Data Appendix to "Infant Mortality among US Whites in the 19th Century: New Evidence from Childhood Sex Ratios"

Data Sources

Categories of "race" follow usage in the source, unless otherwise noted.

Sources for Figures & Analyses[1]

Sources for Figure 1 (Infant Mortality Rates, 1840-1990)

US white IMR are from HSUS (2006) Series Ab921. In that series, IMR values at decennial census benchmarks 1850-1910 are from Haines (1998: 158-60, 167). As discussed in our text, the values for 1850-1900 come from Haines' model life tables (1979: 307; 1998: 158-60), based on census data on age 5-20 mortality and population by age. The value for 1910 (Haines 1998: 154, 167) is an indirect estimate of the IMR, based on census data on population-by-age and children ever-born and surviving (maternal recall). Although presented for 1910 in the HSUS series, Haines (1998: 154) reports that the value is for circa 1904 (as discussed in our text). For 1915 to 1990, the Series Ab921 data are annual, based

[1] See also "References for data sources", at the end of this appendix.

on vital statistics (registrations of births and of infant deaths). These annual data are for the "Birth Registration Area" (BRA) which covered about 1/3 of the US white population in 1915 and expanded over time, reaching complete coverage in 1933 (Linder and Grove 1947: 840-857, 1036). Figure 1 background IMR are three-year moving averages for European populations, with data from *IHS* (2013), except for England & Wales and Scotland which are from the UK Office of National Statistics (UK ONS). Austria (1840-1990), Belgium (1840-1990), Denmark (1840-1990), Finland (1867-1989), France (1840-1990), Germany (1840-1939), West Germany (1947-1988), East Germany (1947-1988), Ireland (1865-1990), Italy (1864-1990), Netherlands (1841-1990), Norway (1840-1990), Sweden (1840-1990), Switzerland (1872-1990): *IHS* Series A7.

England & Wales 1840-1990: IMR calculated from births and infant deaths from datasets at the UK ONS; Births in England and Wales: summary tables and Vital statistics in the UK: births, deaths and marriages. Scotland (1856-1990): Vital statistics in the UK: births, deaths and marriages. UK ONS datasets downloaded 19 November 2024.

Sources for Figure 2: Infant mortality by age 5–20 mortality.

Rates of infant and age 5-20 mortality are HMD estimates, from HMD life tables (Human Mortality Database, data downloaded on 2022 July 17). The data cover Australia (1921, 1925), Belgium (1841 and quinquennially 1845-1910, 1920, 1925), Canada (1921, 1925), Denmark (quinquennially 1855-1925)[2], England and Wales (1841 and quinquennially 1845-1925), Finland (1878 and quinquennially 1880-1925), France (1816 and

[2] HMD estimates of IMR in Denmark before 1855 appear to be substantially overstated (Torres 2019); see below in Sources for Figure 3.

quinquennially 1820-1925), Italy (1872 and and quinquennially 1875-1925), Netherlands (quinquennially 1850-1925), Norway (1846 and quinquennially 1850-1925), New Zealand (1901 and quinquennially 1905-1925), Scotland (quinquennially 1855-1925), Spain (1908 and quinquennially 1910-1925), Sweden (1751 and quinquennially 1755-65 and 1775-1925), and Switzerland (1876 and quinquennially 1880-1925).

The age 5-20 mortality rates for the shaded band in Figure 2 are the values from the life tables of Haines (1998:158-64) for the decennial benchmark years 1850 to 1880.

Sources for Figure 3 (Under-five sex ratios by infant mortality) and for regression analysis of the CSR:IMR relationship

The dataset for Figure 3, and used for regression analysis, is comprised of highly credible data on infant mortality rates and on childhood sex ratios. These data are direct estimates of infant mortality, taken from vital statistics, combined with under-five sex ratios from censuses or population registries. We have national-level data for Sweden (1757–1960), Denmark (1840–1960), Belgium (1846–1961), the Netherlands (1859–1960), Scotland (1861–1960), New Zealand (1867–1961), Australia (1880–1961), Switzerland (1880–1960), Finland (1885–1960), Norway (1890–1960), France (1901–1954), Italy (1911–1961), South Africa (1918–1921), Germany (1925–1960), England & Wales (1926-1961), and Austria (1930-1961). Our sub-national data include Prussian districts (1849–1910), divisions of England & Wales (1851-1921), Bavarian districts (1867–1880), Austrian Provinces (1865–1910), the State of Massachusetts (1860–1915), and various aggregates within the United States (1900–1940). Details follow.

For many historical populations, the *Human Mortality Database* provides access to official statistics for infant mortality rates and under-five sex ratios. We expand our geographic scope by also drawing on vital statistics and census data from various official sources for populations not included in the HMD.[3] In many cases, the national-level data are available from *International Historical Statistics* (Palgrave Macmillan (Ed.) 2013), which we abbreviate as *IHS* below. Specific sources and methods by polity follow.

Australia (1876–1961)

Infant mortality rates for 1876–1901 are from McDonald et al. (1987:58, series MFM 154). Rates for 1901–1961 are from Australian Bureau of Statistics (ABS), Historical Population.[4] Under-five sex ratios are paired with 5-year-average IMRs calculated from the values for the 6 years ending in the census year, weighted to reflect the census date.[5]

Under-five populations by sex are census values, with decennial data from 1881–1921 and single-year values for 1933, 1947, 1954, and 1961. The data for 1881 and 1891 are reported in Caldwell (1987:33–34). The 1901 and 1911 data are from the 1911 Census of

---

[3] The HMD is restricted to national populations "where death registration and census data are virtually complete" (HMD Overview). We include cases where the data are less "complete", requiring data only for infant mortality rates and childhood sex ratios. We also data from sub-national aggregates (the HMD has national data).

[4] Data downloads, Deaths, Table 5.4 "Infant mortality rates, states and territories, 1901 onwards", released 2019-04-18 (downloaded 2021-06-21).

[5] For example, the 1881 Census date was April 3, about ¼ into the year. For the 5-year-average IMR, the 1881 value is weighted ¼, the 1875 value weighted ¾, and those for 1876-1880 weighted 1. Australian census dates are from ABS Historical population methodology.

Australia.[6] Data for 1921, 1933, 1947, 1954, and 1961 are reported in the Census of 1966.[7] The data for Australia refer to non-Aboriginal populations.[8]

Austria, Provinces (1865–1910)

Infant mortality rates for 1865-1880 are calculated from births and infant deaths, reported annually in issues of Austria's *Statistisches Jahrbuch*.[9] Data for 1886-1910 are reported annually in the volumes of *Österreichische Statistik, Bewegung der Bevölkerung*.[10] For Provinces of Austria, we have under-five populations by sex for 1869, 1880, 1890, 1900, and 1910, from Statistics Austria.[11] There are 27 observations for Austrian Provinces in the dataset.

Austria, national data (1930-1961)

Infant mortality rates (1930–1961) are from *IHS* (2013: 3577,3580,3583), Series A7. Under-five populations by sex are for the years 1934, 1951, and 1961, reported in Statistik

---

[6] *Census of the Commonwealth of Australia taken for the night between the 2nd and 3rd April, 1911*, Vol. II, Part 1 – Ages, pp. 10-11.

[7] Commonwealth Bureau of Census and Statistics (1970), *Census of Population and Housing, 30 June 1966 Commonwealth of Australia*. Volume 1. Population: single characteristics, part 1. Age, pp. 10–11.

[8] Until 1967, population and vital statistics for Australia excluded the Aboriginal people "as required by Section 127 of the Constitution prior to its repeal in 1967" (Australian Bureau of Statistics 1968:1).

[9] E.g. the 1865 data are in *Statistisches Jahrbuch der Österreichischen Monarchie - Für das Jahr 1866* (Wien, 1868), pp. 18, 20-21. The Jahrbuch issues, whose titles vary somewhat, are available from austrian literature online.

[10] For example, the 1886 data are in *Österreichische Statistik, Bewegung der Bevölkerung der im Reichsrathe vertretenen Königreiche und Länder im Jahre 1886*. The volumes for 1886–1890, 1896–1900, and 1906–1910 are available online in the Österreichische Statistik, 1880- section of the Österreichische Nationalbibliothek.

[11] STATcube – Statistical Database of STATISTICS AUSTRIA, Dataset: Population census data since 1869 by age and Provinces, downloaded 2023-02-20.

Austria, *Statistisches Jahrbuch 2010*.[12] *IHS* (2013: 3441, Series A2) also reports these age-sex population data, but rounded down to the nearest thousand.

Belgium (1842–1961)

Infant mortality rates for Belgium for 1842–1961 are HMD estimates (downloaded 2021-10-26). Under-five populations by sex are census data, decennially 1846–1866 and 1880–1910, with single-years 1930, 1947, and 1961. The data were obtained through the HMD (downloaded 2021-07-01). The population data up to 1930 are reported in volumes of *Annuaire Statistique de la Belgique*.[13] The data for 1947 and 1961 are found in the censuses of those years. [14]

Denmark (1835–1960)

Infant mortality rates for Denmark for 1835-1854 are from *IHS* (2013: 3575, 3577; Series A7) and rates for 1855-1960 are HMD estimates (downloaded on 2021-10-26).[15] Under-five populations by sex are quinquennial 1840–1860 and 1900–1960, and decennial 1870–1890. The data were obtained through the HMD (downloaded on 2021-07-01). The

---

[12] 2.08 Bevölkerung 1869 bis 2001 nach fünfjährigen Altersgruppen und Geschlecht (Population 1869 to 2001 by five-year age groups and sex), p. 45; PDF "02 Bevölkerung" downloaded 2021-07-05.

[13] The counts for 1846 to 1890 are in the *1893 Annuaire Statistique* (p. 64); those for 1866-1900 are in the *1910 Annuaire* (pp. 143-146); the *1923-24 Annuaire* has the counts for 1910 and 1920; the 1940 Annuaire (p. 34) has the 1930 counts. These volumes are available online from the Belgian Data Portal.

[14] Institut National de Statistique (1951), *Recensement Général de la Population, de L'Industrie et du Commerce au 31 décembre 1947, tome V, Répartition de la population par âge*, Tableau 1 - Répartition des habitants par âge et sexe …" (p. 10). Bruxelles: Imprimerie Fr. Van Muysewinkel (online at KU Leuven libraries). Institut National de Statistique (1965). *Recensement Général de la Population, 31 décembre 1961, tome V, Répartition de la population par âge.* Bruxelles: Institut National de Statistique (online at KU Leuven libraries).

[15] We rely on Torres (2019) to adopt the *IHS* series for 1835-1855. After 1855 the two IMR series are very close, but for 1835-1855 the HMD estimates are some 25-50 (per thousand) above those of *IHS*. According to Torres, the HMD estimates for 1835-1855 are overstated because infant deaths included stillbirths.

Danish censuses until 1921 were taken as of February 1 of the census year (Andreev 2002: 14-15); we take the childhood sex ratios from those census data as measures for the prior year. Censuses after 1921 were taken in the second half of the year, for measures of the current year sex ratio. The data for 1840-1900 are from Danmarks Statistik (1905:55, Tabel 46). According to the HMD (DNKref.pdf, 2023 September), the data for 1901–1960 are "population estimates … produced by Danmarks Statistik", which were "obtained directly from the statistical office."

England & Wales

English sub-national data (1846-1921)

Infant mortality rates for the eleven Registration Divisions are calculated from births and infant deaths in the Annual Reports of the Registrar-General until 1910 (specific page references are available in our replication datafiles). The 1911 child-sex ratio data are paired with an average infant mortality rate for the years 1906-1910 because 1911 data were not available.[16]

We have under-five populations for the eleven Registration Divisions of England & Wales decennially from 1851 to 1911. For 1921 the data are for individual or grouped Administrative Counties (see below).

---

[16] The 1911 data on births and infant deaths were not published for the Registration Districts; starting in 1911, vital statistics reporting shifted from Registration areas to Administrative areas (1911 Annual Report of the Registrar-General, pp. vii-viii). The 1911 census data was April 2; to approximate the average infant mortality from April 2, 1906 to April 2, 1911 we take a weighted average of 1906-1910, weighting 1906 by 0.8 and the other four years by 1.05.

From 1851 to 1911 the census dates in England were on or near April 1, so about 1/4 of the census year had elapsed. Accordingly, our 5-year average infant mortality rates were constructed to reflect 1/4 of the census year, 3/4 of the year 5 years prior, and the full years in between.

Infant mortality rates for 1916-1921 are for Administrative Counties and County Boroughs, reflecting the change in 1911 from Registration to Administration areas (Seventy-Fourth Annual Report of the Registrar General (1911), pp. vii-viii) for vital statistics reporting. Infant mortality rates are calculated from births and infant deaths in the Registrar General Annual Reports, 1916-1921. With the census referring to the population as of June 20, 1921, for the prior five-year infant mortality rate we give 0.45 weight to 1921, 0.55 to 1916, and 1 to each of 1917-1920.

Under-five populations by sex for the eleven Registration Divisions are from the following publications: Census of Great Britain, 1851, Population Tables, I, Numbers of the Inhabitants, Report and Summary Tables (London 1852), pp cxcii; Census of England and Wales for the year 1861, Population Tables, Vol. II, "Ages, Civil Condition, …" (London 1863), p, xiv (Summary Tables, Table II); Census of England and Wales, 1871, Population Abstracts, "Ages, Civil Condition, …" (London 1873), p. xvi (Summary Tables, Table II); Census of England and Wales, 1881, Volume III, "Ages, Condition as to marriage, occupations …" (London 1883), pp. 3, 31, 81, 125, 165, 215, 277, 319, 375, 425, 463; Table 1 for each of the eleven Registration Divisions, in the "Divisional Tables" of Volume 3 of the 1891 Census of England and Wales ("Ages, Condition as to Marriage, …" (London: 1893),

pp. 3, 29, 85, 137, 177, 223, 289, 329, 399, 453, 491); Census of 1901, Summary Tables, Table XXVIII "Ages of persons, males and females, in registration divisions and counties" (1903: pp. 162-171); Census of England, 1911, Vol. VII, Ages and Condition as to Marriage, Table 11 (London: 1913, pp. 312-373).

The 1921 data (for Administrative Counties and County Boroughs, see above). under-five populations by sex are from the 1921 Census (General Tables, Table 37 (pp. 145-150)). Smaller counties or boroughs are aggregated with adjacent units for under-five populations over 30 thousand, referring to the "Geographical Divisions" in the 1921 census (General Tables, Table 33, pp. 140-41.[17] For England in 1921 we have 33 observations, including 6 urban areas (Birmingham, Leeds, Liverpool, London, Manchester, and Sheffield).

England & Wales, national data (1922–1961).

Infant mortality rates (1922–1961) are from the ONS Dataset "Vital statistics in the UK: births, deaths and marriages" (downloaded 2021-09-27). Under-five populations by sex for England and Wales are quinquennial for 1926–1961, from the Historic Mortality Datasets of the National Archives.[18] Five-year average values of the IMR are paired with the under-five sex ratios.

[17] R code for constructing the aggregates is included with the paper's supplementary materials.

[18] RG 69/2, Historic Mortality: 1901–1995 dataset, Population, 1901–1995 (file POPLNS.csv), downloaded 2021-06-18.

Finland (1881–1960)

Infant mortality rates (1881–1960) are HMD estimates (downloaded 2021-10-26). Under-five populations by sex are quinquennial from 1885 to 1960, obtained through the HMD (downloaded 2022-02-28). The HMD identifies Statistics Finland as the source of the data.[19]

France (1897–1954)

Infant mortality rates (1897–1968) are HMD estimates (downloaded on 2021-10-26). Under-five populations by sex are quinquennial 1901–1946, and the year 1954. The data were obtained through the HMD (downloaded on 2021-07-01), which identifies the source as Vallin & Meslé (2001).[20]

German polities

Districts of the Kingdom of Bavaria (1863-1880)

Infant mortality and under-five population data by Regierungsbezirk (Rb). We have infant mortality rates in the years 1863-1880 and under-five sex ratios for 1867 and 1880. IMRs for 1863-1867 are from data on births and infant deaths in Mayr (1878: 2-4, 14-22). For

[19] Under-five populations for 1885–1940 and 1945–1970 were received as computer files by the HMD from Statistics Finland: "Population estimates for years 1866–1940," and "Population estimates for years 1941–1995." This according to the "Data Sources" (https://mortality.org/hmd/FIN/DOCS/ref.pdf – login required) on the Finland page of the HMD website (accessed 2022-03-02.)

[20] The "Data sources" (https://mortality.org/hmd/FRATNP/DOCS/ref.pdf – login required) on the HMD data page for France describe the source as follows: "Vallin, J. and F. Meslé. (2001). Tableau I-C-1: Population par sexe et âge (de 0 à 100 ans), au 1 janvier, de 1899 à 1998, avec deux estimations selon le territoire pour les années de changement de territoire [revised post-publication]. In: Tables de mortalité françaises pour les XIXe et XXe siècles et projections pour le XXIe siècle. Paris: Institnational d'études démographiques.cite Table Tableau I-C-1: Population par sexe et âge (de 0 à 100 ans), au 1 janvier, de 1899 à 1998" (accessed 2022-03-03).

1876-1880, IMRs are from data on births and infant deaths from Bayern Statistisches Bureau (1881: 191, 198). Under-five populations by sex are from the 1867 census of Bavaria (Mayr 1872: 76-79).[21] The 1880 census data for under-five populations by sex are from Bayerisches Statistisches Landesamt (1882: 102-160).[22]

Districts of the Kingdom of Prussia (1849, 1871-1910)

We have data at the level of the Regierungsbezirk (district), with infant mortality rates for 1849-1910 and child populations by sex for the year 1849 and quinquennially from 1875 to 1910.[23] All but the 1849 data are from the "Galloway Prussia Database 1861 to 1914". The 1849 data are from the Prussian statistics bureau (Königlich Preußisches Statistisches Bureau, 1851a,b). We have a single year of infant mortality data for 1849, with births and infant deaths from Königlich Preußisches Statistisches Bureau (1851b: 10-308). Under-5 populations by sex are from Königlich Preußisches Statistisches Bureau (1851a: 274).

The Galloway database provides births and infant deaths for infant mortality annually for 1875-1910. The database has child populations by sex: under-five for 1875 and 1880, under-six quinquennially from 1895 to 1910.

With the exception of 1849 and 1875, we pair under-five sex ratios with the 5-year rolling means of infant mortality. For 1849, we have the single year of infant mortality rates to pair with the under-five sex ratios. For 1875, we pair the under-five sex ratios with the average

---

[21] Taf. III. Die Bevölkerung der einzelnen Regierungsbezirke nach Geschlecht, Civilstand und fünfjährigen Altersklassen.

[22] Tab. II. Die Bevölkerung der Regierungsbezirke und des Königreichs Bayern nach Geschlecht, Staatsangehörigkeit, Civilstand und einzelnen Geburtsjahren.

[23] We exclude the very small Sigmaringen from our data set; all the other Regierungsbezirke have under-5 populations over 25 thousand.

of the 1875 and 1876 infant mortality rates (preferring the average to a single year as an estimate for the years 1871-75).

German Republic (1921–1933)

Infant Mortality Rates (1921–1933) are from *IHS* (2013: 3577, 3580), Series A7. Under-five populations by sex are census values for 1925 and 1933; the data are from the *Statistisches Jahrbuch* of 1929 and 1939.[24] *IHS* (2013:3454, Series A2) also reports these age-sex population data, but rounded to the nearest thousand.[25]

West Germany (1956-1960)

Infant mortality rates (1956–1960) are HMD estimates (downloaded on 2021-10-26). under-five populations by sex for 1960 were obtained through the HMD (downloaded on 2021-10-26), which identifies the source as Statistisches Bundesamt.[26]

[24] The 1925 data from *1929*, p. 14; 1933 from *1939*, p. 14.

[25] The *IHS* value for 1933 differs from ours; we use the value from the 1933 census (June 16); the *IHS* values for 1933 are consistent with the estimates for Dec. 31, 1933, found in *Statistisches Jahrbuch 1936*, p. 12.

[26] Annual population estimates as of December 31st, by age (0–94, 95+) and sex. Unpublished data.

Italy (1907–1961)

Infant mortality rates (1907–1961) are from Istat Time Series.[27] Under-five populations by sex decennially 1911–1931 and 1951–1961, as well as 1936, are from Istat (Italian National Institute of Statistics), Time Series.[28]

[27] Health, Infant mortality rate by age at death and sex; perinatal mortality rate by sex - Years 1863-2013 (Table_4.8.xls).

[28] Population, Population by age class and sex, aging ratio and dependency ratio at Census from 1861 to 2011 according to reference year borders (Table_2.2.1.xls, downloaded 2021.0907).

New Zealand (1863–1961)

Infant mortality rates are for the non-Maori population from 1863–1945 and for the total population from 1947–1960. Data for 1863–1936 are from Stats NZ Store House.[29] The data for 1936–1945 are from The New Zealand Official Year-book 1957.[30] Data for 1947–1961 are for the total population (including Maori), from Stats NZ Inforshare.[31]

Under-five census populations by sex are for the non-Maori population until 1951. We have data for the single years 1867, 1874, 1881, 1936, and 1945, and quinquennially 1886–1926 and 1951–1961. The data for 1867, 1874, and 1881 are found in the 1881 census.[32] Quinquennial data for 1886–1916 are reported in the 1916 census.[33] The data for 1936, 1945, and quinquennially 1951–1961, are from the Stats NZ Store House.[34]

Netherlands (1855–1960)

---

[29] A2.7 Infant mortality rate and infant mortality number (spreadsheet), Thorns/Sedgwick non-Maori (column 3).

[30] Section 4 – Vital Statistics. European Infant Mortality.

[31] Population, Death Rates - DMM, Infant mortality rate (Annual-Dec).

[32] *Results of a census of the colony of New Zealand, taken for the night of the 3rd of April, 1881*, Chapter 28, Table 1, "Showing the Increase of Persons of Both Sexes, Males, and Females (exclusive of Maoris), at different Ages, in the Intervals between the various Censuses, from December, 1864, to April, 1881."

[33] Results of a census of the Dominion of New Zealand … 1916, Part II Ages, p. 1.

[34] Spreadsheet (182.xls) titled A1.6 Population by age and sex (Long-term data series; Population;), spreadsheet A1.6 (citing Bloomfied (1984), "Census Reports: Table II.6. Age Groups … 1874-1976").

Infant mortality rates (1855–1960) are HMD estimates (downloaded on 2021-10-26). Under-five populations by sex (1859, 1869, and quinquennially 1875–1970) were obtained through the HMD (downloaded on 2021-07-01), which identifies the sources as the NIDI mortality database for 1859–1949 and Statistics Netherlands (Centraal Bureau voor de Statistiek) for 1950–1960.

Norway (1886–1960)

Following Backer (1961), we deem credible IMR data for Norway to start with the year 1876.[35] Infant mortality rates (1886–1970) are from *IHS* (2013: 3578, 3581, 3585); Statistics Norway online data on births and infant-deaths corroborate the *IHS* infant mortality data.[36]

Under-five populations by sex are census values, decennially 1890–1930 and 1950–60, and for the year 1946.[37] Data for 1890–1900 are from Statistics Norway (1910).[38] Data for

---

[35] Although counts of births and infant deaths start with the year 1836, we are guided by the judgment of Julie E. Backer, writing as "former chief of the Population Statistics Division, Central Bureau of Statistics of Norway". According to Backer (1961, p. 36), until 1876 infants who died early inflated counts of the stillborn, with live-births and infant deaths correspondingly understated. Although some early publications from Statistics Norway report IMR data from before 1876, their *Historical Statistics* of 1978, 1994, and 2000 present 5-year average values of IMR starting with 1876. In our view, that corroborates our conclusion that 1876 marks the start of reliable IMR data for Norway.

[36] Statistisk sentralbyrå, Historisk statistikk, 3.13 Folkemengde, fødte, døde, ekteskap, flyttinger og folketilvekst.

[37] The census values refer to January 1 of a year so we treat them as the prior year's ending value (so our 1890 U5 counts are from the January 1, 1891 census). *IHS* and the HMD list Norway's population data with the census years (so our 1890 value is listed in HMD as 1891).

[38] *Norges Folkemængde fordelt paa de enkelte aldersaar, 1846-1901*, Norges Officielle Statistik. V. 113, pp. 32, 34.

1910–1930 are reported in the 1930 census.[39] The rest of the age-sex data for Norway are taken from published census volumes from the respective years: 1946 from Statistics Norway (1951), *Folketellingen 1946, Hefte 3*[40]; 1950 from Statistics Norway (1953), *Folketellingen 1950, Hefte 2.*[41]; 1960 from Statistics Norway (1963), *Folketellingen 1960, Hefte 2.*[42]

Scotland (1857–1961)

Infant mortality rates (1857–1961) are HMD estimates (downloaded on 2021-10-26). Under-five populations by sex are decennial 1861–1901 and quinquennial 1911 to 1961. The data were obtained through the HMD (downloaded on 2021-07-01); original sources are as follows. The quinquennial data for 1861 to 1881 are published in the 1881 census.[43] Data for 1891–1901 are in the 1901 census.[44] Quinquennial data for 1911 to 1936 are from

---

[39] Statistics Norway (1934), Folketellingen 1930, Hefte 5. *Folkemengden fordelt efter kjønn, alder og ekteskapelig stilling*, p. 2.

[40] *Folkemengden etter kjønn, alder og ekteskapelig stilling, …*, Tabeller p. 2.

[41] *Folkemengden etter kjønn, alder og ekteskapelig stilling …* (Population census December 1, 1950, Second volume, Population by sex, age, and marital status …), Tabeller p. 2.

[42] *Folkemengden etter kjønn, alder og ekteskapelig status*.

[43] Scotland Census Office (1883), *Ninth decennial census of the population of Scotland … 1881 … Vol. II*, Appendix tables; with the 1861 and 1871 data in Table XXII, "Population of Scotland in 1861 and 1871, in sexes and ages …" (p. xxxii) and the 1871 and 1881 in Table XXI, "Population of Scotland in 1871 and 1881, in sexes and ages …" (p. xxxii). The volume is available online from HathiTrust.

[44] Scotland Census Office (1903), *Eleventh decennial census of the population of Scotland … 1901 … Vol II*, Appendix Tables, Table 1, "Population of Scotland in 1891 and 1901, distinguishing males and females at each year of life …" (p. xxxii). Available online from Google Books.

the General Register Office for Scotland.[45] Quinquennial data for 1941 to 1961 are from General Register Office for Scotland.[46]

South Africa (1913–1921)

Data are for the white population. Infant mortality rates (1914–1921) are from *IHS* (2013:219) Series A7. Five-year average IMR are paired with under-5 childhood sex ratios. We have under-five census populations by sex for 1918 and 1921, reported in the 1922 and 1925 volumes of the *Official Yearbook* of South Africa.[47]

Sweden (1753–1960)

Infant mortality rates (1753–1960) are from Statistics Sweden.[48] We have under-five populations by sex for 1757, 1763, 1850, and quinquennially for 1785-1805, 1815-1835 and 1860-1960. Data for 1860-1960 are from Statistics Sweden.[49] For years before 1860,

---

[45] Mid-year population estimates by sex and five year age group, 1911-1938. The HMD reports these as "Retrieved 15 May 2008" http://www.gro-scotland.gov.uk.

[46] Mid-year population estimates by sex and single year of age until the last age 85+ (1939-1970) or 90+ (1971-2001); unpublished data received by HMD via email on 28 February 28, 2007.

[47] The 1918 data are in Union office of census and statistics (1923), *Official Yearbook of the Union and of Basutoland, Bechuanaland Protectorate and Swaziland, No. 5 –1922* (pp. 158–59); Pretoria: The Government Printing and Stationary Office. The 1921 data are in Union office of census and statistics (1927), *Official Yearbook of the Union and of Basutoland, Bechuanaland Protectorate and Swaziland, No. 8 –1925* (p. 868); Pretoria: The Government Printing and Stationary Office.

[48] Statistical Database, Population, Population statistics, Deaths, Live births, stillbirths and infant mortality rates by sex. Year 1749–2020 (accessed 2023-09-15).

[49] Statistical Database, Population, Population statistics, Number of inhabitants, Population by age and sex. Year 1860–2021 (accessed 2022-02-28).

we relied on Sundbärg's (1909) work.[50] We use "official" counts (1909: 180), but exclude those we deem unreliable, based on comparisons to Sundbärg's (1909: 208, 216, 224) "corrected" counts.[51] The latter figures are used as "Input Data" for Sweden by the HMD.

Switzerland (1876–1960)

Infant mortality rates (1876–1960) are calculated from data on births and infant-deaths from Historical Statistics of Switzerland, Marriage, Birth, and Death.[52] These IMRs are corroborated by *IHS* (2013: 3578,3582) Series A7. We have under-five populations by sex for 1880, 1888, and 1941; and decennially 1900–1930 and 1950-1960. The data are from Historical Statistics of Switzerland, Population.[53]

Areas of the United States

---

[50] We relied on a variety of internet translation sites to access Sundbärg's tables and discussion, which are in Swedish.

[51] We compared the under-five sex ratios of the official and corrected counts; we deemed data unreliable when the sex ratios differed by more than 0.5% (log basis).

[52] HSSO, 2012. Tab.C.41. hsso.ch/2012/c/41 (Total Deaths (Excluding Stillborn Births) by Age Group 1867–1995) and HSSO, 2012. Tab.C.5a hsso.ch/2012/c/5a (Marriage, Birth, and Death 1867–1995: General Overview).

[53] HSSO, 2012. Tab. B.8a. hsso.ch/2012/b/8a (Total Residential Population by Age in Five Year Increments (Approximate Ages), 1860–1990)

Except as otherwise noted, we use the 20th-century US vital statistics definition of urban, referring to cities with population 10,000 or more.

State of Massachusetts (1856–1925)

We use state totals quinquennially 1860-1895 and 1905-1915. We do not use the state totals for 1900, 1920, and 1930; for those years Massachusetts areas are included in various regional aggregates (see below). The state-level data are for the total population (white and nonwhite).[54] Infant mortality rates (1856–1920) for the state are from *HSUS (2006)* Series Ab928.[55] Massachusetts state censuses provide under-five populations by sex decennially 1865-1925.[56] The US federal censuses include the state's data decennially for 1860-1890 and 1910.[57] We average the values from published federal census volumes with the available IPUMS full count data (1860-1880, 1910).[58]

---

[54] The nonwhite population of Massachusetts was too small to affect the patterns of interest and appropriate vital statistics (births and infant deaths) often are not available by race.

[55] The data for 1856-1941 are from Massachusetts vital statistics; after 1942, data are from US vital statistics.

[56] *Abstract of the Census of Massachusetts, 1865*, p. 2; *The census of Massachusetts: 1875, Volume I, Population and social statistics*, p. 269 (the published total for age-one females corrected from 15589 to 13589 via pp. 263-68); *The census of Massachusetts: 1885, Volume I, Population and social statistics, Part 1*, p. 434; *Census of the Commonwealth of Massachusetts: 1895, Volume II, Population and social statistics*, p. 422; *Census of the Commonwealth of Massachusetts 1905, volume 1, population and social statistics*, p. 480; *The decennial census 1915*, p. 478. These are available online

[57] *Ninth Census, Volume II, The Vital Statistics of the United States*, Table XXIII, pp. 563, 575 (data for 1860 as well as 1870). *Statistics of the population of the United States at the tenth census (June 1, 1880)*, Table XXI, p. 592. *Report on the population of the United States at the eleventh census: 1890, Part II*, Table 3, pp. 104–105. *Twelfth census of the United States, taken in the year 1900, Population Part II* (Census Reports Volume II), Ages, Table 3, pp. 110–111. *Thirteenth census of the United States taken in the year 1910, volume 1, population 1910, General Report and Analysis*, Table 43, p. 380.

[58] Ruggles et al. (2024). The 1890 census manuscripts have not survived, so there is no full count data for that year.

Other states or areas of the United States (1900, 1920, 1930, 1940)

US areas in 1900 include 23 observations, built from data in the census of 1900, which includes data for infant mortality rates and under-5 sex ratios for states and some cities, with urban and rural breakdowns of state-level data.[59] The areas consist of: rural counties of Northern New England (ME, NH, VT); rural counties of Southern New England (CT, MA, RI); Boston MA; other MA urban; other New England urban; NY rural; Brooklyn NY; Manhattan NY; other New York City; other NY urban; NJ rural; NJ urban; Philadelphia PA; other PA cities (registration cities with population over 4,000); MI rural; MI urban; Cleveland & Cincinnati; Chicago; Milwaukee & Minneapolis & St Paul; St Louis; other Midwestern cities (registration cities with population over 4,000); registration cities of the South; registration cities of the West. The 1900 aggregates were formed to achieve a minimum under-five population over 49,000 in order to reduce the role of random variation in sex-ratio data.

Infant mortality rates are single-year values for 1900 calculated from births and infant deaths reported in US Census Office (1902), Twelfth Census, Census Reports Volume III, Vital Statistics Part 1, Table 19. Under-five populations by sex are from the same source.

US areas in 1920 include 37 observations. They include rural and urban parts of MA, NY, PA, MD, IN, MI, OH, WI, and CA. The urban parts of the states are exclusive of larger cities, which are included separately. The largest cities enter individually: Boston, Brooklyn, Chicago, New York City, Philadelphia, and Pittsburgh. Smaller cities are included in urban aggregates, as follows: other MA urban, urban CT, other urban New England; urban KS &

[59] R code for constructing the aggregates is included with the paper's supplementary materials.

MN; urban areas of the South; urban WA & OR. We also have: rural northern New England (ME, NH, VT), rural CT & RI, the rural parts of each of KS, MN, and VA; rural WA & OR; and the state of UT.[60] The 1920 aggregates were formed to achieve a minimum under-five population over 49,000 in order to reduce the role of random variation in sex-ratio data.

For 1920, infant mortality rates are calculated from on births and infant deaths for 1915–1919, taken from annual reports of birth statistics for the BRA.[61] The 1920 US census data refer to population as of January 1, 1920 so we take the simple averages (of births and of infant deaths) for the 5 years from 1915 to 1919.

Under-five populations by sex are built up from the IPUMS 1920 full count data (Ruggles et al. 2024).

US areas in 1930 include 66 observations. The aggregates were formed to achieve a minimum under-five population over 49,000 in order to reduce the role of random variation in sex-ratio data. The aggregates include rural and urban parts of New Jersey, New York, Pennsylvania, Illinois, Indianna, Michigan, Ohio, Wisconsin, Iowa, Missouri, Washington, and California. The urban parts are exclusive of larger cities, which are included separately. The largest cities were entered individually: New York City, Chicago, Detroit, Philadelphia, Los Angeles, Cleveland, Boston, Pittsburgh, St Louis. Smaller cities were grouped to varying degrees, as follows: Minneapolis & St Paul; San Francisco & Oakland; Baltimore & Washington DC, and other southern cities (New Orleans, Louisville,

[60] R code for constructing the aggregates is included with the paper's supplementary materials.

[61] US Bureau of the Census, *Birth statistics for the registration area of the United States* : 1915, first annual report (Washington: GPO, 1917); 1916, second annual report (1918); and *Birth statistics for the birth registration area of the United States* 1917, third annual report (1919); 1918, fourth annual report (1920); 1919, fifth annual report (1921). These are available online at HathiTrust).

Atlanta, Memphis, Nashville). Cities smaller than those above are included in various urban aggregates, as follows: urban Massachusetts excluding Boston; urban New England excluding Massachusetts; West North Central urban (excluding Iowa and Missouri, included above); South Atlantic urban; other urban South (urban areas of states in the East South Central and West South Central census Divisions, exclusive of cities mentioned above). For 1930, we also have the rural parts of the states of Kansas, Minnesota, Nebraska, North Dakota, Virginia, Alabama, Arkansas, Florida, Georgia, Louisiana, Mississippi, North Carolina, South Carolina, Kentucky, Tennessee, and West Virginia.[62] Rural aggregates (for under-five populations over 49,000) include northern New England rural (ME, VT, NH), southern New England rural (CT, MA, RI), and rural Maryland & Delaware. With very small urban populations, we aggregated the smaller states Idaho & Utah, and Montana & Wyoming. Finally, for each of Colorado, New Mexico, and Oregon we use the entire state, because the urban portions fell well below our 49,000 population-size threshold.[63]

The 1930 data for California, Colorado, and New Mexico refer to total populations (white and nonwhite). Colorado births and infant deaths are not presented by race in 1930. For the other states, total populations are used because the 1930 census (unlike other censuses) classified persons deemed "Mexican" as non-white.[64]

---

[62] The urban parts of these states fell below our 49,000 population threshold, so they are included in urban aggregates (described above).

[63] R code for constructing the aggregates is included with the paper's supplementary materials.

[64] See e.g. the 1940 Census (1943), *Population Volume 2, Characteristics of the population …, Part 1: United States Summary …* , p. 3). The 1940 census includes various corrected counts for the 1930 census, with "Mexicans" classified as "white" as in the census years other than 1930.

For 1930, infant mortality rates are calculated from births and infant deaths for 1925–1930, taken from annual reports of birth statistics for the BRA.[65] The 1930 US census data refer to the population as of April 15, 1930; for an appropriate average IMR, we take weighted averages (of births and of infant deaths) across the 6 years 1925-1930; 1925 is weighted 260/365 of one-fifth, 1930 is weighted 105/365 of one-fifth, and the other 4 years each weighted one-fifth (treating April 15 as 105 days of the year). Under-five populations by sex are from the IPUMS 1930 full count data (Ruggles et al. 2024).

US states 1940. We have 46 observations after setting a minimum under-five population of 25,000 to reduce the role of random sex-ratio variation (Nevada, Delaware, and Wyoming being too small).

Under-five (white) populations by sex for 1940 are the simple average of the values from the published census and the IPUMS full count sample (Ruggles et al. 2024). The published census values are taken from the 1940 census, *Population Volume II, Characteristics of the Population, Sex, Age, Race, …*. That volume is presented in seven Parts; Table 7 for each state has data for the under-five white population by sex, DC's data appears in DC's Table 3.

Infant mortality rates for 1935–40 are taken from Linder & Grove (1947: Table 28 (pp. 578–605)). The 1940 census date was April 1, so for the 5-year average IMR we weight 1940 by 1/4 of one-fifth, 1935 by 3/4 of one-fifth, and the years 1936-39 by one-fifth each.

---

[65] US Bureau of the Census, *Birth, stillbirth, and infant mortality statistics for the birth registration area of the United States* 1925, eleventh annual report, part 1 (Washington: GPO, 1927); 1926, twelfth annual report, part 1 (1929); 1927, thirteeth annual report, part 1 (1930); 1928, fourteenth annual report (1930); 1929, fifteenth annual report (1932); 1930, sixteenth annual report (1934). These are available online at HathiTrust)

Sources for other Figures (4-8)

See text for sources for Figures 4, 5, 6, 7, and 8, which draw on data detailed above or below.

Sources for Childhood Sex Ratios in the US

We draw on the decennial US censuses for under-five populations by sex, with two broad sources: published US census volumes and IPUMS "full count data" (Ruggles et al. 2024). IPUMS full count data are available for decennially 1850-1880 and 1900-1940. For these years, we average the census volume and the IPUMS full count values of under-five populations by sex, taking each as a plausible tally of the underlying census manuscripts.

Under-five populations by sex for 1850, 1860, and 1870 for the United States and for states, are reported in the *Ninth Census – Volume II. The Vital Statistics of the United States*: Tables XXIII (all races), XXVI (whites), and XXIX (Blacks).[66]

Under-five populations by sex for 1880 are reported in *Statistics of the Population of the United States at the Tenth Census (June 1, 1880)*.[67] National totals (white and nonwhite) are reported for single years of age in Table XX.[68] Table XXI reports state totals for these data.[69]

---

[66] US Census Bureau 1872, pp. 563, 575, 610, 619, 649, 658. "Race" categories follow usage in the source.

[67] The US Census website refers to this volume as "1880 Census: Volume 1. Statistics of the Population of the United States".

[68] Table XX. Population of the United States, by specified age, sex, race, … 1880; pp. 48-49

[69] Table XXI. Population, by specified age, sex, race, … 1880; pp. 552-645

IPUMS "full count samples" (Ruggles et al. 2024) are available decennially for 1850–1880, for non-slave populations, and decennially from 1900–1940 (the 1890 census manuscripts have not survived (Blake 1996)). The access dates for the full count samples employed for the paper are as follows: 1850, 1860, 1870: 2022.0204; 1880, 1900: 2024-12-04; 1910 2023-09-07; 1920 2024-12-04; 1930,1940: 2024-12-03.

References for Data Sources

Andreev, K.F. (2002). *Evolution of the Danish population from 1835 to 2000*. Monographs on Population Aging, 9. Odense: Odense University Press.

Australian Bureau of Statistics (1968). Summary of vital and population statistics, Dec 1968 (Publication 3212.0, archived content).

Backer, J.E. (1961), *Dødeligheten og dens Årsaker i Norge 1856–1955 (Trend of mortality and causes of death in Norway 1856–1955).* Samfunnsøkonomiske studier nr. 10, Statistisk Sentralbyrå, Oslo. Available in Statistics Norway's digitized publications, SØS 10 (accessed 2021-10-05).

Bayern Statistisches Bureau (1881). *Zeitschrift des Königlich-Bayerischen Statistischen Bureaus* (13). München: Adolf Ackermann. Available online from MDZ.

Bayerisches Statistisches Bureau (1883). *Bayerische Bevölkerung nach Geschlect, Alter und Civilstand.* Beiträge zur Statistic Bayerns, Heft 46. München: Wolf & Sohn. Available online via HathiTrust.

Blake, K. (1996). "First in the Path of the Firemen": The fate of the 1890 population census. *Prologue | Genealogy Notes 28(1)* [electronic resource]. Washington: National Archives. https://www.archives.gov/publications/prologue/1996/spring/1890-census-1.html

Caldwell, J. C. (1987). Chapter 2, Population. In: Vamplew, W. (ed.). *Australians: Historical Statistics*. Broadway, New South Wales: Fairfax, Syme & Weldon: 32–41.

Danmarks Statistik (1905). *Befolkningsforholdene i DK i det 19. Aarhundrede.* København: Bianco Lunos Bogtrykkeri. Available online from Danmarks Statistik.

Galloway, P. R. (2007). Galloway Prussia Database 1861 to 1914", https://www.patrickrgalloway.com/prussia.htm (downloaded 2021-07-01).

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.

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.

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

Königlich Preußisches Statistisches Bureau (1851a). Tabellen und amtliche Nachrichten über den Preußischen Staat. 1849 1. Berlin: A.W. Hayn. Available online from MDZ.

Königlich Preußisches Statistisches Bureau (1851b). Tabellen und amtliche Nachrichten über den Preußischen Staat. 1849 2. Berlin: A.W. Hayn. Available online from MDZ.

Linder F. E. and Grove, R. D. (1947). *Vital Statistics Rates in the United States 1900-1940.* Washington: GPO (available online from the NBER).

Mayr G. (1872). Die Volkszählung im Königreiche Bayern vom 3. December 1867. 2, Die bayerische Bevölkerung nach Alter, Civilstand und Geschlecht. München: Gotteswinter & Mössl. Available online from MDZ.

Mayr G. (1878). Die bewegung der bayerischen bevölkerung in den jahren 1862/63 bis 1875. München: Ackermann Available online from MDZ.

McDonald P., Ruzicka, L. & Pyne, P. (1987). Chapter 3, "Marriage, fertility and mortality". In Vamplew (1987:42-61).

Mitchell, B. R. & Deane, P. (1962). *Abstract of British Historical Statistics*. Cambridge, England: Cambridge University Press.

Palgrave Macmillan (Ed.) (2013). *International Historical Statistics* (eBook). London, England: Palgrave MacMillan Ltd. DOI 10.1057/978-1-137-30568-8.

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 [dataset]. Minneapolis, MN: IPUMS. https://doi.org/10.18128/D014.V4.0

Sundbärg, Gustav (1909). "VII. Folkmängden i åldrarna 0/1 år, 1/3 år och 3/5 år". In *Fortsatta bidrag till en svensk befolkningsstatistik; Statistisk tidskrift, Häft. 146, Årgångens No 3.* Stockholm: Kungl. Statistiska Centralbyrån (pp. 177-85).

Torres, C. (2019). Research note: Infant mortality in Denmark from 1835 to 1854 using the Human Mortality Database [unpublished manuscript]. Interdisciplinary Centre on Population Dynamics, CPop, University of Southern Denmark. doi:10.31235/osf.io/guq6r.

US Bureau of the Census (1975). *Historical statistics of the United States, colonial times to 1970, Bicentennial edition, part 1*. Washington, D.C.: US Government Printing Office. Available online at the United States Census Bureau.

US Census Bureau (1872). *Ninth Census – Volume II. The Vital Statistics of the United States,* (Washington: GPO).

US Census Bureau (1883). *Statistics of the Population of the United States at the Tenth Census (June 1, 1880)* (Washington: GPO).[70]

US Department of Health, Education, and Welfare (1954). History and organization of the vital statistics system. Chapter 1 (pp. 2–19) of *Vital Statistics of the United States 1950 Volume 1*. Washington D.C.: United States Government Printing Office.

US Department of Health, Education, and Welfare (1996). *Vital Statistics of the United States 1992, Volume II – Mortality*.

[70] The US Census website refers to this volume as "1880 Census: Volume 1. Statistics of the Population of the United States".