DEMOGRAPHIC RESEARCH
A peer-reviewed, open-access journal of population sciences

DEMOGRAPHIC RESEARCH

VOLUME 52, ARTICLE 10, PAGES 303–350
PUBLISHED 18 FEBRUARY 2025
https://www.demographic-research.org/Volumes/Vol52/10
DOI: 10.4054/DemRes.2025.52.10

Research Article

Infant mortality among US whites in the 19th century: New evidence from childhood sex ratios

Jesse McDevitt-Irwin

James R. Irwin

Contents

Demographic Research: Volume 52, Article 10
Research Article

Infant mortality among US whites in the 19th century: New evidence from childhood sex ratios

Jesse McDevitt-Irwin[1]

James R. Irwin[2]

Abstract

BACKGROUND

Basic facts of infant mortality in the 19th-century United States are largely unknown due to a lack of data on births and infant deaths. Contradictory views have emerged from previous research. Estimates from life table exercises with US census data, published in the most recent (2006) *Historical Statistics of the United States*, suggest that infant mortality among US whites circa 1850–1880 was substantially worse than in much of contemporary Europe. However, a broader range of historical evidence indicates that US whites were among the healthiest 19th-century populations.

METHODS

We offer a new basis for estimating infant mortality: childhood sex ratios. Because of the female survival advantage in infancy, high rates of infant death tend to be reflected in female-skewed childhood sex ratios. We verify the empirical relationship between infant mortality and childhood sex ratios in historical populations with credible data on both and demonstrate that sex ratios can reveal broad patterns of infant mortality.

RESULTS

Turning to the US census for under-5 sex ratios, we find that white infant mortality circa 1850–1880 was in the range of 60–110 deaths per 1,000 – well under half the values presented in *Historical Statistics of the United States* and below contemporary European levels. By 1900, infant mortality in the United States had increased substantially, pointing to the challenges that modernization posed to population health.

CONTRIBUTION

We demonstrate a novel method of characterizing infant mortality, using childhood sex ratios. With census data often available where vital statistics are not, our method promises to shed new light on historical patterns of population health. Applied to the 19th-century

[1] Center for Economic History, Northwestern University, Evanston, IL, USA.
Email: jesse.mcdevittirwin@northwestern.edu.
[2] Central Michigan University, Mount Pleasant, MI, USA.

United States, our method shows that infant mortality among the white population was much lower than previously suggested.

1. Introduction

Infant mortality is a key indicator of population health and living standards, especially historically, when differences in infant mortality across populations were far greater than they are today.[3] Unfortunately, the basic facts of infant mortality in the 19th-century United States have yet to be established because of a lack of records on births and infant deaths.[4] One view, emerging from research constructing life tables for the 19th-century United States (Haines 1979, 1998; Hacker 2010), suggests that US whites' infant mortality circa 1850 to 1880 was high by the standards of contemporary Europe. However, a wide range of existing evidence on historical infant mortality in Europe and the United States points to the opposite conclusion: that 19th-century US whites had relatively low infant mortality.

In this paper we offer a new empirical basis for characterizing broad patterns of infant mortality, using childhood sex ratios. It is well-known that males are biologically more vulnerable than females to infant mortality; the corollary that we build from is that high rates of infant death will skew the sex ratio among survivors toward females.[5] Assembling historical data from European and settler populations, we document a striking empirical relationship between infant mortality and childhood sex ratios. That relationship informs our simple model to estimate infant mortality from childhood sex ratios.

[3] In the 19th century, across Europe alone, infant mortality rates ranged from less than 100 to more than 300 per 1,000 (Mitchell 1998: 120–122). The world has seen a collapse of infant mortality since the early 20th century. By 2020, more than one-third of the world population lived in places with infant mortality rates below 10 deaths per 1,000 births, and two-thirds lived in places with rates below 30 (based on authors' tabulations from country data for 2020 reported at World Bank 2022a, 2022b).

[4] The empirical record of infant mortality in much of 19th-century Europe is reasonably complete (see data sources in the appendix), based on records of births and infant deaths that are simply unavailable for most of the United States before the early 20th century (Haines 2006; US Bureau of the Census 1975: 44). Without such records, estimates of infant mortality for the 19th-century United States as a whole are conjectures from model life tables. Most prominent are those of Haines (1979, 1998) and Hacker (2010); note that those studies seek to characterize 19th-century US mortality across the life span and are not focused on estimating infant mortality. Hacker (2010: 76) explicitly points out the need for further research on infant and childhood mortality during the period.

[5] Of course, this effect could be offset if extremes of sex discrimination reversed females' biological survival advantage, as seen in cases of "missing women," à la Sen (1989). Further to this point, see below. The expected relationship between infant mortality and childhood sex ratios has been previously used to identify cases of "missing women," for example in Beltrán Tapia and Raftakis (2022).

Applying our model to sex ratio data from the decennial US censuses, we estimate that infant mortality among the white population was in the range of 60 to 110 deaths per 1,000 births circa 1850–1880. Our findings sharply contradict the life table values of Haines (1998), which appear in the most recent edition of *Historical Statistics of the United States* (HSUS). Those infant mortality rates range from 176 to 217 for the period 1850 to 1880 – substantially worse than contemporary England or France. But our evidence shows that US whites were among the healthiest populations of the 19th century.

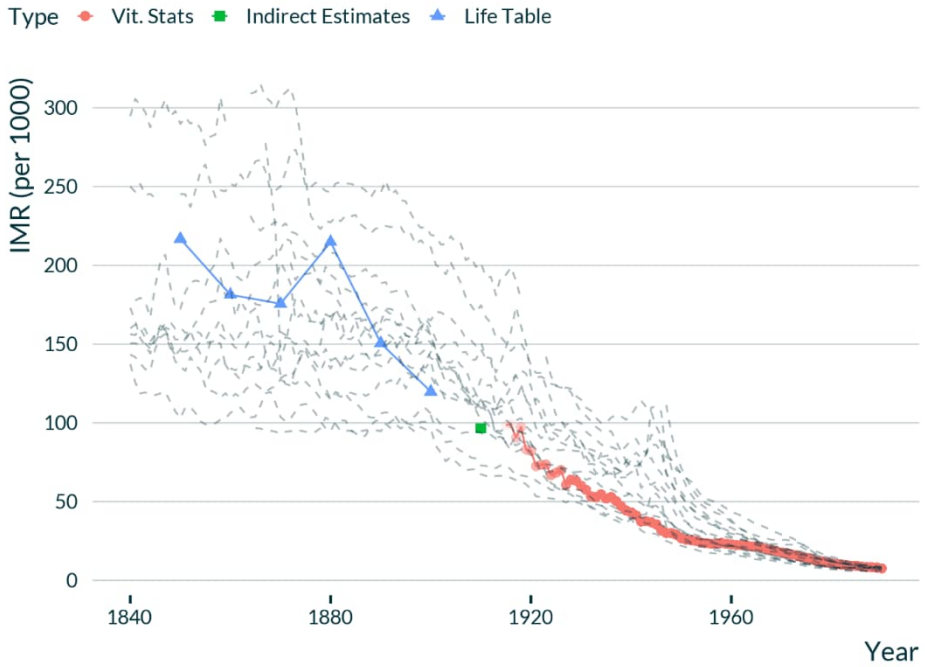
2. Historical background

The basic facts of infant mortality for US whites since the mid-19th century may appear to be reasonably complete.[6] The most recent (2006) edition of HSUS presents the infant mortality rate (IMR) for the white population at decennial benchmarks from 1850 to 1910 and annually starting in 1915.[7] Figure 1 plots this series against the backdrop of IMR data available for a cross section of European populations from 1840 to 1990. The HSUS series features high levels of infant mortality across the four census benchmarks from 1850 to 1880: between 175 and 217 deaths per 1,000 births, with an average just below 200. After 1880, the series traces out a long gradual decline in infant mortality, dropping below 100 by 1910, well below 30 points by 1950, and below 10 points by the early 1980s. Looking across Figure 1, the HSUS series falls well within the range of European infant mortality experiences. Arguably, what stands out is a general pattern of massive improvement in infant mortality since the late 19th century.

[6] The empirical record of non-white and black infant mortality is clearly incomplete (HSUS Series Ab922 and Ab923 [Carter et al. 2006], with estimates for 1850, 1900, 1910, 1916, and 1918–1998) and largely outside the scope of this paper.

[7] HSUS Series Ab921 (Carter et al. 2006). The annual IMR series, from 1915 on, was presented in previous editions of HSUS (US Bureau of the Census [1949: Series C40; 1960: Series B108; 1975: Series B140]). The most recent edition (Carter et al. 2006) added the census benchmark values of IMR for 1850 to 1910, from Haines (1998). See notes 17 and 18 for further discussion of the HSUS series but note that the value presented for the year 1910 is an estimate for circa 1904 based on 1910 census data (Haines 1998: 154, 167; Carter et al. 2006: Table Ab1-10, Footnote 2). Note also that prior editions of the HSUS were produced and published by the US Bureau of the Census (Carter et al. 2006: 1-xi); the current (2006) edition was "prepared by the academic community" (Chiswick 2006: 5_819), with Michael R. Haines the editor of the chapter on vital statistics (Haines 2006).

Figure 1: Infant mortality rates, 1840–1990



Note: The colored points are the HSUS series for US whites; the dashed lines are three-year rolling averages of various European populations.
Sources: HSUS Series Ab921 (Carter et al. 2006) and see appendix. US white infant mortality by year from HSUS, colored by the type of estimate. "Life Table" refers to values that are extrapolations from older-age mortality using life tables (Haines 1979, 1998). The indirect estimate is from Haines (1998: Table AI), which used the surviving-children method with data from the 1910 census (maternal recall and population by age), building from Haines and Preston (1997). The value presented for 1910 is actually an estimate for circa 1904, but we have plotted the values as reported in the HSUS series; further to this point, see Figure 5 and related text (below). The vital statistics series relies on data from birth and infant death records in the US Birth Registration Area (BRA); these values are shaded according to the proportion of the US white population covered by the BRA, going from one-third in 1915 to 100% in 1933 (Linder and Grove 1947: 840–857, 1036).

The substantial decline in the HSUS infant mortality series after 1880 might be viewed as just another facet of the widely studied "mortality transition" (Caldwell 2001), which has seen life expectancies soar and mortality rates plummet across the globe since the early 20th century.[8] The mortality transition forms a dominant paradigm for historical demographic research, framing a wealth of research investigating the emergence of the

[8] Global average life expectancy at birth rose from 32 in 1900 (Riley 2005: Table 1) to 73 in 2019 (UN Population Division 2022: 16). Over a similar period, under-5 mortality plummeted from roughly 1 in 2 (Hill 1995) to less than 1 in 25 (UNICEF 2023).

very low mortality regimes enjoyed in the "developed" world.[9] Central to this paradigm is a presumed historical fact – that rates of infant and child mortality were inevitably high in the preindustrial and early industrial past. Central also to the paradigm is a broad historical explanation for the collapse of mortality: that the scientific and industrial revolutions were "transmuted" into mortality decline (Caldwell 2006: 157–180).[10] The collapse in mortality thus plays a central role in broader narratives of progress surrounding 19th- and 20th-century modernization and industrialization. In this context, the trajectory of the HSUS series is unremarkable.[11]

However, the high *levels* of infant mortality found in the series for the 19th century are implausible in light of a diverse range of evidence. The HSUS series has US white infant mortality rates between 175 and 218 across the period 1850–1880, levels that would rank US whites among the worst of contemporary European populations (see Table 1 and Figure 1). That ranking is inconsistent with the well-known fact that US whites were among the tallest of contemporary populations (Fogel et al. 1983: 463), as there was a strong negative association between adult height and infant mortality in 19th-century populations (see Table 1). With average heights exceeding 173 cm, the United States appears similar to other settler societies, such as Australia and New Zealand, where infant mortality rates were substantially lower, and people taller, than in contemporary Europe. Granted, the relationship of adult height and infant mortality is not always straightforward, with concerns over the role of selection (Alter 2004; Deaton 2007). However, according to Schneider (2023: Figure 8), 19th-century US whites also had the lowest rates of childhood stunting in the world, strongly pointing toward good health in early childhood. Stunting prevalence among US whites was around one-quarter in the mid-19th century, a level that European populations would not reach until the 1920s and that many countries still haven't reached today (Ssentongo et al. 2021: Figure 1).

Data on IMRs from US cities in the 19th century provide even stronger evidence that the 19th-century HSUS values are implausible. Preston and Haines (1991: 53–57) present a variety of infant mortality estimates for US cities circa 1850–1880, with rates around 165–175 for Philadelphia, 180 for Brooklyn and Chicago, and 170–200 for Boston. It is well established that there was a substantial "urban penalty" (Kearns 1988) in infant mortality in the 19th century, meaning that the IMR in the rest of the country should have

[9] For classic contributions on the mortality transition, see Preston and Van de Walle (1978) for Europe and Condran and Cheney (1982) for the United States.

[10] Caldwell (2006) titles chapter 8 "Transmuting the industrial revolution into mortality decline."

[11] This view is reinforced by Hacker's (2010) life tables for the 19th-century US white population. The IMRs from these life tables also average about 200 for the period 1850–1870, and Hacker's life table IMR estimates exhibit a strong downward trend from the 1860s, fitting the mortality transition paradigm even better than does the HSUS series.

been lower than in the largest cities.[12] Indeed, death data from around the turn of the century clearly show that the highly urbanized Northeast region had higher infant mortality than the rest of the country.[13] It is simply implausible that the United States as whole, which was four-fifths rural in 1860, had infant mortality as high as that seen in its largest cities, but this is the implication of the HSUS life table values.

Table 1: Height and IMR in 19th-century Europe and settler colonies

Country	Birth Years	Height (cm)	IMR (per 1,000)
Italy	1861–1870	163.1	227
Great Britain	1861–1865	166.3	147
France	1856–1860	166.4	203
Netherlands	1856–1860	166.5	218
Belgium	1861–1865	166.8	165
Germany	1856–1860	167.3	287
Sweden	1856–1860	168.4	144
Denmark	1856–1860	168.5	134
Norway	1856–1860	168.9	97
Australia (whites)	1876–1880	171.7	121
New Zealand (whites)	1880s–1890s	172.5	86
United States (whites)	1830–1840	173.2	?

Sources: Data on European heights are from Hatton and Bray (2010: Appendix B); data on US heights are from Fogel et al. (1983: 463). IMRs are from the Human Mortality Database (HMD), except for those for Germany and Italy, which are from Mitchell (1998: 121). For Great Britain, IMR reflects combined values from the HMD for England and Wales and for Scotland. New Zealand heights are from Inwood, Oxley, and Roberts (2010: Table 1), and IMR is from Stats NZ Store House. Australia heights are from Whitwell and Nicholas (2001: Figure 1) and IMR is from McDonald, Ruzicka, and Pyne (1987: 58).

This implausibility of the HSUS IMR values for the 19th century is seen most clearly when comparing them to the one state with credible vital statistics dating back to the mid-19th century: Massachusetts (Abbott 1897: 714; Haines 2006: 385). We have already established that the highly urbanized Northeast had a relatively high IMR, as expected by the wide literature on the urban penalty. Yet in Massachusetts we see infant mortality averaging less than 160 for the period 1860–1880 – a remarkable 40 points below the level of infant mortality in the HSUS series for US whites as a whole. But birth and infant

[12] For example, in 1890 England, urban infant mortality was about 220, while the rural number was just under 100 (Woods, Watterson, and Woodward 1988: 353). On the urban mortality penalty in the 19th century, among many possible, see also Davis (1973: 102–104), Williamson (1982), Haines (2001), and Cain and Hong (2009).
[13] See Condran and Crimmins (1979, 1980) for 1890 and 1900 infant death rates by state. In 1900 the Northeast was 66% urban and the Midwest 39% (US Census Bureau 2012: 20, 22). The 1900 Death Registration Area data show much higher rates of infant death in states of the Northeast (35–38) than in those of the Midwest (23–25) (authors' calculations). A similar pattern is seen in the 19th century; Lynch, Mineau, and Anderton (1985: Table 4) find that infant mortality in Utah from 1850 to 1880 was around 100, just half the level found in the HSUS national series. Haines's (1977) results for upstate New York in 1865 point in a similar direction. Using census data on maternal recall, Haines (1977: Table 4) estimates rural under-5 mortality at 18%–19% and urban at 25%–26%.

death records from 1890 and 1900, covering many more states, clearly identify Massachusetts as a high mortality state, as expected from its relatively high level of urbanization.[14] In sum, it is simply not credible that infant mortality among 19th-century US whites exceeded that of the state of Massachusetts. The high rates of IMR presented in HSUS are puzzling if not simply incredible.

However, this puzzle has a simple solution: the HSUS infant mortality rates for the 19th century are better described as conjectures than estimates. They are life table extrapolations from older-age mortality, with no basis in data on births or infant deaths. In sharp contrast, the annual HSUS series (from 1915 onward) are official statistics – direct estimates of infant mortality from records of births and infant deaths.[15]

In terms of sources and methods, the HSUS series (Ab921) includes three different types of estimates. First, the annual values (1915 on) are direct estimates of infant mortality from registration of births and infant deaths.[16] Second, the value for 1910 is a standard indirect estimate of infant mortality, using census data on maternal recall of children born and surviving.[17] Third, the decennial benchmark values from 1850 to 1900 are from model life tables, with no basis in data on births and infant deaths.[18] Lacking requisite data for direct or indirect estimates of infant mortality, Haines (1979, 1998) fit model life tables to census mortality data for ages 5 to 20.[19] The estimated life tables include the level of infant mortality for each census year (1850 to 1900), which appear in HSUS Series Ab921 (Carter et al. 2006).[20]

[14] For example, in 1900, for which more vital statistics data are available, the infant death rate in Massachusetts (86% urban) was 182, while in Michigan (40% urban) it was 128 (authors' calculations based on data in Condran and Crimmins 1980: Table 1). More generally, Massachusetts's infant mortality appears to have been typical of the highly urbanized Northeast region (Condran and Crimmins 1980).

[15] Recall from note 7**,** above, that previous editions of HSUS included the annual IMR series from 1915 onward; the decennial IMR values for 1850 to 1910 in the most recent HSUS are from Haines (1998).

[16] Nationwide birth and infant death records start in 1933. From 1915 to 1932, the estimates cover just part of the country, the BRA, covering about one-third of the white population in 1915, increasing to 95 percent coverage in 1932 (Linder and Grove 1947: 840–857, 1036).

[17] Though labeled 1910, the estimate is for circa 1904, based on maternal recall in the 1910 census (Preston and Haines 1991: 74; Haines 1998: 154; Carter et al. 2006: Table Ab1-10, Footnote 2).

[18] HSUS Series Ab9 (Carter et al. 2006) also presents the decennial benchmark values of white infant mortality, with the following footnote: "For the expectation of life at birth and the infant mortality rate, the values for 1900 and 1910 are from approximately 1895 and 1904, respectively" (Carter et al. 2006: Table Ab1-10, Footnote 2). The footnote is correct for 1910 (see note above) but not for 1900. Although Haines (1998: 154, 165) includes an indirect estimate of white infant mortality for circa 1894–1895, based on maternal recall in the 1900 census, the HSUS IMR series has the model life table value for 1900 from Haines (1998: 160; 1979: 307). On the indirect estimates, see also Haines and Preston (1997).

[19] Haines restricts his data to ages 5–20, reasoning that these census mortality data are more accurate than the data for other ages (Haines 1977: 327, 330; 1979: 290).

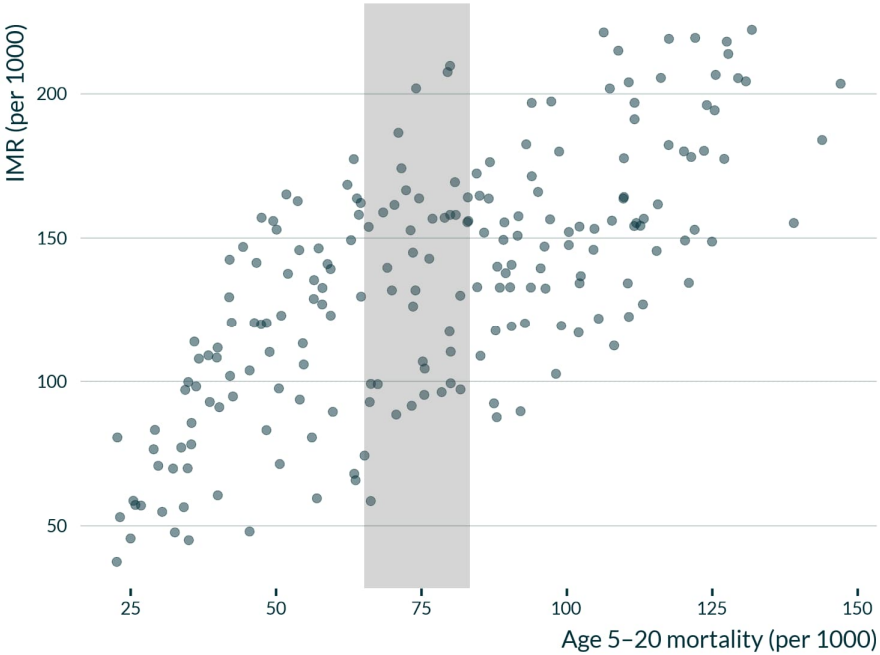
[20] Hacker (2010) similarly estimates life tables for the 19th-century US white population, but based on existing estimates of life expectancy at age 20 rather than census mortality data. The associated infant mortality values are broadly comparable to those from Haines (1979).

However, infant mortality has no necessary relationship with mortality at older ages.[21] As emphasized by Woods (1993: 217), indices of mortality in infancy, early childhood, and adulthood are all "indispensable" for characterizing a population's mortality, because "each one captures a distinctive aspect of the mortality pattern and their empirical interrelations clearly were not predictable in the past." For example, in England from 1840 to 1880, age 5–20 mortality declined by half while infant mortality was roughly constant. Here, extrapolating from age 5–20 mortality would produce severe overestimates of past infant mortality. More generally, the highly credible life tables of the HMD show a wide range of infant mortality rates for given levels of mortality at older ages.[22] Figure 2 plots infant mortality rates against age 5–20 mortality rates from HMD life tables, covering a range of European (or European-descent) populations in the period 1835–1925; the shaded area shows the range of age 5–20 mortality rates in Haines's (1998) life tables that produced the HSUS IMR estimates. With this range of age 5–20 mortality, infant mortality rates in the historical life tables ranged from below 70 to above 200.

Over a decade ago, Hacker (2010: 76) noted the problems associated with estimating infant mortality from mortality at older ages, concluding that "empirical research on infant and child mortality in the United States is sorely needed." So far, a lack of birth and infant death records has stood in the way of such research, and the basic facts of infant mortality in the 19th-century United States remain to be determined. Here we offer a new empirical approach. Building off well-known facts of biology and demography, we have devised a new method for characterizing patterns of infant mortality using readily available census data on childhood sex ratios. Relative to values presented in the current HSUS, our evidence points to dramatically lower infant mortality rates among 19th-century US whites, much as one would expect given known patterns of population health in both Europe and the United States at the time.

[21] Hacker (2010: Table 6) makes this point within the context of the 19th-century United States, illustrating the wide range of infant mortality possible when extrapolating from adult mortality.
[22] HMD life tables are for "populations where death registration and census data are virtually complete" (HMD "Scopes and basic principles," accessed October 27, 2023).

Figure 2: Infant mortality by age 5–20 mortality



Sources: HMD life tables, 1835–1925. The shaded area is the range of estimates of age 5–20 mortality for US whites from 1850 to 1880, as given by Haines (1998: 156–164).

3. Infant mortality and childhood sex ratios

It has long been known that biologically, girls are less vulnerable than boys to infant mortality.[23] The corollary we highlight, and build from, is that high rates of infant mortality skew the sex ratio among surviving children toward females. Our work marks

[23] Current knowledge is conveniently encapsulated by the editors of *PLOS Medicine* in their summary of Sawyer (2012): "Newborn girls survive better than newborn boys because they are less vulnerable to birth complications and infections and have fewer inherited abnormalities. Thus, the ratio of infant mortality among boys to infant mortality among girls is greater than one, provided both sexes have equal access to food and medical care." Knowledge of excess male infant mortality dates back at least to the 18th century; see, for example, Struyck (1740), Wargentin (1755), and Clarke (1786). For discussion see Théré and Rohrbasser (2006). The female survival advantage in infancy is attributed to multiple factors: females have fewer congenital diseases owing to their redundant X chromosome, and they are also more resistant to infectious disease. For a review see Waldron (1998: 64–83).

a sharp pivot away from the wide literature on "missing women" (Sen 1989; Coale 1991; Klasen 1994; Das Gupta 2005; Beltrán Tapia and Raftakis 2022), where male-biased sex ratios feature prominently. Our focus instead is on populations in which the "biological vulnerability" of infant boys is not outweighed by the "social vulnerability" of girls (Thompson 2021: 467).[24] In such populations, the degree of female skew in the child population is indicative of the level of infant mortality. Whereas existing work has characterized the expected relationship between infant mortality and childhood sex ratios to identify male-biased populations (Beltrán Tapia and Raftakis 2022), we develop a novel use: inferring infant mortality rates from observed childhood sex ratios.[25]

The effect of infant mortality on childhood sex ratios is apparent in both historical populations and familiar model life tables. For example, in 1900 infant mortality in Austria was above 200 and there were similar numbers of boys and girls under the age of 5. By 1970 infant mortality had plummeted to 20 deaths per 1,000 and there were about 5% more boys than girls, a value typical of the sex ratio at birth in healthy populations (Maconochie and Roman 1997; Grech, Savona-Ventura, and Vassallo-Agius 2002). A similar pattern is found in a wide range of polities (see Figure 3). In familiar model life tables, the relationship is also evident. For example, in the UN general model, moving from a life expectancy at birth (e_0) of 65 years for both males and females to an e_0 of 35, infant mortality increases from 54 deaths per 1,000 to 183, skewing the sex ratio among survivors to age 1 (l_1) roughly 5 percentage points towards females (UN 1982: 258–260).[26]

We use standard life table modeling to illustrate the impact of infant mortality on the sex ratio among surviving children. Consider the childhood sex ratio as the natural logarithm of a hypothetical population of survivors to age 1, $SR1 \stackrel{\text{def}}{=} ln(\frac{l_1^f}{l_1^m})$.[27] The natural logarithm provides both analytical and presentation advantages over more familiar representations of sex ratios, and we use it throughout the paper.[28] Although the

[24] The 19th-century United States is such a case, as existing evidence on child mortality shows a clear female survival advantage (Haines 1977: Table 7; Kunze 1979: Table 14; Lynch, Mineau, and Anderton 1985: Table 4). We discuss this point at length in the "Robustness" section.

[25] This contrast is shown clearly in Beltrán Tapia and Raftakis (2022: Figure 2), where the authors make a similar scatterplot as our Figure 3, with infant mortality on the x-axis and sex ratios on the y-axis. While they are interested in outliers, we are focused on the typical pattern. See Beltrán Tapia and Gallego-Martínez (2017), Beltrán Tapia and Szoltysek (2022), and Beltrán Tapia and Cappelli (2024) for further examples of the use of the expected relationship of IMR and childhood sex ratios to identify populations with "missing girls."

[26] This effect would be even stronger if we accounted for the well-known pattern that women tend to live longer than men, particularly in our period of interest (Tabutin and Willems 1998).

[27] We adapt notation from Preston, Heuveline, and Guillot (2001: Chapter 3).

[28] The most common existing representations of sex ratios are: (1) the number of males per 1,000 females, as in academic demography dating back to at least Jastrzebski (1919); (2) the number of females per 1,000 males, as in most South Asian academic and policy publications (e.g., Oldenberg 1992); and (3) the proportion of males, as in much of the human biology literature (e.g., Orzack et al. 2015). In addition to additive separability

childhood populations reported in census data correspond to person-years in an age interval (most often ${}_5L_0$ in our case), we model the sex ratio among survivors to age 1 (l_1). This simplification clarifies the key factors determining childhood sex ratios without sacrificing the validity of our model, as sex ratios based on ${}_5L_0$ and l_1 are roughly equivalent in both historical and model life tables.[29] With B^j the number of births and q_0^j the infant mortality rate of sex j, we can express the sex ratio of survivors to age 1 as follows:

$$SR1 \stackrel{\text{def}}{=} ln(\frac{l_1^f}{l_1^m}) = ln(\frac{B^f \cdot (1 - q_0^f)}{B^m \cdot (1 - q_0^m)}).$$

With a few steps of algebra, and defining $SRB \stackrel{\text{def}}{=} ln(\frac{B^f}{B^m})$, we get the following expression:

$$SR1 = SRB + [ln(1 - q_0^f) - ln(1 - q_0^m)]. \quad (1)$$

Here we see that the sex ratio at age 1 is determined by two additively separable terms: the sex ratio at birth and the relative survival of girls and boys. The additive separability comes directly from using the logarithm of the sex ratio and is a clear advantage over alternative representations of sex ratios, in which the sex ratio at birth and the effect of mortality are multiplicative in determining the sex ratio among survivors. This additive separability implies that as infant mortality approaches 0, so does the second term, and the childhood sex ratio approaches the sex ratio at birth. It follows that the empirical implications of our model are limited for populations with low infant mortality, in which childhood sex ratios will reflect sex ratios at birth.

For an empirically tractable expression, we take Taylor series approximations ($ln(1 + x) \approx x$). Defining q_0 as overall infant mortality and $\mu = \frac{q_0^m - q_0^f}{q_0}$ as excess male mortality, we obtain:

(discussed below), our logarithmic representation also has the advantage of being symmetric with respect to whether males or females are the reference group. It is also naturally expressed in easily interpreted values: percentages. Therefore, throughout this text, we report sex ratios as natural logarithms expressed as percentages. For a population with 1,050 boys and 1,000 girls, with ln(1050/1000) = 0.048790 or 4.88%, we would describe the sex ratio as about 4.9% more male than female or, equivalently, as about 4.9% less female than male.

[29] Using HMD data from 1970 and earlier, we find an R^2 of 0.98 between the l_1 and ${}_5L_0$ sex ratios. Similarly, taking all eight families of model life tables from the UN Population Division (2011), we find an R^2 of 0.98 between the l_1 and ${}_5L_0$ sex ratios among levels with at least 30 deaths per 1,000. This equivalence is largely because excess male mortality is so much stronger in infancy than later in childhood.

$$SR1 \approx SRB + \mu \cdot q_0. \quad (2)$$

Equation (2) clarifies that infant mortality and excess male mortality combine to move the childhood sex ratio toward girls, away from the sex ratio at birth. The greater the excess male mortality, μ, the more infant mortality skews the sex ratio among survivors. Importantly, this effect is roughly proportional to the level of infant mortality, and as suggested above, the effect will be negligible for populations with low infant mortality (e.g., rates below 20). However, the effect will be substantial in populations with high infant mortality. For example, with excess male infant mortality of 20%, well within the relevant historical range (Hill and Upchurch 1995), if infant mortality decreased from 150 to 100, the sex ratio would shift about 1 percentage point toward boys.

Equation (2) provides a simple model for understanding the drivers of childhood sex ratios and guides our basic empirical approach below. However, several considerations complicate this simple model. Most simply, the magnitude of excess male mortality is not constant across populations or times (Drevenstedt et al. 2008), which could attenuate or exaggerate the observed effect of IMR on childhood sex ratios (CSR). Less simply, insults to maternal health tend to push the sex ratio at birth toward females (Fukuda et al. 1998; Catalano 2003; Almond and Edlund 2007).[30] As maternal and infant health are closely linked (Kramer 1987), this process of fetal loss could reinforce the observed relationship between infant mortality and childhood sex ratios.[31] Working in the opposite direction is the likelihood that fetal loss would be selective, meaning that women in poor health would give birth to more robust infants who were less vulnerable to infant mortality (Catalano and Bruckner 2006; van Dijk, Nilsson, and Quaranta 2024). In sum, any "structural" (Goldberger 1972) interpretation of Equation (2) thus runs into concerns of endogeneity. Fortunately, our goal is prediction, not estimating parameters. The extent to which childhood sex ratios reflect, and therefore can predict, infant mortality is an empirical question, which we address with historical data from populations where both variables are available.

[30] The apparent mechanism is maternal stress hormones, which increase the probability of miscarriages, which are disproportionately male (James and Grech 2017: 51). The sex ratio at birth has been used as an indicator for maternal health and fetal loss (Davis, Gottlieb, and Stampnitzky 1998; Grech and Masukume 2016; Shifotoka and Fogarty 2013; Sanders and Stoecker 2015; Valente 2015; Guimbeau, Menon, and Musacchio 2022).

[31] Klasen (1994: 1064–1066) noted this relationship between sex ratio at birth and infant mortality in the context of "missing women."

4. Data

To characterize the empirical relationship between childhood sex ratios and infant mortality, we assemble data from Europe, the United States, and other settler societies, mostly from the mid-19th century onward.[32] Data for childhood sex ratios are taken from censuses or population registries and those for infant mortality are taken from official sources, International Historical Statistics, and the HMD.

We pair a childhood sex ratio with an average rate of infant mortality in preceding years.[33] We generally use the under-5 population, but other age groupings yield the same basic results. Using the under-5 age group has a number of important advantages over using younger ages while still letting us avoid problems introduced at older ages by migration. First, the data are more widely available from published sources. Second, the five-year age span increases the sizes of childhood populations, reducing the role of random variation in sex ratios.[34] Finally, pooling across ages reduces the impact of sex-biased age heaping. The starting points for our series are dictated by the availability of data. We end our series at the start of the 1960s; by then, rates of infant mortality in our sample populations were too low to materially affect childhood sex ratios and ultrasound, which spread in the 1970s (Campbell 2013), was not yet a factor in sex ratio patterns. We restrict our dataset to under-5 populations of at least 25,000. We have 571 observations for Europe and settler societies other than the United States. For the United States, we have eight observations from the state of Massachusetts for the 19th century and 177 observations from a variety of other aggregates (urban, rural, and mixed) for the years 1900 to 1940.[35] Thus a typical observation in our dataset pairs the (ln) under-5 sex ratio from a particular year with the average infant mortality rate for the preceding five years for some country or subnational unit (plotted in Figure 3). Summary statistics for these data can be found in the appendix.

[32] See the appendix for a fuller discussion of our sample. In brief, our non-US data cover Sweden (1757–1960), Denmark (1840–1960), Belgium (1846–1960), England and Wales (1851–1961), the Netherlands (1859–1960), Scotland (1861–1960), New Zealand (1867–1961), Austria (1865–1961), Australia (1880–1961), German polities (1849–1961), Switzerland (1880–1960), Finland (1885–1960), Norway (1890–1960), France (1901–1954), Italy (1911–1961), and South Africa (1918–1921). For the United States we have Massachusetts from 1860 to 1960 and an increasing number of states from 1900 to 1940.

[33] With some exceptions, we pair the under-5 sex ratio with the prior five-year mean of the infant mortality rate. We have under-6 populations for 140 Prussian cases of 1890–1910. For Prussian districts in 1849 and some US states in 1900, we have just one year of infant mortality data (see the appendix for details).

[34] Random variation in sex ratios will not be small unless populations are large. To illustrate, model the sex proportion as a binomial random variable, as in Visaria (1967: 33), with mean 1/2. With 10,000 children, the 90% CI is 6 percentage points, which is very large relative to the effects we seek to measure. With 50,000 children, the 90% CI shrinks to about 3 percentage points.

[35] By 1950, US infant mortality had fallen below 30 deaths per 1,000 and state-level differences in white infant mortality were too small to be useful for our study.

5. Results

5.1 Childhood sex ratios reveal infant mortality

Above we have established the theoretical basis for childhood sex ratios reflecting infant mortality. Figure 3 provides a first test of the empirical relevance of our model, plotting under-5 sex ratios (SR5) against IMR. The empirical correspondence is striking, demonstrating that childhood sex ratios are closely related to infant mortality both in theory and practice. High rates of infant mortality imply relatively more girls, and low rates imply relatively more boys. Given this strong empirical relationship, childhood sex ratios can shed new light on infant mortality in populations lacking data on births or infant deaths.[36] We proceed in three steps to draw inferences about IMR from childhood sex ratios. First we use least squares regression to predict IMR from under-5 sex ratios; then we use quantile regression and Bayesian modeling to quantify the uncertainty in our predictions.

To characterize infant mortality based on childhood sex ratios, we start with a simple regression of the form of our model's Equation (3):

$$SR1 \approx SRB + \mu \cdot q_0 \tag{3}$$

With the under-5 sex ratio proxying for SR1 and using weighted least squares with the sample data (from Figure 3), we obtain:[37]

$$\widehat{SR5} = -0.0492 + 0.225 \cdot IMR.$$

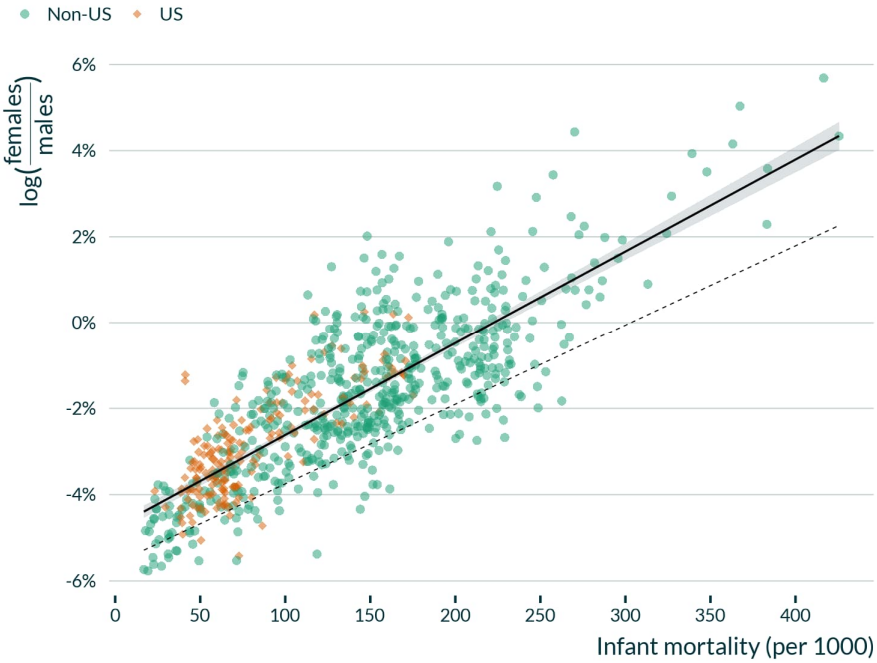
Our estimated equation fits very well with the theoretical predictions from our model above. The regression intercept (–0.0492) corresponds to a sex ratio at birth with about 5% more boys than girls, as expected for populations with very low infant mortality (Maconochie and Roman 1997; Grech, Savona-Ventura, and Vassallo-Agius 2002). The slope coefficient (0.225) falls well within the 15%–25% range of excess male mortality

[36] Moreover, the European and US data follow very similar patterns, giving us confidence in extrapolating to the 19th-century United States. See Figure 3, where the available US states and regions (mostly from 1900 onward but also from Massachusetts for 1860 onward) are shaded orange. They fit well within the broader CSR–IMR pattern seen in contemporary Europe.

[37] We use a priori efficient regression weights, equal to 1 over the square root of the sampling variance of each observation. This sampling variance is a classical measurement error coming from an underlying binomial distribution that generates the sex ratio for a given population. The magnitude of the sampling variance is inversely related to the population size; we calculate it for each observation via simulation. In the Bayesian section, below, we model this underlying measurement error, incorporating it into prediction intervals. Our results are robust to other weighting (see Figure A-1).

typical of the populations in our sample.[38] Put simply, in our sample, a 45-point increase in infant mortality is associated with a 1-percentage-point shift toward girls ($0.225 \cdot 0.045 \approx 0.01$).

Figure 3: Under-5 sex ratios by infant mortality



Notes: The black line is the regression of under-5 sex ratios on infant mortality; the dashed line is the 10th percentile regression. See text below. Data mainly from Europe and the United States (see "Data" section).

While it might be tempting to develop a richer empirical model exploiting other data available for the polities in Figure 3, our goal here is prediction. To what extent can we predict infant mortality from sex ratios alone? In our simple specification, infant mortality accounts for more than two-thirds of the variation of childhood sex ratios within our

[38] Infant mortality by sex is available from the HMD for most of our sample populations, with the major exceptions being the German and Austrian Empires in the 19th century. In the available populations, excess male mortality (μ) was generally 15%–20% in the 19th century before increasing to around 25% as mortality declined in the 20th century. These values line up well with those found in existing work (Hill and Upchurch 1995; Drevenstedt et al. 2008).

sample, with $R^2 = 0.670$. Inverting the regression results above gives us a simple estimator of infant mortality from childhood sex ratios:[39]

$$\widehat{IMR} = \frac{SR+0.0492}{0.225} \quad (4)$$

As proof of concept, we apply our prediction method to Massachusetts – the only US state with reasonably complete records on births and infant deaths going back to the mid-19th century. We drop the Massachusetts data, re-estimate our regression, and then predict IMR from under-5 sex ratios in Massachusetts. As plotted in Figure 4, we find a striking, if rough, correspondence between predicted infant mortality and the actual values (five-year averages). The Massachusetts example illustrates the promise of childhood sex ratios for characterizing the approximate level of infant mortality in a population.

To some extent, plots like Figures 3 and 4 convey the approximate nature of estimates of infant mortality based on childhood sex ratios. However, formal approaches offer some precision about the degree of uncertainty in our predictions of infant mortality from childhood sex ratios. We take two approaches to quantifying the uncertainty in our estimates. First, adopting a frequentist perspective, we use quantile regression to establish an upper bound on plausible IMR given observed sex ratios. Second, we construct posterior predictive intervals, using Bayesian techniques to model the distribution of IMR as a function of observed sex ratios. Although coming from two distinct conceptual frameworks, these approaches result in the same basic conclusions about the degree of uncertainty in inferences about infant mortality from data on childhood sex ratios.

For a frequentist perspective, we use the data in Figure 3 to construct a range of plausible infant mortality rates given observed childhood sex ratios. We characterize the conditional distribution of sex ratios on infant mortality using quantile regression, allowing us to infer the likelihood of an observed sex ratio given hypothesized levels of infant mortality. Estimating a conditional quantile, we can then construct hypothesis tests, ruling out unlikely levels of infant mortality. In Figure 3 we plot the 10th percentile of the under-5 sex ratio conditional on infant mortality:

$$\hat{q}_{SR|IMR}(10\%) = \hat{\gamma} + \hat{\delta} \cdot IMR = -0.0596 + 0.208 \cdot IMR.$$

For an observed under-5 sex ratio of SR_i, we reject all infant mortality beyond the level that corresponds to this 10th percentile: reject if $IMR > \underline{IMR}$, where

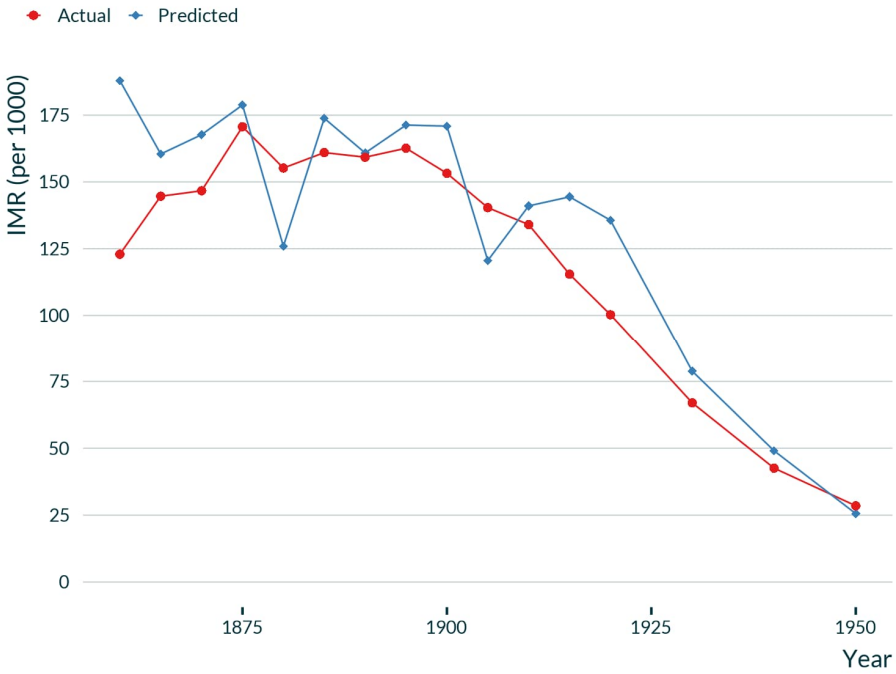
[39] Of course, Equation (4) could generate predictions of negative infant mortality rates, highlighting the point made below that these methods are not intended to be useful for populations with very low IMR, as seen in much of the world today.

$$\underline{IMR} = \frac{SR+0.0596}{0.208}. \quad (5)$$

Graphically, given an observed sex ratio, with 90% confidence we reject all infant mortality to the right of the dashed line plotted in Figure 3. Here we have an upper bound for estimates of infant mortality, in the spirit of classical hypothesis testing (with 10% significance, 1-tail test). For example, given an observed sex ratio of 3% more boys than girls, we would obtain an upper bound on infant mortality of roughly 140 deaths per 1,000:

$$(\underline{IMR} = \frac{-0.03+0.0596}{0.208} \approx .14).$$

Figure 4: Out-of-sample prediction of infant mortality from childhood sex ratios for Massachusetts



Notes: Predicted values calculated from a regression of under-5 sex ratios on infant mortality (five-year averages). Regression data from Figure 3, excluding Massachusetts data. Massachusetts sex ratios from state and federal censuses; IMR from HSUS Series Ab928 (Carter et al. 2006).

Bayesian techniques provide a different path to describing the uncertainty in our estimates of IMR: posterior predictive intervals.[40] We find the probable distribution of infant mortality, conditional on an observed childhood sex ratio, with a model estimated on the data from Figure 3. Details of the model and estimation are in the appendix.[41] This Bayesian approach incorporates three distinct sources of uncertainty in our predictive intervals: regression uncertainty (standard errors on coefficients), variance in infant mortality not explained by sex ratios (model regression residuals), and the statistical noise inherent to finite population sex ratios (measurement error). Given an under-5 sex ratio and the underlying population size, we generate a predictive interval from estimated iterations of the posterior distribution. For example, with 3% more boys than girls among 250,000 children, the 50% posterior predictive interval is roughly 70 to 120 deaths per 1,000. For a smaller population, say 25,000 children, the interval would be wider, from roughly 60 to 130. For much smaller populations, sex ratios are of little use due to their inherent noisiness.

Both of these approaches illustrate the approximate nature of predicting infant mortality from childhood sex ratios. The degree of uncertainty in these predictions makes them too coarse a tool for populations with very low rates of infant mortality, like most of the world today. Much of this uncertainty likely comes from the fact that excess infant male mortality varied over time and place (Drevenstedt et al. 2008). Future research might tighten these intervals for specific cases by allowing the slope to vary across countries, for example in a hierarchical model. But our simple, bivariate approach is more than sufficient for characterizing broad patterns of infant mortality in the 19th-century United States, where IMR might have been anywhere from below 100 to above 200 deaths per 1,000.

5.2 US infant mortality 1850–1880

Having established the usefulness of childhood sex ratios for inferring infant mortality, we now apply these methods to the 19th-century US white population. We draw on the four decennial censuses from 1850 to 1880, using data from both the published census volumes and the full count IPUMS samples (Ruggles et al. 2024).[42] We exclude the 1890

[40] We thank an anonymous referee for suggesting Bayesian posterior predictive intervals for our analysis.
[41] We only note here that we use "weakly informative priors" (à la Gelman et al. 2008) and that we estimate the model using brms in the R package, which calls the C++ program Stan.
[42] We use the average of the IPUMS and published census values, viewing both as plausible tallies of the underlying manuscripts. The two sources give very similar under-5 sex ratios; 1850 shows the biggest discrepancy, with the full count IPUMS ratio being 0.36% more male than the census volume's ratio.

census because age reporting in that year was inconsistent with practices in the rest of the censuses, biasing childhood sex ratios in 1890 toward males.[43]

As discussed above, the HSUS series places US white infant mortality in the period 1850 to 1880 at around 200 deaths per 1,000. Such high infant mortality has strong and simple implications for childhood sex ratios. Referring to our model above, Equation (1), supposing a modest degree of excess male mortality – 20% – and a typical sex ratio at birth – 5% more boys than girls – an infant mortality rate of 200 would result in a childhood sex ratio of parity. Referring to our simplest empirics (Figure 3), we see that for populations with infant mortality of 200, under-5 sex ratios are similarly concentrated in the range of 1 percentage point on either side of parity. The observed under-5 sex ratios of US whites flatly contradict these implications, with values ranging from 3.1% more boys than girls in 1870 to nearly 3.5% in 1850 (see Table 2). Within our sample, such sex ratios are generally associated with infant mortality around 80 deaths per 1,000. More precisely, the estimator (Equation [4]) from our reference dataset provides a new perspective on historical infant mortality of US whites. Plugging under-5 sex ratios into Equation (4), we obtain estimates of the five-year mean infant mortality for census benchmarks. For example, the 1860 census yields IMR for 1856–1860.

Table 2: US white childhood sex ratios and associated estimates

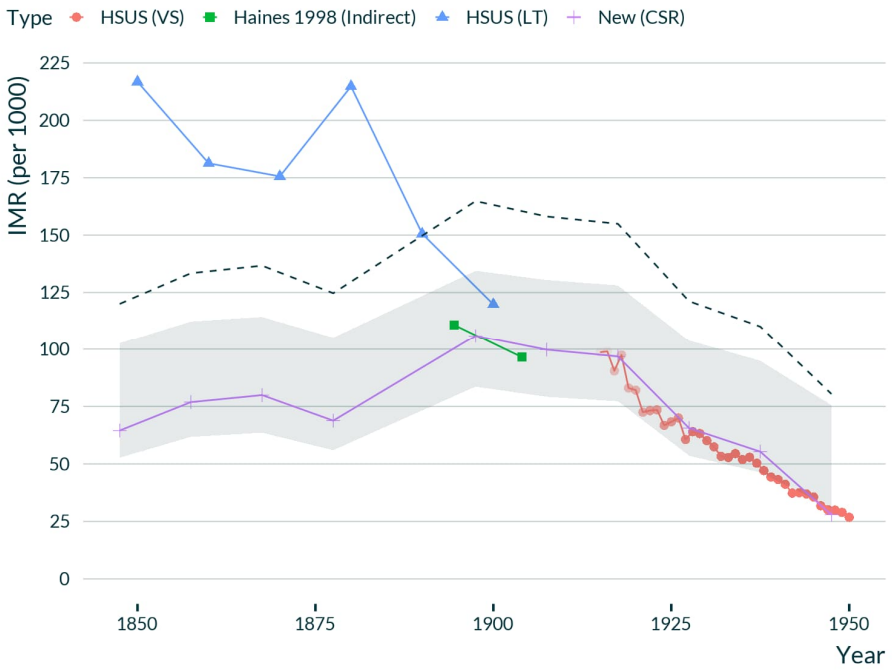
Census year	Under-5 sex ratio (100*log F/M)	IMR estimate (from CSR) deaths per 1,000	IMR interval estimate (IQR) deaths per 1,000	IMR upper bound (90%) deaths per 1,000
1850	–3.47%	65	53–103	120
1860	–3.18%	77	62–112	133
1870	–3.12%	80	64–114	137
1880	–3.37%	69	56–105	125
1890	–3.73%	N/A	N/A	N/A
1900	–2.54%	106	84–135	165
1910	–2.67%	100	79–130	158
1920	–2.74%	97	77–128	155
1930	–3.44%	66	54–104	121

Notes: Under-5 sex ratios from US census, both census volumes and PUMS; see "Data" section, above. IMR estimate and IMR upper bound calculated from under-5 sex ratios; see Equations (3) and (4). The year 1890 is excluded from estimation due to enumeration concerns, as discussed in note 43.

[43] The 1890 census recorded "age at nearest birthday" instead of "age at last birthday," which was used from 1850 to 1880 and from 1900 forward (US Census Office 1902: xlviii; and see the questionnaires for subsequent censuses at US Census Bureau 2021). A child approaching 5 years of age would be enumerated as age 5 in 1890 but in the other censuses would be enumerated as age 4. Thus older 4-year-olds would be under-represented in the 1890 census under-5 cohort (compared to the other censuses), biasing that cohort's sex ratio toward males (because the sex ratio among 4-year-olds is less male than among infants, a result of excess male infant mortality). This pattern is evident in the census counts of the US-born populations of 1890 and 1900: the under-5 cohort of 1890 numbered 6.49 million with a sex ratio 3.7% male; ten years later, the age 10 through 14 cohort numbered 6.65 million with a sex ratio 2.3% male. Further to this point, see the "Robustness" section.

We plot our new estimates of IMR for US whites alongside those from Haines (1998) and HSUS (Carter et al. 2006) in Figure 5. From 1900 onward, our values line up well with existing estimates based on maternal recall (1895 and 1904) from Haines (1998), as well as vital statistics (1915 onward) from HSUS.[44] But for 1880 and earlier – years for which the HSUS series is based solely on extrapolation from age 5–20 mortality – our new estimates are much lower than those of HSUS.

Figure 5: Estimates of US white infant mortality: 1850–1950



Notes: Plus signs are our new estimates, five-year average rates based on childhood sex ratios, using Equation (4). The dashed line gives a 90% upper bound, using Equation (3). The shaded area is the 50% (interquartile range) Bayesian posterior predictive interval. As in Figure 1, the existing estimates from HSUS Series Ab921 (Carter et al. 2006) are broken up by type. The left segment (1850–1900) comes from life tables, the right segment from births and infant death records (shaded by degree of coverage). In the middle we plot the two indirect estimates from Haines (1998), for circa 1894–1895 (based on the 1900 census) and for circa 1904 (based on the 1910 census). The latter value appears in HSUS Series Ab921 and Ab9 (Carter et al. 2006) for the year 1910, but here we plot it with the correct year.

[44] As discussed above, the HSUS series presents life table values from Haines (1998) for 1850–1900 and official vital statistics for 1915 onward. Figure 5 also presents the indirect IMR estimates for 1894–1895 and 1904 from Haines (1998: 165–167; see also Haines and Preston 1997: 80, 88). As noted above, the indirect estimate for circa 1904 is presented in the HSUS series for the year 1910; we correct that date in Figure 5.

For the period 1850–1880, our point estimates of US infant mortality fall between 66 and 82 deaths per 1,000. The 50% Bayesian posterior prediction interval (interquartile range) is roughly 60–110 across the same period. Building from our quantile regression above, which characterizes the 90th percentile of childhood sex ratios given infant mortality, we use Equation (3) to construct an upper bound on US white infant mortality. At the 90% confidence level, we can consistently reject infant mortality approaching 140 deaths per 1,000, with our upper bound ranging from 120 in 1850 to 137 in 1870.[45] Thus we reject the HSUS life table values, which range from 167 to 218 across the period. We would also reject the hypothesis that US whites had infant mortality approaching that of, for example, England (IMR around 150) during the period.

6. Robustness

Among observed populations, under-5 sex ratios of some 3% more boys than girls are associated with relatively low rates of infant mortality (see Figure 3). In this sense, our qualitative result that US white infant mortality was relatively low in the 19th century is very robust and is not sensitive to modifications to our empirical specification (such as allowing for non-linearity, alternative regression weights, or allowing the intercept to vary across countries; see Figure A-1 in the appendix). Similarly, infant mortality rates around 200 deaths per 1,000 are associated with childhood sex ratios within 1% of parity, so the HSUS infant mortality values for the period 1850–1880 are simply inconsistent with the sex ratio evidence. However, we must address several concerns.

A central concern is the quality of the sex ratio data and the possibility that our findings are an artifact of census enumeration error.[46] Under-enumeration of young children is a common problem in historical censuses, including for the 19th-century United States (e.g., Coale and Zelnik 1963: 10–11; Hacker 2013). If enumeration of young children was biased toward males, then observed childhood sex ratios would tend to understate the level of infant mortality.[47]

To test this possibility, we compare the under-5 sex ratio in one census to two alternative indicators of childhood sex ratios. First we use the age 10–14 sex ratio in the census ten years later – essentially following the cohort across the decade for a second measure of the under-5 sex ratio (a "forward" measure). We look at the US-born white population of the nation as a whole so that immigration and interregional migration are

[45] An 80% Bayesian posterior predictive interval tends to span 40–130 deaths per 1,000 in the period 1850–1880, lining up well with the 90% upper bound.
[46] Recall that anomalous enumeration of ages in the 1890 US census strongly biased the under-5 sex ratio toward boys, which could be mistaken for evidence of very low IMR circa 1890.
[47] We thank George Alter (personal communication) for both alerting us to this problem and suggesting the use of forward sex ratios. We also thank an anonymous referee for suggesting we use sex ratios for ages 5–9.

not at play. The age 10–14 sex ratio in one census promises to be a good proxy for the under-5 sex ratio in the previous census ten years earlier: under-enumeration was much lower for ages 10–14 than for the under-5 age group (Hacker 2013, Figure 3), and child mortality after age 4 is generally both dramatically lower and less male biased than infant mortality (Hill and Upchurch 1995). We also use the age 5–9 sex ratio from the concurrent census. By the same logic outlined above, the age 5–9 sex ratio should be similar to the under-5 ratio. If a relative undercounting of infant girls in the 19th century is biasing our under-5 sex ratios toward boys – for a false impression of low infant mortality – then we should observe a relatively more female sex ratio among ages 10–14 ten years later and among ages 5–9 in the same year.

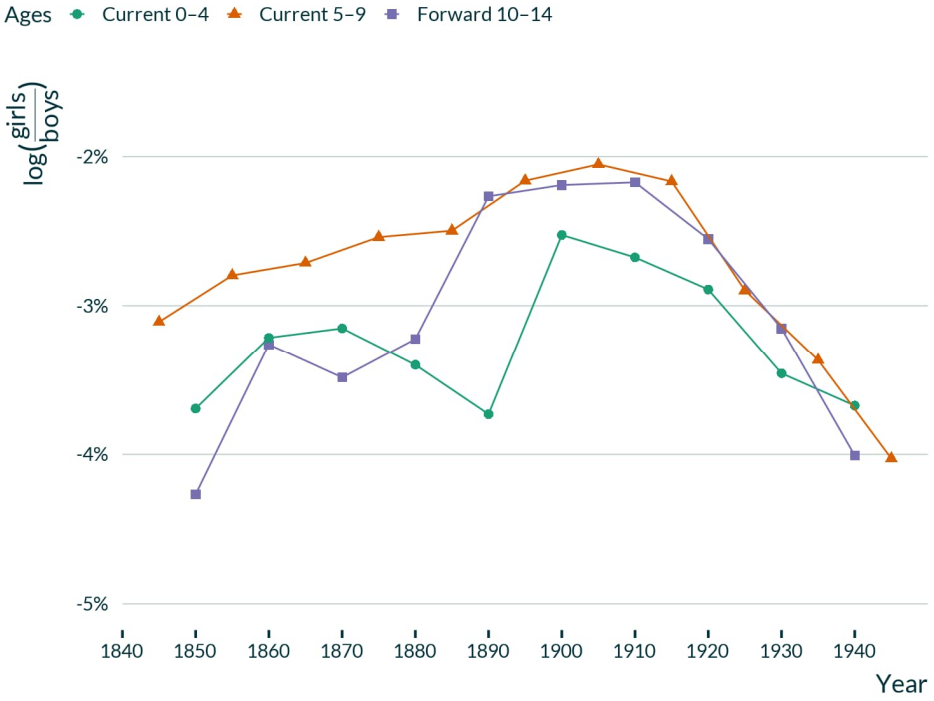
Looking across the censuses corroborates our findings: the older-age sex ratios do not tilt toward females. For each decennial census year from 1850 to 1940, Figure 6 plots the under-5 sex ratio and the age 5–9 sex ratio, along with the age 10–14 sex ratio from the next census (ten years later). All three sex ratios line up well in terms of their general pattern, with the exception of the census year 1890. There, the under-5 sex ratio is much more male, powerfully signaling a male-biased enumeration of those under 5 in 1890. As noted above, that bias is an expected result of the anomalous age question used in the 1890 census ("age at nearest birthday" instead of "age at last birthday").[48]

Comparing the three sex ratios strongly corroborates our basic results. All three measures show a population distinctly more male in the mid-19th century, with some 3%–4% more boys than girls, signaling relatively low rates of infant mortality. Furthermore, Figure 6 shows a striking inverted-U shape for both the 10–14 and 5–9 sex ratios. Infant mortality appears to have deteriorated over the second half of the 19th century before beginning its well-documented improvement in the 20th. This fits well with the prevailing view that US population health deteriorated across much of the 19th century.[49]

[48] See note 43. The wording of the age question in 1890 meant that ages 4.5 and up were excluded from the 1890 under-5 cohort, biasing its sex ratio toward male. For the 10–14 category, this age recording issue is inconsequential insomuch as the sex ratio at age 9 is very similar to that at age 14. In contrast, because of excess male infant mortality, the age 4 sex ratio is more female than the age 0 ratio.

[49] Fogel (1986), Pope (1992), and Hacker (2010) all find that life expectancy declined from 1800 to 1850. Margo and Steckel (1983) and Komlos (1987) find that adult male heights declined over the same period.

Figure 6: Alternative measures of childhood sex ratios at census benchmarks, US native-born whites 1850–1940



Note: The green line connects the under-5 sex ratio in each census year. The purple line connects the sex ratio of ages 10–14 in the following census year (ten years later). The orange line connects the age 5–9 sex ratio, displaced five years backward to reflect the birth years of the cohort. As discussed in the text, the 1890 under-5 sex ratio is male biased due to the anomalous enumeration practices of the 1890 census. Data from IPUMS and census volumes.

Another robustness concern is missing women (à la Sen 1990) and the possibility that the relative male tilt of childhood sex ratios in the United States circa 1850–1880 reflected excess female mortality in early childhood.[50] However, a wide range of evidence contradicts this possibility. For all cases where we have infant and child mortality estimates for the 19th- and early 20th-century United States, we see a clear female survival advantage in infancy.[51] Moreover, excess female mortality typically

[50] Sources of excess female mortality of infants and children range from female infanticide to sex bias in allocation of household resources, such as food and health care. See Visaria (1969: 53–54) and D'Souza and Chen (1980), among many others.

[51] For an extensive list of IMR estimates by gender in the 19th-century United States, see Haines (1977: Table 7). See US Census Office (1885: Table IV), Lynch, Mineau, and Anderton (1985: Table 4), and Ferrie (2003:

reflects extreme parental preferences toward males – son preference (Das Gupta 1987). This was not a feature of the 19th-century United States. Jones, Millington, and Price (2023) provide an authoritative account of parental gender preferences across the entire period 1850 to 1940 using full-count census microdata. They find a consistent parental preference for a mix of genders, with families more likely to have another child if their first two were the same gender, whether boys or girls (2023: Tables 2 and 5). In sum, mortality evidence shows a clear female survival advantage, and fertility patterns demonstrate a preference for mixed genders, not for sons. "Missing girls" is therefore not a concern in our study of the United States. That said, those applying our method to other settings should check for patterns of sex discrimination and son preference before naively interpreting male-skewed sex ratios as evidence of low infant mortality.

7. Discussion

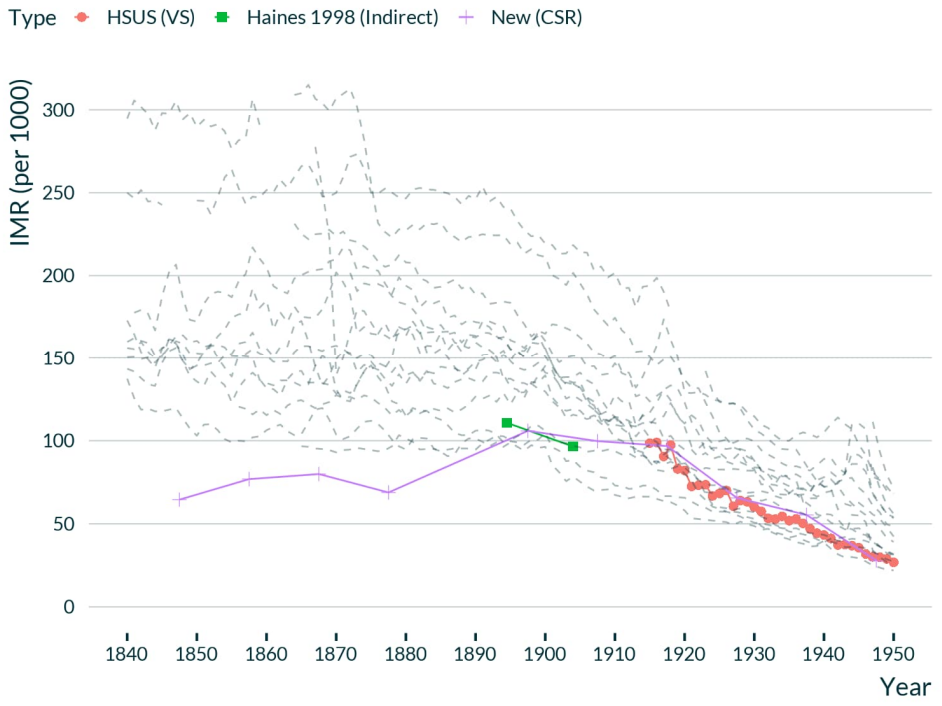
Childhood sex ratios allow us to overcome the challenge of the lack of vital statistics for the 19th-century United States and characterize levels of infant mortality. With boys outnumbering girls by more than 3% at each of the decennial censuses from 1850 to 1880, we have clear evidence that US whites were a healthy population by the standards of the 19th century. Figure 7 presents our estimates against the backdrop of the well-documented IMR of contemporary Europe. We include the HSUS estimates from 1915 forward, as well as the indirect estimates of Haines (1998) for circa 1895 and 1904. We exclude the HSUS decennial life table values (1850–1900).[52] As demonstrated above, these are both a priori uninformative and empirically implausible.

Our new estimates for the 19th century place US white infant mortality well below levels typical of contemporary Europe. Given our discussion above, this is hardly surprising, as available historical evidence, from heights to mortality, places US whites as healthier than contemporary Europeans. With infant mortality under 100 deaths per 1,000 in the 19th century, US whites appear similar to other settler populations, like that of New Zealand (with IMR averaging less than 90 in the 1880s–1890s; see appendix). Our estimates thus line up with a broad historical understanding of the 19th-century United States as a healthy place by contemporary standards (for the white population).

Table 9) for other examples of a clear female survival advantage in infancy and childhood in the 19th-century United States. For reference to "normal" rates of excess male infant and child mortality, see (among many) Hill and Upchurch (1995), Drevenstedt et al. (2008), and Alkema et al. (2014).

[52] Recall that the value for 1910 in the HSUS series (Carter et al. 2006: Series Ab921) is Haines's (1998) indirect estimate for circa 1904 (see note 17).

Figure 7: Estimates of US infant mortality against a backdrop of European experiences



Notes: The purple crosses are the new point estimates, based on under-5 sex ratios, using Equation (4). The green squares are indirect estimates of IMR from Haines (1998); the red circles are the HSUS vital statistics series. We have excluded the HSUS life table values, as discussed above. See Figure 1 for more information.

We further find that infant mortality among US whites increased across the second half of the 19th century. Our results suggest that the "antebellum puzzle" (Margo and Steckel 1983) extended to a broader "industrialization puzzle" (Komlos 1998), as US population health deteriorated during a period of tremendous economic growth.[53] US per-

[53] At least in terms of infant mortality and maternal health. Childhood sex ratios reflect *infant* mortality because the female survival advantage is greatest among the very young (neonatal). They say little about *child* mortality, which is much less male skewed and need not follow the same pattern as infant mortality (see discussion of Woods 1993, above). Further research is needed to identify plausible levels of child mortality. Recent developments in the use of linked census data to estimate child mortality (e.g., Hacker, Dribe, and Helgertz 2023) promise to fill this gap, making them a natural complement to work on childhood sex ratios. A divergence between infant mortality and older-child mortality could explain why Haines's (1979, 1998) estimates of IMR

capita income doubled from 1875 to 1910 (Carter et al. 2006: Series Ca11), yet infant mortality increased, casting doubt on simple narratives of progress, such as the McKeown thesis.[54] Instead, our results point to the challenges that modernization posed to population health. It was only after 1900 that we see the path of US infant mortality turn downward, lining up with the advent of sanitation measures in US cities (Cain and Rotella 2008). This provides another piece of evidence for the growing consensus that investments in public health measures, rather than economic growth, drove the modern mortality transition.[55]

This point is made abundantly clear when we split US whites into urban and rural populations and use Equation (4) for separate estimates of infant mortality, plotted in Figure 8. In line with existing work (e.g., Kearns 1988), we find a pronounced urban health penalty in the 19th century, with infant mortality some 80 to 100 points higher in urban than rural areas. After 1900 the urban penalty began to decline, and by 1930 it had almost disappeared, as found by Haines (2001: 47). It is clear from Figure 8 that the early 20th-century decline in infant mortality was primarily an urban phenomenon. As this was a period of rapid urbanization (the urban share of the population increased from 35% in 1890 to 56% in 1930; US Census Bureau 2012: Table 10), the reduction in urban infant mortality was the critical factor for improvements in infant mortality overall.

The estimates from Figure 8 also show that the exceptional health of US whites in the mid-19th century was largely a rural phenomenon. While US urban infant mortality was comparable to that of contemporary London (around 150 deaths per 1,000), rural US whites enjoyed levels of infant mortality well below those any major European population. Under-5 sex ratios in the rural US place infant mortality under 60 deaths per 1,000 circa 1850. For comparison, in the 19th century the healthiest Scottish counties had infant mortality above 80 (Lee 1991: Table 1), and infant mortality in rural England was around 100 (Woods, Watterson, and Woodward 1988: 353). The combination of a mostly rural population (72% in 1880; US Census Bureau 2012: Table 10) and low infant mortality in rural areas produced a rate of infant mortality for the US white population that was exceptionally low by 19th-century standards.[56]

---

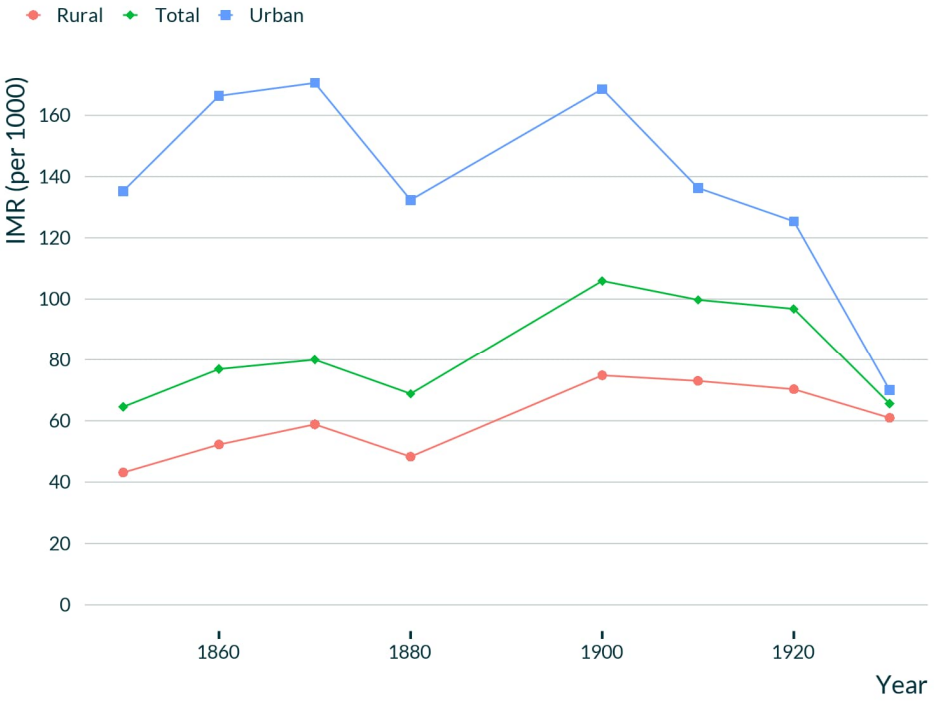
are so different from ours, as his are based on extrapolations from age 5–20 mortality. We thank an anonymous referee for pointing this out.

[54] Although mostly disregarded in public health circles (Colgrove 2002), the McKeown thesis (that economic growth drove mortality improvements in the 19th century; see McKeown 1976) continues to be evoked by economists (e.g., Anderson, Charles, and Rees 2022).

[55] See, for example, Aykroyd and Kevany (1973), Preston and Van de Walle (1978), Cain and Rotella (2001), Cutler and Miller (2005, 2022), and Alsan and Goldin (2019).

[56] Note that by our estimates, white rural infant mortality rose some 30 points from 1850 to 1900. Although still low by contemporary standards, the increased rural infant mortality merits further research. Potential explanations are found in research on the "antebellum puzzle" (Carson 2020 and references therein), which connects economic growth and expanding market access to deteriorating health in the 19th-century United States through the spread of disease (Haines, Lee, and Weiss 2003) and worsening nutrition (Komlos 1987).

Figure 8: Infant mortality estimates from under-5 sex ratios for US whites, urban vs. rural



Notes: Calculated via Equation (4). Sex ratio data from the US census: published volumes and IPUMS.

While a full explanation is beyond our scope, exceptional health is arguably unsurprising for a relatively egalitarian population in a prosperous, land-abundant economy.[57] Most simply, low population density tends to retard the spread of disease.[58] Less simply, absent coerced labor (e.g., slavery), a high land-to-labor ratio would

[57] See, for example, the discussion of Engerman and Sokoloff (2013). Smith (1776: 391) and Malthus (1798: 27) both highlighted "the plenty of good land" in the United States. This plenty, of course, was only for white settlers, as the land was violently seized from indigenous peoples (Carlos, Feir, and Redish 2022) – to say nothing of the enslaved black population of the South (see discussion below).
[58] We thank an anonymous referee for highlighting this explanation.

promote abundant nutrition and high labor incomes, both of which would enhance health.[59]

Whatever the sources of low infant mortality among US whites, they did not extend to the contemporary Black population. While childhood sex ratios provide clear evidence of relatively low infant mortality among 19th-century US whites, they corroborate the most pessimistic views of Black infant mortality under slavery.[60] In 1850 and 1860, the under-5 sex ratio of the slave population was remarkably skewed toward females, with over 2% more girls than boys, while among white children boys outnumbered girls by more than 3% (Table 2). While at the extreme of our sample, the female-skewed childhood sex ratios of the enslaved suggest an infant mortality rate of 300 or more (McDevitt-Irwin 2024). Before the abolition of slavery, the 19th-century United States featured an extreme contrast in terms of population health, with whites enjoying one of the lowest infant mortality rates in the world while enslaved Blacks suffered one of the highest. The extremes of infant mortality found in the rural United States at mid-century harshly illustrate the importance of social structures for population health, as well as the range of infant mortality possible in the preindustrial era.

8. Conclusion

Infant mortality is a key indicator of historical population health and living conditions more generally. But until now, establishing even approximate levels of infant mortality for the 19th-century United States has been an intractable problem due to a lack of data on births and infant deaths. Life table exercises (Haines 1979, 1998) have suggested a high rate of infant mortality for US whites: between 175 and 220 deaths per 1,000 in the period 1850–1880. Although published in the most recent edition of *Historical Statistics of the United States* (Carter et al. 2006: Series Ab921), such values appear implausibly high in light of a range of other evidence and known patterns of historical infant mortality.

This paper provides a partial solution to the problem of a lack of data for standard estimates (direct or indirect) of infant mortality in the 19th-century United States. We offer a new method for characterizing broad patterns of infant mortality, using childhood sex ratios from census data. Because of the well-known biological survival advantage of infant females, high rates of infant mortality tend to skew the surviving population toward girls. This theoretical relationship is strikingly evident in historical data from Europe and the United States, providing a simple means to infer infant mortality rates from under-5

[59] This point is reminiscent of Nieboer (1910: 418–419). More concretely, Ferrie (2003) and Hacker, Dribe, and Helgertz (2023) have highlighted the importance of socioeconomic status in determining mortality in the 19th-century United States.

[60] For a fuller treatment of this point, see our ongoing work (McDevitt-Irwin 2024).

sex ratios. We use quantile regression to place bounds on plausible rates of infant mortality given observed sex ratios.

The US census reveals roughly 3% more males than females under the age of 5 for 19th-century US whites. These childhood sex ratios suggest that US white infant mortality in the period 1850–1880 was less than half of the HSUS life table values: in the range of 60 to 110 deaths per 1,000 rather than 200. Using hypothesis testing, we reject at the 10% significance level an average infant mortality greater than 130 for US whites across the period 1850–1880, thus rejecting the HSUS life table values. Our results place US whites among the healthiest populations of the 19th century, with infant mortality substantially below levels found in Europe. The relative good health of US whites stood in sharp contrast to the experience of the black population under slavery: childhood sex ratios suggest that black infant mortality rates were some 250 points higher than those of the white population.

On our evidence, the history of infant mortality in the United States was not any simple variation on well-documented European patterns. In the 'pretransition' period, US whites experienced much lower infant mortality than Europeans. Moreover, the 20th-century mortality decline in the United States was preceded by a substantial deterioration of maternal and infant health. Rising infant mortality in the closing decades of the 19th century – a period of rapid economic growth and development – contradicts simple narratives of progress like the McKeown thesis.[61] Instead, our results point to the importance of public health initiatives for overcoming the challenges of mass urbanization.

Childhood sex ratios can provide a basis for characterizing infant mortality in populations lacking data on births and infant deaths. With census data often available when vital statistics are not, childhood sex ratios promise to substantially expand knowledge of infant mortality in historical populations, a fundamental indicator of population health. However, future applications must be acutely sensitive to the challenge of distinguishing between gender discrimination and low infant mortality as causes of male-skewed sex ratios.[62] More generally, when census microdata are available, childhood sex ratios offer possible insights that vital statistics do not. While birth and infant death records are often available by sex, race, and location, they are generally not at the individual level for historical populations. Sex ratios, on the other hand, can be

[61] This point echoes elements of Easterlin (1999), that economic growth alone did not lead to improved health, and Engerman (1997), that modern economic growth came with meaningful trade-offs to population well-being.
[62] We explore this issue in more depth in an existing manuscript (McDevitt-Irwin and Irwin 2022), with the example of India under the Raj. There were 6% more boys than girls in Punjab in 1910. Naively plugging this value into our estimator, you would get a very low rate of IMR. Of course, these male-skewed sex ratios reflect sex discrimination against girls, not low infant mortality. However, because such discrimination goes against girls, female-skewed sex ratios are an unambiguous sign of high infant mortality. Drawing again from our previous manuscript, there were 6% more girls than boys under the age of 5 among black South Africans in 1911, a striking indicator of extremely high infant mortality.

tabulated directly from census microdata, along any measured dimensions and along their intersection. This flexibility makes them an ideal dependent variable for quantitative social science research, particularly for the United States, where full-count census microdata are available from IPUMS for the century 1850–1950.[63]

9. Acknowledgments

This paper has evolved as part of a broader body of research on historical sex ratios as indicators of infant mortality, in process since May 2020. For helpful comments on this paper and previous versions, we thank Douglas Almond, George Alter, Vincent Bagilet, Lou Cain, Joel E. Cohen, Joseph Ferrie, Kris Inwood, Moneek Madra, Catherine McDevitt, Joel Mokyr, Benjamin Orlove, Claire Palandri, Samuel Preston, Jeffrey Shrader, Rodrigo Soares, Jenny Trinitapoli, and anonymous referees. We thank participants at the 2021 and 2022 CEA meetings, the Northwestern Economic History Workshop, the University of Chicago Demography Workshop, and the 2023 PAA meetings for their comments.

[63] For example, in our ongoing work, we study the extent to which residential and occupational patterns can explain observed racial and ethnic differences in sex ratios. To do so, we construct sex ratios by the intersection of residence, occupation, and race, something that cannot be done with vital statistics, as they are already aggregated along each of these margins.

References

Abbott, S.W. (1897). The vital statistics of Massachusetts: A forty years' summary. In: *Twenty-Eighth Annual Report of the Massachusetts State Board of Health*. (Public Document 34). Boston: Wright and Potter, State Printers: 713–829.

Alkema, L., Chao, F., You, D., Pedersen, J., and Sawyer, C.C. (2014). National, regional, and global sex ratios of infant, child, and under-5 mortality and identification of countries with outlying ratios: A systematic assessment. *The Lancet Global Health* 2(9): E521–E530. doi:10.1016/S2214-109X(14)70280-3.

Almond, D. and Edlund, L. (2007). Trivers–Willard at birth and one year: Evidence from US natality data 1983–2001. *Proceedings of the Royal Society B: Biological Sciences* 274(1624): 2491–2496. doi:10.1098/rspb.2007.0524.

Alsan, M. and Goldin, C. (2019). Watersheds in child mortality: The role of effective water and sewerage infrastructure, 1880–1920. *Journal of Political Economy* 127 (2): 586–638. doi:10.1086/700766.

Alter, G. (2004). Height, frailty, and the standard of living: Modelling the effects of diet and disease on declining mortality and increasing height. *Population Studies* 58(3): 265–279. doi:10.1080/0032472042000272339.

Anderson, D.M., Charles, K.K., and Rees, D.I. (2022). Reexamining the contribution of public health efforts to the decline in urban mortality. *American Economic Journal: Applied Economics* 14(2): 126–157. doi:10.1257/app.20190034.

Aykroyd, W.R. and Kevany, J.P., (1973). Mortality in infancy and early childhood in Ireland, Scotland and England and Wales 1871 to 1970. *Ecology of Food and Nutrition* 2(1): 11–19. doi:10.1080/03670244.1973.9990310.

Beltrán Tapia, F.J. and Cappelli, G. (2024). Missing girls in Liberal Italy, 1861–1921. *The Economic History* Review 77(1): 185–211. doi:10.1111/ehr.13257.

Beltrán Tapia, F.J. and Gallego-Martínez, D. (2017). Where are the missing girls? Gender discrimination in 19th-century Spain. *Explorations in Economic History* 66: 117–126. doi:10.1016/j.eeh.2017.08.004.

Beltrán Tapia, F.J. and Raftakis, M. (2022). Sex ratios and gender discrimination in modern Greece. *Population Studies* 76(2): 329–346. doi:10.1080/00324728.2021.1923787.

Beltrán Tapia, F.J. and Szołtysek, M. (2022). 'Missing girls' in historical Europe: Reopening the debate. *The History of the Family* 27(4): 619–657. doi:10.1080/1081602X.2022.2132979.

Bohnert, N., Jåstad, H.L., Vechbanyongratana, J., and Walhout, E. (2012). Offspring sex preference in frontier America. *The Journal of Interdisciplinary History* 42(4): 519–541. doi:10.1162/JINH_a_00303.

Cain, L. and Hong, S.K. (2009). Survival in 19th century cities: The larger the city, the smaller your chances. *Explorations in Economic History* 46(4): 450–463. doi:10.1016/j.eeh.2009.05.001.

Cain, L. and Rotella, E. (2001). Death and spending: Urban mortality and municipal expenditure on sanitation. *Annales de Démographie Historique* 101: 139–154. doi:10.3917/adh.101.0139.

Cain, L. and Rotella, E. (2008). Epidemics, demonstration effects, and investment in sanitation capital by U.S. cities in the early twentieth century. In: Rosenblum, J.L. (ed.). *Quantitative economic history: The good of counting.* New York: Routledge: 34–53. doi:10.4324/9780203928134.ch3.

Caldwell, J.C. (2001). Population health in transition. *Bulletin of the World Health Organization* 79(2): 159–160.

Caldwell, J.C. (2006). *Demographic transition theory*. Dordrecht: Springer Netherlands.

Campbell, S. (2013). A short history of sonography in obstetrics and gynaecology. *Facts, Views and Vision in ObGyn* 5(3): 213–229.

Carlos, A.M., Feir, D.L., and Redish, A. (2022). Indigenous nations and the development of the US economy: Land, resources, and dispossession. *The Journal of Economic History* 82(2): 516–55. doi:10.1017/S0022050722000080.

Carson, S.A. (2020). Net nutrition, insolation, mortality, and the antebellum paradox. *Journal of Bioeconomics* 22: 77–98. doi:10.1007/s10818-020-09293-6.

Carter, S.B., Gartner, S.S., Haines, M.R., Olmstead, A.L., Sutch, R., and Wright, G. (eds.) (2006). *Historical statistics of the United States*: *Millennial edition*. New York: Cambridge University Press.

Catalano, R.A. (2003). Sex ratios in the two Germanies: A test of the economic stress hypothesis. *Human Reproduction* 18(9): 1972–1975. doi:10.1093/humrep/deg370.

Catalano, R. and Bruckner, T. (2006). Secondary sex ratios and male lifespan: Damaged or culled cohorts. *Proceedings of the National Academy of Sciences* 103(5): 1639–1643. doi:10.1073/pnas.0510567103.

Chiswick, C.U. (2006). Appendix 3: Origin of Historical Statistics of the United States. In: Carter, S.B., Gartner, S.S., Haines, M.R., Olmstead, A.L., Sutch, R., and Wright, G. (eds.). *Historical statistics of the United States: Millennial edition*. New York: Cambridge University Press: 819–824.

Clarke, J. (1786). Observations on some causes of the excess of the mortality of males above that of females. *The London Medical Journal* 9 (Pt 2): 179.

Coale, A.J. (1991). Excess female mortality and the balance of the sexes in the population: An estimate of the number of 'missing females.' *Population and Development Review* 17(3): 517–523. doi:10.2307/1971953.

Coale, A.J. and Zelnik, M. (1963). *New estimates of fertility and population in the United States*. Princeton: Princeton University Press. doi:10.1515/9781400874934.

Colgrove, J. (2002). The McKeown thesis: A historical controversy and its enduring influence. *American Journal of Public Health* 92(5): 725–729. doi:10.2105/AJPH.92.5.725.

Condran, G.A. and Cheney, R.A. (1982). Mortality trends in Philadelphia: Age- and cause-specific death rates 1870–1930. *Demography* 19(1): 97–123. doi:10.2307/2061131.

Condran, G.A. and Crimmins, E. (1979). A description and evaluation of mortality data in the federal census: 1850–1900. *Historical Methods* 12(1): 1–23.

Condran, G.A. and Crimmins, E. (1980). Mortality differentials between rural and urban areas of states in the northeastern United States 1890–1900. *Journal of Historical Geography* 6(2): 179–202. doi:10.1016/0305-7488(80)90111-5.

Cutler, D. and Miller, G. (2005). The role of public health improvements in health advances: The twentieth-century United States. *Demography* 42(1): 1–22. doi:10.1353/dem.2005.0002.

Cutler, D. and Miller, G. (2022). Reexamining the contribution of public health efforts to the decline in urban mortality: Comment. *American Economic Journal: Applied Economics* 14(2): 158–165. doi:10.1257/app.20190711.

D'Souza, S. and Chen, L.C. (1980). Sex differentials in mortality in rural Bangladesh. *Population and Development Review* 6(2): 257–270. doi:10.2307/1972730.

Das Gupta, M. (1987). Selective discrimination against female children in rural Punjab, India. *Population and Development Review* 13(1): 77–100. doi:10.2307/1972121.

Das Gupta, M. (2005). Explaining Asia's 'missing women': A new look at the data. *Population and Development Review* 31(3): 529–535. doi:10.1111/j.1728-4457.2005.00082.x.

Davis, D.L., Gottlieb, M.B., and Stampnitzky, J.R. (1998). Reduced ratio of male to female births in several industrial countries: A sentinel health indicator? *JAMA* 279(13): 1018–1023. doi:10.1001/jama.279.13.1018.

Davis, K. (1973). *Cities: Their origin, growth, and human impact*. San Francisco: W.H. Freeman and Company.

De Jastrzebski, S. (1919). The sex ratio at birth. *The Eugenics Review* 11(1): 7–16.

Deaton, A. (2007). Height, health, and development. *Proceedings of the National Academy of Sciences* 104(33): 13232–13237. doi:10.1073/pnas.0611500104.

Drevenstedt, G.L., Crimmins, E.M., Vasunilashorn, S., and Finch, C.E. (2008). The rise and fall of excess male infant mortality. *Proceedings of the National Academy of Sciences* 105(13): 5016–5021. doi:10.1073/pnas.0800221105.

Easterlin, R.A. (1999). How beneficent is the market? A look at the modern history of mortality. *European Review of Economic History* 3(3): 257–294. doi:10.1017/S1361491699000131.

Engerman, S.L. (1997). The standard of living debate in international perspective: Measures and indicators. In: Steckel, R.H. and Floud R. (eds.). *Health and welfare during industrialization*. Chicago: University of Chicago Press: 17–46.

Engerman, S.L. and Sokoloff, K.L. (2013). Five hundred years of European colonization: Inequality and paths of development. In Lloyd, C., Metzer, J., and Sutch, R. (eds.). *Settler economies in world history*. Leiden and Boston: Brill: 65–103. doi:10.1163/9789004232655_005.

Ferrie, J.P. (2003). The rich and the dead: Socioeconomic status and mortality in the United States, 1850-1860. In: Costa, D.L. (ed.). *Health and labor force participation over the life cycle: Evidence from the past*. Chicago: University of Chicago Press: 11–50. doi:10.7208/chicago/9780226116198.003.0002.

Fogel, R.W. (1986). Nutrition and the decline in mortality since 1700: Some preliminary findings. In: Engerman, S.L. and Gallman, R.E. (eds.). *Long-term factors in American economic growth*. Chicago: University of Chicago Press: 439–556. doi:10.3386/w1802.

Fogel, R.W., Engerman, S.L., Floud, R., Friedman, G., Margo, R.A., Sokoloff, K., Steckel, R.H., Trussell, J.T., Villaflor, G., and Wachter, K.W. (1983). Secular changes in American and British stature and nutrition. *Journal of Interdisciplinary History* 14(2): 445–481. doi:10.2307/203716.

Fukuda, M., Fukuda, K., Shimizu, T., and Møller, H. (1998). Decline in sex ratio at birth after Kobe earthquake. *Human Reproduction* 13(8): 2321–2322. doi:10.1093/humrep/13.8.2321.

Gabry, J., Simpson, D., Vehtari, A., Betancourt, M., and Gelman, A. (2019). Visualization in Bayesian workflow. *Journal of the Royal Statistical Society Series A: Statistics in Society* 182(2): 389–402. doi:10.1111/rssa.12378.

Gelman, A., Hill, J., and Vehtari, A. (2021). *Regression and other stories*. Cambridge University Press. doi:10.1017/9781139161879.

Gelman, A., Jakulin, A., Pittau, M.G., and Su, Y-S. (2008). A weakly informative default prior distribution for logistic and other regression models. *The Annals of Applied Statistics* 2(4): 1360–1383. doi:10.1214/08-AOAS191.

Gelman, A., Meng, X-L, and Stern, H. (1996). Posterior predictive assessment of model fitness via realized discrepancies. *Statistica Sinica* 6(4): 733–760.

Gelman, A., Simpson, D., and Betancourt, M. (2017). The prior can often only be understood in the context of the likelihood. *Entropy* 19(10): 555. doi:10.3390/e19100555.

Goldberger, A.S. (1972). Structural equation methods in the social sciences. *Econometrica: Journal of the Econometric Society* 40(6): 979–1001. doi:10.2307/1913851.

Grech, V. and Masukume, G. (2016). The sex ratio at birth in South Africa may be a sentinel health indicator. *Early Human Development* 103: 225–227. doi:10.1016/j.earlhumdev.2016.10.008.

Grech, V., Savona-Ventura, C., and Vassallo-Agius, P. (2002). Unexplained differences in sex ratios at birth in Europe and North America. *British Medical Journal* 324(7344): 1010–1011. doi:10.1136/bmj.324.7344.1010.

Guimbeau, A., Menon, N., and Musacchio, A. (2022). Short- and medium-run health and literacy impacts of the 1918 Spanish flu pandemic in Brazil. *The Economic History Review* 75(4): 997–1025. doi:10.1111/ehr.13155.

Hacker, J.D. (2010). Decennial life tables for the white population of the United States, 1790–1900. *Historical Methods* 43(2): 45–79. doi:10.1080/01615441003720449.

Hacker, J.D. (2013). New estimates of census coverage in the United States, 1850-1930. *Social Science History* 37(1): 71–101. doi:10.1215/01455532-1958172.

Hacker, J.D., Dribe, M., and Helgertz, J. (2023). Wealth and child mortality in the nineteenth-century United States: Evidence from three panels of American couples, 1850–1880. *Social Science History* 47(3): 333–366. doi:10.1017/ssh.2023.12.

Haines, M.R. (1977). Mortality in nineteenth century America: Estimates from New York and Pennsylvania census data, 1865 and 1900. *Demography* 14(3): 311–331. doi:10.2307/2060789.

Haines, M.R. (1979). The use of model life tables to estimate mortality for the United States in the late nineteenth century. *Demography* 16(2): 289–312. doi:10.2307/2061144.

Haines, M.R. (1998). Estimated life tables for the United States, 1850–1910. *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 31(4): 149–169. doi:10.1080/01615449809601197.

Haines, M.R. (2001). The urban mortality transition in the United States, 1800–1940. *Annales de Démographie Historique* 101: 33–64. doi:10.3386/h0134.

Haines, M.R. (2006). Vital statistics. In: Carter, S.B., Gartner, S.S., Haines, M.R., Olmstead, A.L., Sutch, R., and Wright, G. (eds.). *Historical statistics of the United States*: *Millennial edition*. New York: Cambridge University Press: 380–391.

Haines, M.R. and Preston, S.H. (1997). The use of the census to estimate childhood mortality: Comparisons from the 1900 and 1910 United States census public use samples. *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 30(2): 77–96. doi:10.1080/01615449709601177.

Haines, M.R., Lee, A.C., and Weiss, T. (2003). The short and the dead: Nutrition, mortality, and the antebellum puzzle in the United States. *The Journal of Economic History* 63(2): 382–413. doi:10.1017/S0022050703001839.

Hatton, T.J. and Bray, B.E. (2010). Long run trends in the heights of European men, 19th–20th centuries. *Economics and Human Biology* 8(3): 405–413. doi:10.1016/j.ehb.2010.03.001.

Hill, K. (1995). The decline of childhood mortality. In: Simon, J.L. (ed.). *The state of humanity*. Oxford: Blackwell: 37–50.

Hill, K., and Upchurch, D.M. (1995). Gender differences in child health: Evidence from the demographic and health surveys. *Population and Development Review* 21(1): 127–151. doi:10.2307/2137416.

Human Mortality Database (HMD). Max Planck Institute for Demographic Research (Germany), University of California, Berkeley (USA), and French Institute for Demographic Studies (France). Available at www.mortality.org (data downloaded on various dates).

Inwood, K., Oxley, L., and Roberts, E. (2010). Physical stature in nineteenth-century New Zealand: A preliminary interpretation. *Australian Economic History Review* 50(3): 262–283. doi:10.1111/j.1467-8446.2010.00305.x.

James, W.H., and Grech, V. (2017). A review of the established and suspected causes of variations in human sex ratio at birth. *Early Human Development* 109: 50–56. doi:10.1016/j.earlhumdev.2017.03.002.

Jones, T.R., Millington, M.J., and Price, J. (2023). Changes in parental gender preference in the USA: Evidence from 1850 to 2019. *Journal of Population Economics* 36(4): 3057–3070. doi:10.1007/s00148-023-00957-x.

Kearns, G. (1988). The urban penalty and the population history of England. In: Brändström, A. and Tedebrand, L-G. (eds.). *Society, health and population during the demographic transition* (Umeå Universitet, Demografiska Databasen 4). Stockholm: Almqvist & Wiksell: 213–236.

Klasen, S. (1994). 'Missing women' reconsidered. *World Development* 22(7): 1061–1071. doi:10.1016/0305-750X(94)90148-1.

Komlos, J. (1987). The height and weight of West Point cadets: Dietary change in antebellum America. *The Journal of Economic History* 47(4): 897–927. doi:10.1017/S002205070004986X.

Komlos, J. (1998). Shrinking in a growing economy? The mystery of physical stature during the industrial revolution. *The Journal of Economic History* 58(3): 779–802. doi:10.1017/S0022050700021161.

Kramer, M. S. (1987). Determinants of low birth weight: Methodological assessment and meta-analysis. *Bulletin of the World Health Organization* 65(5): 663–737.

Kunze, K. (1979). The effects of age composition and changes in vital rates on nineteenth century population: Estimates from new data. PhD thesis, The University of Utah.

Lee, C.H. (1991). Regional inequalities in infant mortality in Britain, 1861–1971: Patterns and hypotheses. *Population Studies* 45(1): 55–65. doi:10.1080/00324720310001450 86.

Lemoine, N.P. (2019). Moving beyond noninformative priors: Why and how to choose weakly informative priors in Bayesian analyses. *Oikos* 128(7): 912–928. doi:10.1111/oik.05985.

Linder, F.E. and Grove, R.D. (1947). *Vital statistics rates in the United States 1900–1940*. Washington: GPO.

Lynch, K.A., Mineau, G.P., and Anderton, D.L. (1985). Estimates of infant mortality on the western frontier: The use of genealogical data. *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 18(4): 155–164. doi:10.1080/01615440.1985.10594160.

Maconochie, N. and Roman, E. (1997). Sex ratios: Are there natural variations within the human population? *BJOG: An International Journal of Obstetrics and Gynaecology* 104(9): 1050–1053. doi:10.1111/j.1471-0528.1997.tb12065.x.

Malthus, T.R. (1798). *An essay on the principle of population*. London: Printed for J. Johnson, in St. Paul's Churchyard. Project Gutenberg https://www.gutenberg.org/ebooks/4239 (released 2003; updated 2020).

Margo, R.A. and Steckel, R.H. (1983). Heights of native-born whites during the antebellum period. *The Journal of Economic History* 43(1): 167–174. doi:10.1017/S0022050700029144.

McDevitt-Irwin, J.R. (2024). US infant mortality under slavery and after emancipation: New evidence from childhood sex ratios [Unpublished manuscript]. https://jrmcirwin.github.io/sr1/UsSlv.pdf.

McDevitt-Irwin, J.R., and Irwin, J.R. (2022). Childhood sex ratios reveal infant mortality: Evidence from historical populations [Unpublished manuscript]. https://jrmcirwin.github.io/sr1/DemPap_1.pdf.

McDonald, P., Ruzicka, L., and Pyne, P. (1987). Chapter 3: Marriage, fertility and mortality. In: Vamplew, W. (ed.). *Australians: Historical statistics*. Broadway, New South Wales: Fairfax, Syme and Weldon: 42–61.

McKeown, T. (1976). *The modern rise of population*. New York: Academic Press.

Mitchell, B. (1998). *International historical statistics: Europe 1750–1993, 4th ed.* New York: Grove's Dictionaries, Inc. doi:10.1007/978-1-349-12791-7.

Nieboer, H.J. (1910). *Slavery as an industrial system: Ethnological researches*. The Hague: Martinus Nijhoff Publishers.

Oldenburg, P. (1992). Sex ratio, son preference and violence in India: A research note. *Economic and Political Weekly* 27(49/50): 2657–2662.

Orzack, S.H., Stubblefield, J.W., Akmaev, V.R., Colls, P., Munné, S., Scholl, T., Steinsaltz, D., and Zuckerman, J.E. (2015). The human sex ratio from conception to birth. *Proceedings of the National Academy of Sciences* 112(16): E2102–E2111. doi:10.1073/pnas.1416546112.

Pope, C.L. (1992). Adult mortality in America before 1900: A view from family histories. In: Goldin, C. and Rockoff, H. (eds.). *Strategic factors in nineteenth century American economic history: A volume to honor Robert W. Fogel*. Chicago: University of Chicago Press: 267–296.

Preston, S.H. and Haines, M.R. (1991). *Fatal years: child mortality in late nineteenth-century America*. Princeton, NJ: Princeton University Press. doi:10.1515/9781400861897.

Preston, S.H. and Van de Walle, E. (1978). Urban French mortality in the nineteenth century. *Population Studies* 32(2): 275–297. doi:10.1080/00324728.1978.10410715.

Preston, S.H., Heuveline, P., and Guillot, M. (2001). *Demography: Measuring and modeling population processes*. Oxford: Blackwell Publishing.

Riley, J.C. (2005). Estimates of regional and global life expectancy, 1800–2001. *Population and Development Review* 31(3): 537–543. doi:10.1111/j.1728-4457.2005.00083.x.

Ruggles, S., Nelson, M.A., Sobek, M., Fitch, C.A., Goeken, R., Hacker, J.D., Roberts, E., and Warren, J.R. (2024). IPUMS Ancestry full count data: Version 4.0 1850–1940. Minneapolis: IPUMS.

Sanders, N.J. and Stoecker, C. (2015). Where have all the young men gone? Using sex ratios to measure fetal death rates. *Journal of Health Economics* 41: 30–45. doi:10.1016/j.jhealeco.2014.12.005.

Sawyer, C.C. (2012). Child mortality estimation: Estimating sex differences in childhood mortality since the 1970s. *PLOS Medicine* 9(8): e1001287. doi:10.1371/journal.pmed.1001287.

Schneider, E.B. (2023). The determinants of child stunting and shifts in the growth pattern of children: A long-run, global review. *Journal of Economic Surveys* early view version: 1–48. doi:10.1111/joes.12591.

Sen, A. (1989). Women's survival as a development problem. *Bulletin of the American Academy of Arts and Sciences* 43(2): 14–29. doi:10.2307/3824748.

Sen, A. (1990). More than 100 million women are missing. *The New York Review of Books* 37(20): 61–66.

Shifotoka, A. and Fogarty, A.W. (2013). Is the human sex ratio at birth a sentinel health indicator that is associated with HIV/AIDS and tuberculosis prevalences in the 21st century? *Journal of Epidemiology and Community Health* 67(4): 327–331. doi:10.1136/jech-2011-200744.

Smith, A. (1776). *An inquiry into the nature and causes of the wealth of nations*. London: T. Nelson and Sons, Paternoster Row. Project Gutenberg https://www.gutenberg.org/ebooks/3300 (released 2002; updated 2024).

Ssentongo, P., Ssentongo, A.E., Djibril M., Ba, D.M., Ericson, J.E., Na, M., Gao, X., Fronterre, C., Chinchilli, V.M., and Schiff S.J. (2021). Global, regional and national epidemiology and prevalence of child stunting, wasting and underweight in low- and middle-income countries, 2006–2018. *Scientific Reports* 11: 5204. doi:10.1038/s41598-021-84302-w.

Struyck, N. (1740). *Inleiding tot de algemeene geographie, benevens eenige sterrekundige en andere verhandelingen*. Amsterdam: Issak Tirion.

Tabutin, D. and Willems, M. (1998). Differential mortality by sex from birth to adolescence: The historical experience of the west (1750–1930). In: Lopez, A.D. and Ruzicka, L.T. (eds.). *Too young to die: Genes or gender?* New York: United Nations: 17–52.

Théré, C. and Rohrbasser, J-M (2006). Facing death in the early days of life: Inequality between the sexes in enlightenment demographic thought. *The History of the Family* 11(4): 199–210. doi:10.1016/j.hisfam.2006.12.002.

Thompson, A.L. (2021). Greater male vulnerability to stunting? Evaluating sex differences in growth, pathways and biocultural mechanisms. *Annals of Human Biology* 48(6): 466–473. doi:10.1080/03014460.2021.1998622.

UNICEF (2023). Under-Five mortality rate. https://data.unicef.org/topic/child-survival/under-five-mortality/.

United Nations Department of Economic and Social Affairs, Population Division (1982). Model life tables for developing countries (ST/ESA/SER.A/77). New York: United Nations. https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/files/documents/2020/Jan/un_1982_model_life_tables_for_developing_countries.pdf.

United Nations Department of Economic and Social Affairs, Population Division (2011). Abridged life tables (1 Year Increment). New York: United Nations. www.un.org/development/desa/pd/data/model-life-tables.

United Nations Department of Economic and Social Affairs, Population Division (2022). World population prospects 2022: Summary of results. (UN DESA/POP/2022/TR/NO. 3). New York: United Nations.

US Bureau of the Census (1949). Historical statistics of the United States, 1789–1945. Washington, D. C.: US Government Printing Office.

US Bureau of the Census (1960). Historical statistics of the United States, colonial times to 1957. Washington: US Government Printing Office.

US Bureau of the Census (1975). Historical statistics of the United States, colonial times to 1970, Bicentennial Edition. Washington: US Government Printing Office.

US Census Bureau (2012). *2010 Census of Population and Housing, Population and Housing Unit Counts, CPH-2-1, United States Summary*. Washington, D.C.: Government Printing Office.

US Census Bureau (2021). Decennial census of population and housing questionnaires and instructions (https://www.census.gov/programs-surveys/decennial-census/technical-documentation/questionnaires.html accessed 2024-09-16).

US Census Office (1885). *Report on the mortality and vital statistics of the United States as returned at the tenth census (June 1, 1880) by John S. Billings, part 1*. Washington, D.C.: Government Printing Office.

US Census Office (1902). *Population part II* (Twelfth Census of the United States, 1900). Census Reports, Volume 2. Washington, D.C.: Government Printing Office.

Valente, C. (2015). Civil conflict, gender-specific fetal loss, and selection: A new test of the Trivers–Willard hypothesis. *Journal of Health Economics* 39: 31–50. doi:10.1016/j.jhealeco.2014.10.005.

Van Dijk, I.K., Nilsson, T., and Quaranta, L. (2024). Disease exposure in infancy affects women's reproductive outcomes and offspring health in southern Sweden 1905–2000. *Social Science and Medicine* 347: 116767. doi:10.1016/j.socscimed.2024.116767.

Visaria, P.M. (1967). Sex ratio at birth in territories with a relatively complete registration. *Eugenics Quarterly* 14(2): 132–142. doi:10.1080/19485565.1967.9987713.

Visaria, P.M. (1969). The sex ratio of the population of India. *Census of India 1961, Volume I* (Monograph No, 10). New Delhi: Office of the Registrar General.

Waldron, I. (1998). Sex differences in infant and early childhood mortality: Major causes of death and possible biological causes. In: Lopez, A.D. and Ruzicka, L.T. (eds.). *Too young to die: Genes or gender?* New York: United Nations: 64–83.

Wargentin, P. (1755). Anmärkningar om nyttan af årliga förtekningar på födda och döda i et land. *Sjette Stycket. Kongl. Vetenskaps Academiens Handlingar För år 1755* 16: 241–253.

Whitwell, G. and Nicholas, S. (2001). Weight and welfare of Australians, 1890–1940. *Australian Economic History Review* 41(2): 159–175. doi:10.1111/1467-8446.00080.

Williamson, J.G. (1982). Was the industrial revolution worth it? Disamenities and death in 19th century British towns. *Explorations in Economic History* 19(3): 221–245. doi:10.1016/0014-4983(82)90039-0.

Woods, R. (1993). On the historical relationship between infant and adult mortality. *Population Studies* 47(2): 195–219. doi:10.1080/0032472031000146976.

Woods, R.I., Watterson, P.A., and Woodward, J.H. (1988). The causes of rapid infant mortality decline in England and Wales, 1861–1921 Part I. *Population Studies* 42(3): 343–366. doi:10.1080/0032472031000143516.

World Bank (2022a). Mortality rate, infant (per 1,000 live births) [electronic resource]. Washington: World Bank Group; Data. https://data.worldbank.org/indicator/SP.DYN.IMRT.IN.

World Bank (2022b). Population, total [electronic resource]. Washington: World Bank Group; Data. https://data.worldbank.org/indicator/SP.POP.TOTL.

Appendix

Summary statistics

Table A-1: Summary statistics: Data for regressions

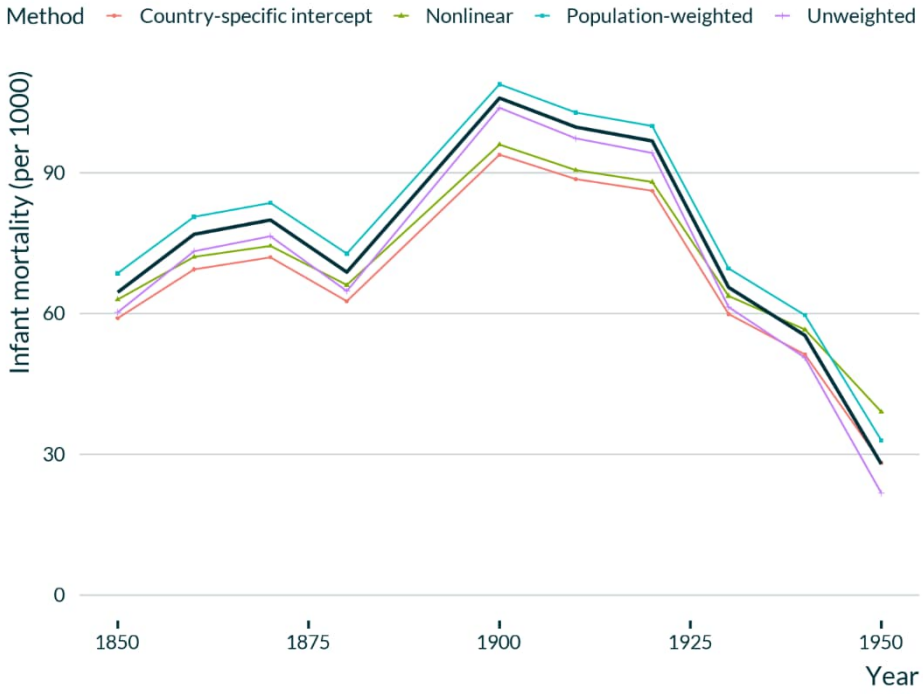
Year	pre-1849	1849–1869	1870–1899	1900–1929	1930–1961
N	41	51	264	294	117

	Min.	1st. Qu.	Median	Mean	3rd Qu.	Max.
Sex ratio (% F/M)	–5.77	–3.24	–2.00	–1.91	–0.83	5.68
IMR (per 1,000)	17	73	131	133	173	426
Population	25,324	92,327	164,589	383,990	326,182	17,358,552

Robustness

Here we present the IMR predictions of various regressions of infant mortality on childhood sex ratios. The base specification, used throughout the paper, uses all data from Figure 3, is least squares, and is weighted by the square root of the under-5 population. We conduct several robustness checks. First we allow the intercept to differ for each country. Then we allow for different weights (unweighted and weighted by total population). Finally we allow for a non-linear relationship between sex ratios and infant mortality (a cubic spline with three knots). All of the results are broadly similar and agree with our basic qualitative result: 19th-century US infant mortality was much lower than previously thought.

Figure A-1: Alternative estimates of US white IMR, based on sex ratios



Note: The black line is our main specification, used throughout in the text. It is weighted least squares; weights are equal to 1 over the square root of the sampling variance of each observation.

Bayesian model of IMR conditional on SR5

Building on Equation (2), we model the underlying data-generating process as a linear relationship between IMR and CSRs, with errors distributed normally $(0,\sigma^2)$. Furthermore, we model the implicit measurement error in observed sex ratios, which are probabilistic draws of an underlying binomial distribution. Thus we have $x_i = x^*_i + \varphi_i$, where $\varphi_i \sim N(0, g(n_i))$, x_i is observed sex ratios, x^*_i is the sex ratio of the underlying binomial draw, φ_i is classical measurement error, and $g(n_i)$ is the variance of the log sex ratio as a function of sample size. We also place a non-negativity bound on infant mortality. The full model then becomes:

$$y_i = \alpha + \beta \cdot x_i + \epsilon_i \text{ , where } y_i \geq 0,$$

$x_i \sim N(x^*_i, g(n_i))$, and $\epsilon_i \sim N(0, \sigma^2)$.

We estimate the model using Markov chain Monte Carlo, four chains with 10,000 iterations each, using the brms package in R (which calls the C++ program Stan). We use "weakly informative priors" — i.e., prior distributions that are specific enough to regularize the estimation problem but vague enough to allow the data to dominate the resulting posterior distributions — following a growing consensus in applied Bayesian statistics (Gelman et al. 2008; Gelman, Simpson and Betancourt 2017; Lemoine 2019; Gabry et al. 2019). We follow the default priors of Stan, described in Gelman, Hill, and Vehtari (2021: 124), where the variance of the priors is scaled by the variance of the data. We plot the prior vs. posterior distributions, showing that our priors are sufficiently diffuse to have minimal effect on our results, as well as posterior predictive checks (Gelman, Meng and Stern 1996), which show that our model is able to roughly reproduce the observed distribution of infant mortality. As noted above, we follow the default recommended priors in Stan (see here for discussion from the developers of Stan) and scale priors by the variance of the observed variables (Gelman, Hill and Vehtari 2021: 124). Table A-2 shows the full priors for our Bayesian model.

Table A-2: Priors for Bayesian model

Parameter	Family	Mean	Variance
Intercept	Normal	0	0.17
Slope	Normal	0	0.096
Sigma (residual variance)	Exponential	14	NA
Standard deviation (measurement error)	Exponential	1	NA
Mean (measurement error)	Normal	0	1

Here we plot prior and posterior predictive checks. In Figure A-2, we plot the prior and posterior distributions of our parameters of interest. We see that the prior distributions of our parameters are an order of magnitude more diffuse than the posterior distributions. In effect, we can see that the priors are sufficiently "weak" that they are not

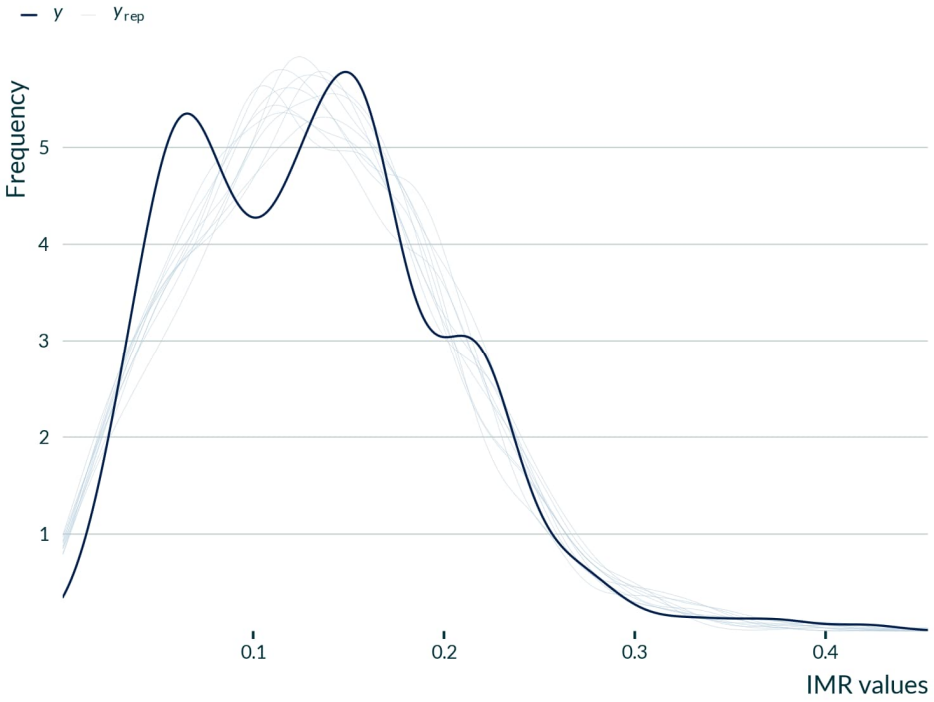
substantially influencing the posterior distributions. Instead they only regularize the estimation problem, as is desirable from "weakly informative priors" (Gabry et al. 2019).

In Figure A-3, we plot posterior predictive checks – i.e., the predicted IMR (y) values from various draws of our posterior distributions against the actual observed IMR values. We see that the model is roughly able to reproduce the underlying distribution of infant mortality.

Figure A-2: Prior vs. posterior distributions

Notes: The blue histograms are for the prior distributions for each parameter; the red are for the posterior. Beta is the slope parameter, and sigma is the residual variance parameter.

Figure A-3: Predictive posterior check for Bayesian model



Notes: The black line is the distribution of the observed IMR data from Figure 3. The shaded blue lines are the distributions of predicted values from 10 draws of our posterior distributions.

US under-5 sex ratios

We use under-5 sex ratios for US whites throughout this paper. The data are an average of IPUMS and published census volume values.

Table A-3: US under-5 sex ratios ($ln(\frac{u5\ girls}{u5\ boys}) \cdot 100$)

Year	Rural	Urban	Total
1850	–3.948	–1.876	–3.467
1860	–3.743	–1.176	–3.189
1870	–3.595	–1.081	–3.120
1880	–3.832	–1.941	–3.370
1900	–3.235	–1.126	–2.536
1910	–3.276	–1.852	–2.675
1920	–3.337	–2.098	–2.741
1930	–3.547	–3.343	–3.443

Software used

Analysis done in R version 4.4.2 (2024-10-31), with the following packages:

Table A-4: Packages

Package	Loaded version	Date	Source
brms	2.22.0	2024-09-23	CRAN (R 4.4.2)
dplyr	1.1.4	2023-11-17	CRAN (R 4.4.2)
flextable	0.9.7	2024-10-27	CRAN (R 4.4.2)
forcats	1.0.0	2023-01-29	CRAN (R 4.4.2)
ggplot2	3.5.1	2024-04-23	CRAN (R 4.4.2)
investr	1.4.2	2022-03-31	CRAN (R 4.4.2)
kableExtra	1.4.0	2024-01-24	CRAN (R 4.4.2)
lubridate	1.9.3	2023-09-27	CRAN (R 4.4.2)
mediocrethemes	0.1.3	2024-12-04	Github (vincentbagilet/mediocrethemes)
posterior	1.6.0	2024-07-03	CRAN (R 4.4.2)
purrr	1.0.2	2023-08-10	CRAN (R 4.4.2)
quantreg	5.99.1	2024-11-22	CRAN (R 4.4.2)
SparseM	1.81	2021-02-18	CRAN (R 4.4.0)
stringr	1.5.1	2023-11-14	CRAN (R 4.4.0)
tibble	3.2.1	2023-03-20	CRAN (R 4.4.0)
tidyr	1.3.1	2024-01-24	CRAN (R 4.4.0)
tidyverse	2.0.0	2023-02-22	CRAN (R 4.4.0)