ethnic Disparities in U.S. Infant Mortality Rates," no. 74: 8.

Mathews, T. J., Marian F. MacDorman, and Fay Menacker. 2013. "Infant Mortality Statistics from the 2013 Period Linked Birth/Infant Death Data Set." https://doi.org/10.1037/e558952006-001.

Mejía-Guevara, Iván, and Shripad Tuljapurkar. 2019. "'Early Rectangularization' of Under-5 Mortality in Sub-Saharan Africa: Compression, Convergence, Inequality." *bioRxiv*, March, 591925. https://doi.org/10.1101/591925.

Mejía-Guevara, Iván, Wenyun Zuo, Eran Bendavid, Nan Li, and Shripad Tuljapurkar. 2019. "Age Distribution, Trends, and Forecasts of Under-5 Mortality in 31 Sub-Saharan African Countries: A Modeling Study." *PLoS Medicine* 16 (3). https://doi.org/10.1371/journal.pmed.1002757.

NBER. 2020. "Linked Birth/Infant Death Cohort Data." Available at: https://www.nber.org/research/data/linked-birthinfant-death-cohort-data.

Purisch, Stephanie E., and Cynthia Gyamfi-Bannerman. 2017. "Epidemiology of Preterm Birth." *Seminars in Perinatology*, Current Preterm Birth Prevention Strategies, 41 (7): 387–91. https://doi.org/10.1053/j.semperi.2017.07.009.

Riddell, Corinne A., Sam Harper, and Jay S. Kaufman. 2017. "Trends in Differences in US

Mortality Rates Between Black and White Infants." *JAMA Pediatrics* 171 (9): 911–13. https://doi.org/10.1001/jamapediatrics.2017.1365.

Tucker, Janet, and William McGuire. 2004. "Epidemiology of Preterm Birth." *BMJ* 329 (7467): 675–78. https://doi.org/10.1136/bmj.329.7467.675.

Vehtari, Aki, Jonah Gabry, Mans Magnusson, Yuling Yao, Paul-Christian Bürkner, Topi Paananen, and Andrew Gelman. 2020. "Loo: Efficient Leave-One-Out Cross-Validation and Waic for Bayesian Models." https://mc-stan.org/loo.

Vehtari, Aki, Andrew Gelman, and Jonah Gabry. 2017. "Practical Bayesian Model Evaluation Using Leave-One-Out Cross-Validation and Waic." *Statistics and Computing* 27 (5): 1413–32.

Wachter, Kenneth W. 2014. *Essential Demographic Methods*. Cambridge, Massachusetts: Harvard University Press.

WHO. 2018. "WHO Fact Sheet: Preterm Birth." World Health Organization. https://www.who.int/news-room/fact-sheets/detail/preterm-birth.