Sex Ratios and Infant Mortality: Evidence from Historical Vital Statistics

Jesse McDevitt-Irwin and James R. Irwin*

18 October, 2021

Abstract

Sex ratios can reveal broad patterns of infant mortality. While extremes of male-biased sex ratios are familiar as indicators of missing women, sex ratios also reflect infant mortality–making them a promising health indicator for populations without vital statistics. We outline the demography and biology of sex ratios, and their relation to maternal health and infant mortality. Using historical data from Europe and its settler colonies, we show that under-five sex ratios can predict infant mortality, and quantify the uncertainty in such predictions. Being widely available from census data, sex ratios can shed new light on infant mortality in a wide range of historical and contemporary settings.

A work in progress. Comments and suggestions welcome.

*Authors' information: Corresponding author is Jesse McDevitt-Irwin, PhD Student in Sustainable Development at Columbia University (jrm2275@columbia.edu). James R. Irwin is Professor Emeritus (Economics), Central Michigan University.

1 Introduction

We present a novel indicator for infant mortality– under-five sex ratios –and demonstrate its effectiveness on historical data.[1] Infant mortality is a key indicator of population health and living standards (Reidpath and Allotey (2003); OECD and Organization (2020)), as well as a driving force in demographic dynamics (Coale 1989). Yet, the level of infant mortality remains virtually unknown for most of the world before 1950, our understanding severely constrained by a lack of data. Developing and maintaining civil registration and vital statistics remains a continuing challenge today, with more than 30% of births and deaths going unregistered (data from the World Bank; or 30% and 60% if we take (Mikkelsen et al. 2015)).

Our proposed indicator, under-five sex ratios, can be calculated from published census data, which is widely available for a wide range of populations who lack vital statistics. For example, India in the late-19th century, much of Latin America in the early 20th, and North America before the 1920s all had censuses but lacked vital statistics. Sex ratios have the potential to substantially advance our understanding of maternal-infant health during industrialization, colonization and modernization.

From the 1960s onward, indirect methods (e.g. model life tables or Brass's Logit model) provide estimates of infant mortality in many countries without vital statistics. However, such methods require sufficient data (such as surveys of women's birth outcomes) and strong assumptions (e.g. over fertility and the age-path of mortality). Sex ratios require less data and alternative assumptions, making them complementary to existing, indirect methods of infant mortality estimation.

We assemble historical data from Europe and settler colonies on infant mortality and under-five sex ratios, restricting attention to data which we deem reliable. A simple scatter-plot illustrates the close relationship between the two variables. Using out-of-sample testing, we show that sex ratios can reveal *broad* patterns of infant mortality, particularly in high-mortality populations. As an illustration of the promise for sex-ratio evidence to reveal maternal-infant misery, we apply our results to the 1911 census of South Africa, finding that infant mortality among the Black South Africans was on the order of 350-400 deaths per 1000 births.

In sum, sex ratios can reveal broad patterns of infant mortality and shed light on times, places and peoples previously understudied. The remainder of this essay goes as follows: Section 2 an overview of the demography and biology behind the relationship between infant mortality and sex ratios, Section 3 an empirical demonstration of this relationship using data from the 19th and 20th centuries, and Section 4 a discussion of future research.

2 Infant Mortality, Sex ratios and Maternal Health

Sex ratios are familiar as indicators of missing women,[2] but they are also closely tied to infant mortality. Because female infants have a biological mortality advantage, high rates of infant mortality

[1]In the course of our research we have found several papers which pay attention to sex-ratio data as evidence of maternal or infant health. Tavassolli et al seek to explore effects of industrial pollution on "fetal death and infant mortality", highlighting male-biased fetal loss from environmental shocks. The work of Beltrán Tapia and co-authors (2017, 2020, 2021a, 2021b) makes extensive note of the relationship between infant mortality and child sex ratios. However, the authors do not pursue the potential of sex ratios for revealing infant mortality and maternal health. Instead, they focus on detecting sex discrimination against girls, making it a natural complement to our focus on infant mortality and maternal health.

[2]Among many possible, see e.g. Sen (1990), Klasen (1994), Klasen & Wink (2002) for research on "missing women".

skew the surviving population toward females – absent extremes in sex discrimination against girls. Scholars have long known that boys are both born and buried more often than girls.[3] This male disadvantage is greatest in the first year of life, with mortality typically 15-30% greater than that of females (Drevenstedt et al. (2008); Alkema et al. (2014)).[4]

The lower mortality of infant girls has powerful demographic implications for sex ratios. If a population has very low infant mortality, say 5 per 1000 like Canada in 2018, the relative numbers of boys and girls during childhood will be nearly the same as it was at birth. If, on the other hand, a population has an infant mortality rate of 200 per 1000, as was common in 19th-century Europe, then the observed sex ratio during infancy will be distinctly more female than it was at birth, male infants having been culled from the population at a greater rate than females. This effect is central to our analysis, and it is evident in model life tables. For example, in the Coale-Demeny Model West life tables, moving from level 22 to level 11, infant mortality rises from 27 to 160 (per thousand), and the infant sex ratio moves 1.55 percentage points toward girls.[5]

Beyond the effect of infant mortality, the sex ratio at birth reflects maternal health. Put simply, poor maternal health tends to tilt the sex-ratio at birth toward females.[6] In healthy populations, there are about 5% more male live births than female, but a growing body of research demonstrates that maternal stress, pollution and poverty all lead to relatively more female births.[7] Therefore, female-skewed sex ratios are signals not only of infant mortality but also maternal health.[8] Given our focus on infant mortality in this paper, we tend to abstract from the potential role of changes in the sex ratio at birth, but it is worth stating our maintained assumption that the health of infants and their mothers were closely tied together in the populations we study.[9]

2.1 Potential Concerns

While sex ratios offer great promise as indicators of maternal-infant health, they have several key limitations. Most important is the role of sex discrimination against girls, including infanticide and mortal neglect, in pushing sex ratios towards boys. Populations with high levels of infant mortality can have healthy-looking infant sex ratios if sex discrimination offsets the biological

[3]This knowledge was nicely summarized in the report on the 1861 census of Scotland: "At birth 106 males appear for every 100 females ... yet, by the law of nature, a law seen in still more powerful operation even in intra uterine life, the tendency to succumb under disease is so much greater in the male than in the female, that by about the 18th year of life the proportion of the sexes becomes equalized, 100 males being alive at that age for every 100 females." These points have been noted since at least John Graunt (1662), followed by Arbuthnot (1710), and Clarke (1786) (cited in Bakwin, 1929). See Grech (2020) for "historical aspects" of the sex-ratio at birth (and excess male mortality, to a lesser extent).

[4]The female mortality advantage in infancy is due to multiple factors: females have fewer congenital diseases owing to their redundant X chromosome, and they are also more resistant to infectious disease. For an authoritative review, see Waldron (1998: 64-83).

[5]Coale and Demeny (1983 pp. 47, 52). The ln-change of the infant sex ratio follows from the ${}_1L_0$ values by sex. The same pattern is evident in a wide range of other life tables, including the UN model life tables.

[6]The apparent mechanism is maternal stress hormones, which increase the probability of miscarriages, which are disproportionately male (James and Grech 2017: 51).

[7]Important contributions include Catalano et. al. (2003, 2005, 2005, 2009, 2013) and Almond and Edlund (2007); see more references there. More recently, Morse and Luke (2021) have argued that population-level differences in sex ratios at birth come from maternal health and fetal loss.

[8]Recent work in health economics has followed this logic, using sex ratios at birth as an indicator of fetal loss. See, for example, Sanders and Stoecker (2015), Valente (2015) and Guimbeau et. al. (2020).

[9]Separate attention to birth sex ratios and survival sex ratios may prove a fruitful avenue for future research. But for now we are satisfied that our sex-ratio evidence is informative on infant mortality narrowly, as well as maternal and infant health more generally.

mortality advantage of infant girls. For example, the 1891 British census of India reports almost 5% more boys than girls under two-years of age in the "Panjáb" province – a value consistent with a healthy population. However, looking across ages, the Panjáb sex-ratio data suggest a severe case of "missing girls." The under-five sex ratio features 8% more boys than girls, and among five-to-ten year-olds, there were 18% more boys than girls. Although we do not know the set of sex-selective mechanisms behind these patterns, the Panjáb in 1891 is a striking example where extremes of sex-discrimination generate a false negative result when testing for maternal-infant misery with the infant sex ratio. It is important to check sex ratios at older ages, as, in general, sex-discrimination is evident in a pattern of increasingly male-biased sex ratios after infancy.[10] On the other hand, because societal gender discrimination operates against girls, not boys, it follows that female-biased sex ratios are an unambiguous sign of poor maternal-infant health.

In addition to the effects of gendered social practices, there are several practical limitations to the use of sex ratios as indicators of maternal-infant health. First, large sample sizes (tens of thousands) are needed for a sex ratio to have reasonably sized confidence intervals.[11] Second, small changes in infant mortality may prove difficult to discern from sex ratios due to random variation in the sex ratio at birth. Third, sex ratios are also vulnerable to sex-biased age heaping. In our own work on the US census (McDevitt-Irwin and Irwin, 2021, unpublished manuscript), this bias manifests at ages 0 and 1. In order to combat these concerns, we focus our attention on the sex ratios for ages 0 to 4, which reduces both age heaping and sampling variation.

3 Results

We assemble infant mortality and population by age and sex from reliable historical sources, comprising most of Europe, North America, Australia, New Zealand and South Africa.[12] Starting points for the data span a large range (from 1755 for Sweden to 1920 for Spain), reflecting the uneven development of vital statistics. To avoid the role of ultrasound in sex-ratio patterns, we end our series in the early 1970s (Campbell 2013). As discussed in the data appendix, we draw primarily on official sources for vital statistics and census data.

In Figure 1, we plot the infant mortality rate against the under-five sex ratio, $log(\frac{girls}{boys})$. For a given year, the sex ratio is paired to the mean of infant mortality in that year and the previous four. We also plot the regression lines of infant mortality on the sex ratio and the reverse. The two variables go hand in hand, with an R^2 of 77%. Populations with higher infant mortality have relatively more girls than those with lower infant mortality, and vice versa. Note the five clear outliers in the

[10] For example, see the work of Beltrán-Tapia and Raftakis (2020) on modern Greece.

[11] Recalling that a healthy modern population's under-five sex ratio will be about 5% more girls than boys, consider an observed ratio of 3% more boys than girls. With 10,000 children the 90% confidence interval is 6 percentage points.; with 50,000 children the 90% CI shrinks to about 3 percentage points (treating the sex-proportion as a binomial random variable – see e.g. Visaria 1967:133). In the smaller population, the female-biased sex ratio is inconclusive as evidence of poor maternal-infant health. In our own data, Iceland illustrates the limits of sex ratios in small populations. With a population of fewer than 10,000 children for most of the relevant period, the Icelandic data is extremely noisy, and we exclude it from our analysis. The issues of sample size, power and noise have previously been discussed in the context of sex ratios: e.g. Gelman and Weakliem (2009).

[12] Australia (1881), Austria (1869), Belgium (1846), Switzerland (1880), Denmark (1840), Spain (1920), Finland (1885), France (1901), Scotland (1861), England and Wales (1851), Iceland (1845), Italy (1871), the Netherlands (1859), Norway (1855), Germany (1880), South Africa (1918), and Sweden (1755). We also have data from states of the USA: Massachusetts (1860-1950) and various others for the period of 1910-1930. Both the population data and infant mortality rates come from officially published sources of the relevant government agencies.

upper-left quadrant: these are from Italy, 1871-1921. Here, infant mortality was quite high, but the infant-sex-ratio values was not skewed towards girls. We discuss their significance below.

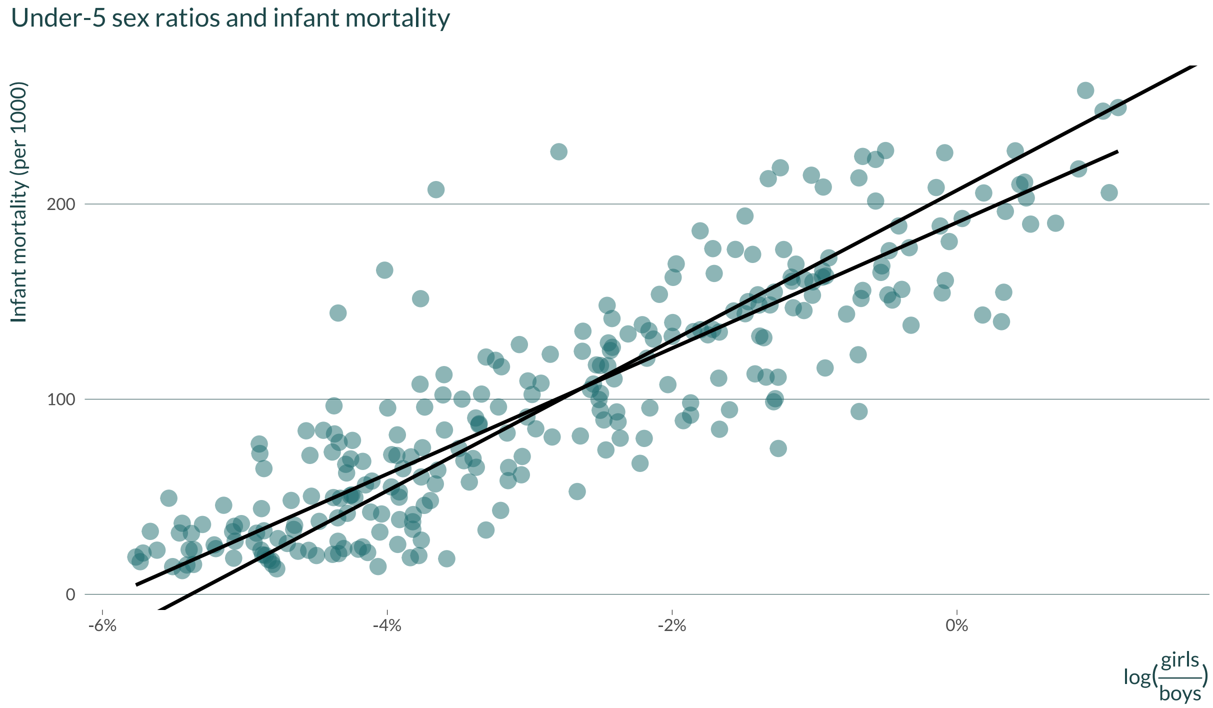

Figure 1: Infant mortality rate by log(girls/boys). The (median) regression of infant mortality on log(girls/boys) and the reverse are plotted as black lines. Data from Europe and settler colonies, up until 1970.

Within our sample, we can predict infant mortality with under-five sex ratios. In order to quantify the uncertainty in such predictions, we use out-of-sample prediction.[13] We drop the observations from one country, regress infant mortality on sex ratios in the remaining data,[14] and then predict the infant mortality for the dropped observations. We plot the errors of these predictions in Figure 2. These leave-one-out residuals are distributed around zero, with the 5th and 95th percentiles at 45 (overestimation) and 50 (underestimation) respectively.

The precision of our predictions depends on the level of infant mortality. In absolute terms, the prediction interval is broadly constant across the sample. However, in real-world applications, we are more interested in the *relative* error than the absolute. In terms of relative error, our predictions are much more precise for infant mortality levels For example, a prediction of 250 deaths per 1000, when the truth is 275, is fairly good. A prediction of 25 when the true value is 50 is not. We plot the proportional error, $log(\frac{Actual}{Residual})$ against infant mortality in Figure 3. In relative terms, our predictions are more precise for populations with greater infant mortality.

[13]For a classic reference, see Butler and Rothman (1980).

[14]We estimate a uni-variate regression using median regression, as it is less sensitive to outliers than least-squares (Kroenker and Hallock 2001). From a conceptual standpoint median regression is also desirable: we are interested in the *typical* infant mortality rate given a sex ratio, rather than the mean.

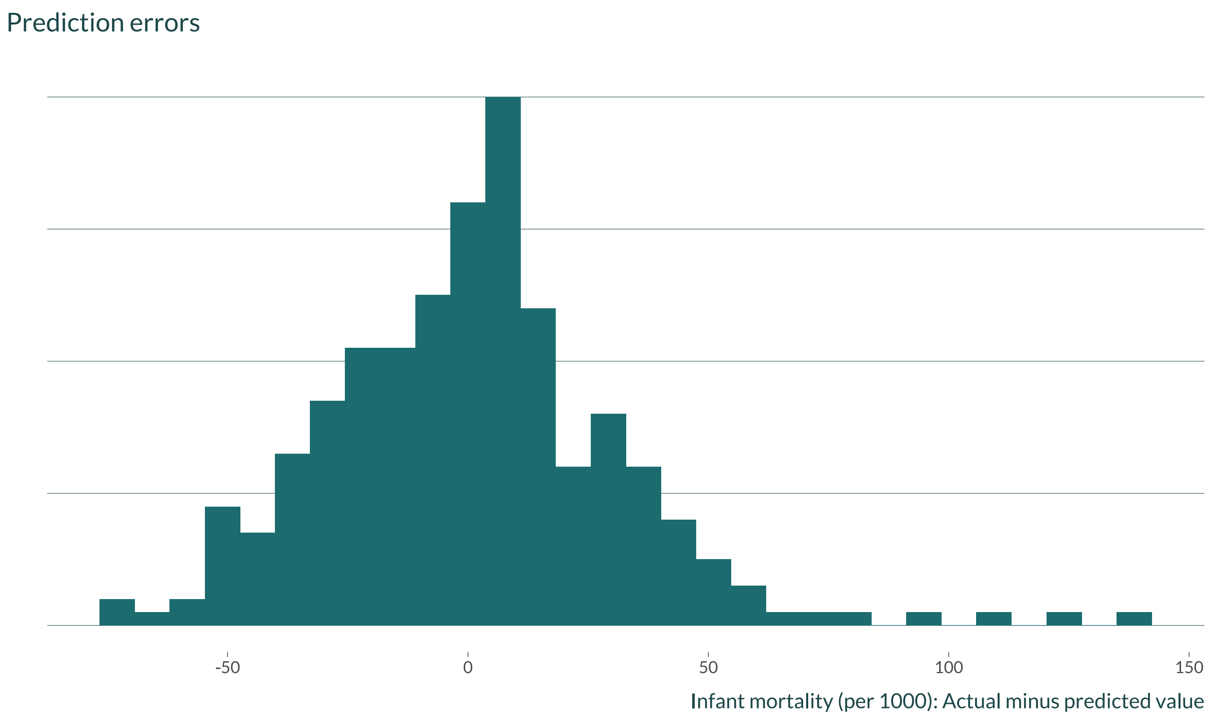

Figure 2: Prediction errors calculated by dropping one country and predicting it with median regression using the remaining values. Predicted minus actual infant mortality. Same data as in Figure 1.

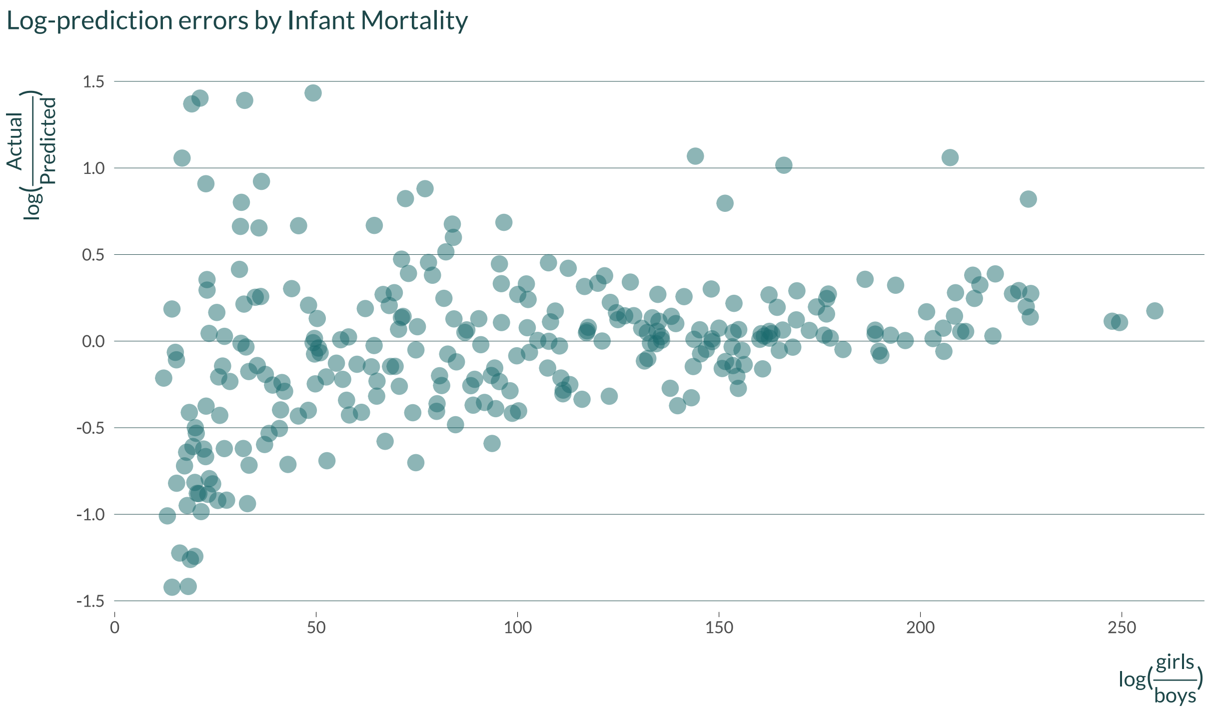

Figure 3: Relative prediction errors: log(Actual/Predicted). Same data as Figure 2, but plotted by actual infant mortality.

4 Discussion

A seen in the figures above, sex ratios can reveal broad patterns of infant mortality. Given population counts by age and sex, as in most published census reports, we can calculate the relative number of girls and boys under the age of five, and use this sex ratio to infer the level of infant mortality. For populations without vital statistics, where we have little to no direct evidence of infant mortality, sex ratios can shed crucial new light on infant health and maternal well-being.

As an illustration of the potential power of sex ratios, consider the non-white population of South Africa under the British Empire, for whom we have no direct evidence on infant mortality and maternal health. The sex ratios tell us that these women and children lived in *misery*. In 1911 South Africa, there were 7% more girls than boys under the age of five recorded in the black population.[15] Such a value is beyond any recorded value in European history. Even allowing for a two-point difference in the sex ratio at birth between white and black populations,[16] our model predicts an infant mortality of 350-400 deaths per thousand births.[17] At the same moment in time, the white population of South Africa enjoyed one of the lowest infant mortality rates in the world, around 90, and, accordingly, a sex ratio of 3% more boys than girls. These results may not

[15] This value refers to the "Bantu" population. The 1911 Census use race classifications of "Bantu", "European or White" and "Mixed and Coloured Races other than Bantu".

[16] See Morse and Luke (2021) for recent analysis. Note that the authors argue convincingly that this difference reflects maternal health and fetal loss. Therefore, our estimate of infant mortality for the black South African population is conservative with respect to population health as a whole, as we are ignoring this maternal health effect.

[17] We take the lines from Figure 1 and evaluate them for a value of 5% more girls than boys.

surprise, but they demonstrate the tremendous potential of sex ratios to reveal maternal-infant misery, particularly in oppressed populations for whom official health statistics have little concern. Populations across the British Empire and Latin America had censuses from the late 19th century onward, usually decades before the development of local vital statistics. Sex ratios can shed new light on these populations, for whom no adequate records of infant mortality exist.

However, predictions of infant mortality from sex ratios are, as shown by our data, coarse. To put it simply, within our sample, we can use sex ratios to distinguish between infant mortality levels of 100 and 150 deaths per thousand, but not between 120 and 140, nor between 10 and 30. This uncertainty comes largely from differences in maternal health and excess male mortality. Although they are closely connected, infant mortality can vary independent of maternal health, for example due to breastfeeding practices, leading to different sex ratios at birth between populations with the same infant mortality. Furthermore, excess male mortality is not fixed, and depends on the disease environment and socio-cultural matrix into which babies are born (Waldron 1998). Beyond these two factors, there is random variation in the sex ratio at birth. All of these factors introduce imprecision into the attempt to predict infant mortality from sex ratios.

We have purposefully presented the simplest possible model between sex ratios and infant mortality: a uni-variate regression. Even in this simplest of models, we can make meaningful inferences about infant mortality from sex-ratio data. However, there is a wide range of contextual information which could be used to improve the precision of these inferences. For example, within our sample, if we allow the intercept in our regression to vary by country, then our 90% prediction interval decreases by about a quarter. Given that cause-of-death patterns differ between rural and urban areas and across time, similar gains in precision might be possible from modeling the effects of these variable. We leave to future research a more precise understanding of the changing relationship of sex ratios and infant mortality.

Beyond any imprecision, sex ratios are, as discussed above, a fundamentally *one-sided* indicator of maternal-infant health, due to female infanticide and mortal neglect. This issue is clearly shown in our data. Every population with more girls than boys has infant mortality of at least 150 deaths per thousand births. However, the association of low infant mortality with male-skewed populations is less universal. In late-19th century Italy, infant mortality was greater than 200, yet there were still 4% more boys than girls. Moreover, as infant mortality decreased in Italy, we see very little change in the relative number of girls and boys. Italy highlights a limitation of our approach and the one-sidedness of sex ratios as an indicator, with male-skewed populations possibly reflecting sex discrimination rather than low mortality.[18] This one-sidedness is reflected in the long right tail of Figure 2. Constructing a 99% confidence interval, we obtain a lower bound (overestimation) of 50 points. Our upper bound (underestimation) is 70 points, and explodes to 100 points and beyond as

[18] Italy's outliers (1871-1921) in Figure 1 are somewhat less surprising in light of evidence that excess male infant mortality in the period was substantially lower in Italy than the rest of Europe, especially in the 19th and early 20th centuries (Drevenstedt et al. (2008); Figures 1 and 2); but the Italian case remains anomalous. The Italian sex-ratios are not as female-biased as we would predict from Italy's sex-specific IMRs. Italy could be a case where the female survival advantage may have been dampened by sex discrimination, a problem recently documented for modern Greece by Beltrán-Tapia and Raftakis (2020). Recall that female infanticide and mortal neglect would lead to relatively more males in the population, offsetting the effects infant mortality. However, we do not find evidence for this effect; unlike in the findings of Beltrán-Tapia and Raftakis, the male bias in Italian sex ratios does not increase with age. So further research is needed to explain Italy's male-biased sex ratios. The issue may simply be a data-quality issue, with either poor reporting of infant mortality or of population counts. Given the note of the Human Mortality Database that "the [Italian] data prior to 1906 should be used with extra caution due to problems of quality," we might have excluded Italy from our analysis. We did not, in order to err on the side of under-stating, rather than overstating, our results, given Italy's anomalous status in or analysis.

we reach the tail of the distribution. Any use of sex ratios should take into account the asymmetric nature of these prediction intervals.

Beyond estimating infant mortality, sex ratios may also prove useful in resolving uncertainty over the rate of stillbirths and neonatal mortality, which are often poorly recorded in both historical and contemporary populations (Dudfield (1912); Breschi et al. (2012); Gourdon and Rollet (2009); and Gourbin and Masuy-Stroobant (1995); as cited in Pozzi and Fariñas (2015)). Peri-natal mortality is heavily skewed toward males (Waldron (1998), 64), meaning that changes in the rate of stillbirth and neonatal mortality should be reflected in sex ratios. Sex ratio analysis could also complement the existing methods of indirect mortality estimation, e.g. the Brass method, which rely on survey questions about child survival.[19] Using sex ratios to estimate infant mortality relies on different assumptions than model-life-table methods, meaning they could be used to corroborate each other.

5 Conclusion

Child sex ratios can reveal *broad* patterns of infant mortality. Such revelations, though coarse, could be extremely useful in shedding light on hitherto under-studied times, places and peoples. Wherever we have censuses, but lack satisfactory vital statistics, sex ratios could provide valuable evidence on population health.

The concern of “missing girls” lurks, rightfully, behind any claims to use sex ratios as an indicator of infant mortality. Indeed, sex discrimination against girls will always bias estimates of infant mortality from sex ratios. However, as we have emphasized throughout, this bias goes in only one direction. Sex discrimination goes *against girls*, meaning that sex ratios provide a *strong* test of *high* infant mortality. When we see 7% more girls than boys, as in the black population of South Africa in 1911, we must conclude that maternal-infant conditions were awful. Sex ratios have the potential to reveal maternal-infant misery in a wide range of hitherto under-studied populations.

Throughout this essay, we focus on infant mortality, as our simple, bi-variate approach cannot say much about maternal health. In particular, our lack of attention to the sex ratio at birth, an important indicator of maternal health, keeps our predictions rather crude. Health during pregnancy is now understood to be, for the infant, a critical phase of physical, emotional and cognitive development.[20] Refining our analysis of sex ratios would enable inferences into *in-utero* health. We have demonstrated the coarse relationship between sex ratios and infant mortality. Future research should look to refine and apply it.

[19]See Hill (1991, 371-374) for a discussion of such methods.

[20]See Barker (1995) for a foundational work, and Almond and Currie (2011) for a review of this literature in the social sciences.

6 References

Alkema, Leontine, Fengqing Chao, Danzhen You, Jon Pedersen, and Cheryl C. Sawyer. 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.

Arbuthnot, John. 1710. "II. An Argument for Divine Providence, Taken from the Constant Regularity Observ'd in the Births of Both Sexes. By Dr. John Arbuthnott, Physitian in Ordinary to Her Majesty, and Fellow of the College of Physitians and the Royal Society." *Philosophical Transactions of the Royal Society of London* 27 (328): 186–90.

Beltrán Tapia, Francisco J., and Domingo Gallego-Martínez. 2017. "Where Are the Missing Girls? Gender Discrimination in 19th-Century Spain." *Explorations in Economic History* 66 (October): 117–26. https://doi.org/10.1016/j.eeh.2017.08.004.

———. 2020. "What Explains the Missing Girls in Nineteenth-Century Spain?" *The Economic History Review* 73 (1): 59–77. https://doi.org/10.1111/ehr.12772.

Beltrán Tapia, Francisco J., Siegfried Gruber, Bartosz Ogórek, and Mikolaj Szoltysek. 2021. "Inferring 'Missing Girls' from Child Sex Ratios in European Historical Census Data: A Conservative Approach."

Beltrán Tapia, Francisco J., and Michail Raftakis. 2021. "Sex Ratios and Gender Discrimination in Modern Greece." *Population Studies*, 1–18.

Breschi, Marco, Massimo Esposito, Stanislao Mazzoni, and Lucia Pozzi. 2012. "The Sardinian Experience of the Lowest Italian Infant Mortality at the Turn of the Twentieth Century. True or False Empirical Evidence?" In *Annales de Démographie Historique*, 63–94. 1. Belin.

Butler, Ronald, and Edward D. Rothman. 1980. "Predictive Intervals Based on Reuse of the Sample." *Journal of the American Statistical Association* 75 (372): 881–89.

Campbell, S. 2013. "A Short History of Sonography in Obstetrics and Gynaecology." *Facts, Views & Vision in ObGyn* 5 (3): 213–29.

Catalano, Ralph A. 2003. "Sex Ratios in the Two Germanies: A Test of the Economic Stress Hypothesis." *Human Reproduction* 18 (9): 1972–5.

Catalano, Ralph A., Katherine Saxton, Tim Bruckner, Sidra Goldman, and Elizabeth Anderson. 2009. "A Sex-Specific Test of Selection in Utero." *Journal of Theoretical Biology* 257 (3): 475–79.

Catalano, Ralph, Tim Bruckner, Elizabeth Anderson, and Jeffrey B. Gould. 2005. "Fetal Death Sex Ratios: A Test of the Economic Stress Hypothesis." *International Journal of Epidemiology* 34 (4): 944–48.

Catalano, Ralph, Tim Bruckner, Jeff Gould, Brenda Eskenazi, and Elizabeth Anderson. 2005. "Sex Ratios in California Following the Terrorist Attacks of September 11, 2001." *Human Reproduction* 20 (5): 1221–7.

Catalano, R., T. Yorifuji, and I. Kawachi. 2013. "Natural Selection in Utero: Evidence from the Great East Japan Earthquake." *American Journal of Human Biology* 25 (4): 555–59.

Coale, Ansley J. 1989. "Demographic Transition." In *Social Economics*, 16–23. Springer.

Coale, Ansley J., and Paul Demeny. 1983. *Regional Model Life Tables and Stable Populations*. Second. Studies in Population. Academic Press.

Drevenstedt, Greg L., Eileen M. Crimmins, Sarinnapha Vasunilashorn, and Caleb E. Finch. 2008. "The Rise and Fall of Excess Male Infant Mortality." *Proceedings of the National Academy of Sciences* 105 (13): 5016–21.

Dudfield, Reginald. 1912. "Still-Births in Relation to Infantile Mortality." *Journal of the Royal Statistical Society* 76 (1): 1–26.

Gelman, Andrew, and David Weakliem. 2009. "Of Beauty, Sex and Power: Too Little Attention Has Been Paid to the Statistical Challenges in Estimating Small Effects." *American Scientist* 97 (4): 310–16.

Gourbin, G., and Godelieve Masuy-Stroobant. 1995. "Registration of Vital Data: Are Live Births and Stillbirths Comparable All over Europe?" *Bulletin of the World Health Organization* 73 (4): 449.

Gourdon, Vincent, and Catherine Rollet. 2009. "Les Mort-Nés à Paris Au XIXe Siècle: Enjeux Sociaux, Juridiques et Médicaux d'une Catégorie Statistique." *Population* 64 (4): 687–722.

Grech, Victor. 2020. "The Sex Ratio at Birth Aspects." *Early Human Development* 140 (January): 104857.

Guimbeau, Amanda, Nidhiya Menon, and Aldo Musacchio. 2020. "The Brazilian Bombshell? The Long-Term Impact of the 1918 Influenza Pandemic the South American Way." National Bureau of Economic Research.

Hill, Kenneth. 1991. "Approaches to the Measurement of Childhood Mortality: A Comparative Review." *Population Index* 57 (3): 368–82. https://doi.org/10.2307/3643873.

Hollingshaus, Mike, Rebecca Utz, Ryan Schacht, and Ken R. Smith. 2019. "Sex Ratios and Life Tables: Historical Demography of the Age at Which Women Outnumber Men in Seven Countries, 18502016." *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 52 (4): 244–53.

James, William H., and Victor Grech. 2017. "A Review of the Established and Suspected Causes of Variations in Human Sex Ratio at Birth." *Early Human Development* 109: 50–56.

Klasen, Stephan. 1994. "'Missing Women' Reconsidered." *World Development* 22 (7): 1061–71.

Klasen, Stephan, and Claudia Wink. 2002. "A Turning Point in Gender Bias in Mortality? An Update on the Number of Missing Women." *Population and Development Review* 28 (2): 285–312.

Maconochie, Noreen, and Eve Roman. 1997. "Sex Ratios: Are There Natural Variations Within the Human Population?" *BJOG: An International Journal of Obstetrics & Gynaecology* 104 (9): 1050–3.

McDevitt-Irwin, Jesse, and Irwin, James R. 2021. "Infant Sex Ratios and Infant Mortality in North America: 1850-1920."

Mikkelsen, Lene, David E. Phillips, Carla AbouZahr, Philip W. Setel, Don de Savigny, Rafael Lozano, and Alan D. Lopez. 2015. "A Global Assessment of Civil Registration and Vital Statistics Systems: Monitoring Data Quality and Progress." *The Lancet* 386 (10001): 1395–1406. https://doi.org/10.1016/S0140-6736(15)60171-4.

Mitchell, Brian. 1998. *International Historical Statistics: Europe 1750-1993*. Springer.

OECD, and World Health Organization. 2020. *Infant Mortality.* Doi:https://doi.org/10.1787/f7b3e551-en.

Pozzi, Lucia, and Diego Ramiro Fariñas. 2015. "Infant and Child Mortality in the Past." In *Annales de Démographie Historique*, 55–75. 1. Belin.

Reidpath, D. D., and P. Allotey. 2003. "Infant Mortality Rate as an Indicator of Population Health." *Journal of Epidemiology & Community Health* 57 (5): 344–46. https://doi.org/10.1136/jech.57.5.344.

Sanders, Nicholas J., and Charles Stoecker. 2015. "Where Have All the Young Men Gone? Using Sex Ratios to Measure Fetal Death Rates." *Journal of Health Economics* 41 (May): 30–45.

Tavassoli, Nahid, Hamid Noghanibehambari, Farzaneh Noghani, and Mostafa Toranji. 2021. "Upswing in Industrial Activity and Infant Mortality During Late 19th Century US." *arXiv Preprint arXiv:2101.02590*. http://arxiv.org/abs/2101.02590.

Valente, Christine. 2015. "Civil Conflict, Gender-Specific Fetal Loss, and Selection: A New Test of the TriversWillard Hypothesis." *Journal of Health Economics* 39: 31–50.

Waldron, Ingrid. 1998. "Sex Differences in Infant and Early Childhood Mortality: Major Causes of Death and Possible Biological Causes." In *Too Young to Die: Genes or Gender?*, 64–83. ST/ESA/SER.A 155. New York: United Nations.

7 Appendix: Data Sources

Appendix: Data Sources

The Human Mortality Database (HMD) provides appropriate data for infant mortality rates and sex ratios for many nations. We expand our geographic scope by drawing on vital statistics and census data from various official sources. In many cases, the data are available from International Historical Statistics (IHS 2013).

The HMD data fall into two categories, census-based or register-based. Data for Finland and Sweden are register-based throughout; there is a shift from census- to register-basis in the data for the Netherlands, Denmark and Iceland after 1870, 1900 and 1950 respectively. In general, we rely on only official estimates of population (i.e. coming from the relevant state authority, as opposed to estimates and reconstructions from demographic research). In general, we calculate sex-ratio values from official population counts by sex and age; for registry-based sex-ratios, we take value at five-year intervals. For infant mortality rates, we rely on official vital statistics except when demographic scholarship offers better estimates [LIST 'em here, once appendix is about complete]. Specific sources and methods by country follow.

Australia (1871-1971)

Under-5 populations by sex. 1871, 1881, 1891: census values reported in J. C. Caldwell, "Population", Chapter 2 of Vamplew (1987); 1901, 1911, and from 1921 to 1971 at five-year intervals to 1971: Australian Historical Population Statistics, 3105.0.65.001, Table 2.1 "Population, age and sex, Australia, 30 June, 1901 onwards" (released 2019 April 18).

Infant mortality rates. 1870-1901: Series MFM 154 in Peter McDonald et al., "Marriage, Fertility, and Mortality," Chapter 3 of Vamplew (1987). 1901-1971: Australian Historical Population Statistics, 3105.0.65.001, Table 5.4 "Infant mortality rates … 1901 onwards" (released 2019 April 18).

Austria (1869-1971)

Under-5 populations by sex. 1869 & 1910: IHS-Europe, A2 p. 1 (Austrian provinces of the Hapsburg Empire). The values here are rounded to the nearest thousand (we would prefer unrounded data, but were unable to locate the data in official sources). 1880, 1890, 1900: Österreichisches statistisches Handbuch … 1886, p. 3; 1893, p. 6; 1909, p. 7. 1923, 1934, 1951, 1961 Statistik Austria, Statistisches Jahrbuch 2010, p. 45 (2.08 Bevölkerung 1869 bis 2001 nach fünfjährigen Altersgruppen und Geschlecht (Population 1869 to 2001 by five-year age groups and sex).

Infant Mortality Rates. 1865-1971: IHS-Europe, A3 pp. 3, 6, 9.

For the years 1865-1910, Austria refers to Austria-Hungary (as in IHS-Europe); for later years Austria refers to the Republic of Austria (the area of which in 1910 embraced less than 1/4 of the population of Austria-Hungary).

Germany (1876-1933)

Under-5 populations by sex. 1880: Statistisches Jahrbuch 1883, p. 10; 1890: Statistisches Jahrbuch 1896, p. 5; 1900: Statistisches Jahrbuch 1903, p.6; 1910: Statistisches Jahrbuch 1919, pp. 6-7; 1925: Statistisches Jahrbuch 1929, p. 14; 1933: Statistisches Jahrbuch 1939, p. 14. IHS reports these data rounded to the nearest thousand (Series A2, p. 3454). For 1933, we use the census values, dated 16 June 1933; the IHS value for 1933 matches the estimates for Dec. 31, 1933, found in Statistisches Jahrbuch 1936, p. 12.

Infant Mortality Rates. International Historical Statistics (2013), Series A7, pp. 3577, 3580.

West Germany (1951-1970) HMD provides under-5 populations by sex for 1956, 1961, and 1970, and infant mortality rates for 1951-56, 1957-61, 1966-70)

East Germany (1959-1971) HMD provides under-5 populations by sex for 1964 and 1971, and infant mortality rates for 1959-64 and 1966-71.

Italy (1881-1981),

Under-5 populations by sex are taken from Cox, P. R. (1958). The aging of populations and its economic and social implications. UN. Pages 142-143.

Infant mortality is taken from ISTAT.

New Zealand (1864-1971)

Under-5 populations by sex 1864, 1867, 1871, 1874, 1878, 1881: Results of a Census of the Colony of New Zealand ... 1874. Chapter 28, Table 1 ("Showing the Increases of Persons ... at different Ages ... the various censuses ... 1864 ... 1881). 1886, 1891, 1896, 1901, 1906, 1911, 1916: Results of a census of the Dominion of New Zealand ... 1916, Part II Ages, page 1. 1936, 1945, 1951, 1956, 1961, 1966, 1971: Stats NZ Spreadsheet A1.6 (citing Bloomfied (1984), "Census Reports: Table II.6. Age Groups ... 1874-1976"). Available in discontinued series of Stat NZ Infoshare, Census of Population and Dwellings, Total Pop - Sex by Single Years of Age. 1874-1945 exclude Maori population.

Norway (1886-1970)

IMR: annual values 1886-1970, as published in IHS, corroborated with Statistics Norway data. Reliable IMR data for Norway start in 1876. 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 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.1 Although early some early publications from Statistics Norway report IMR data from before 1876, the 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.

U5 census populations by sex: 1890, 1900, 1910, 1920, 1930, 1946, 1950, 1960, 1970. These data can be found in the HMD for Norway. The data for 1900 and before are published in Statistics Norway (1910, pp. 32-34). The data for later years are in the respective census volumes (available online at Statistics Norway). Note that U5 population data are also available for 1800, 1845, 1855, 1865, 1875; however, as discussed above, reliable infant mortality data starts in 1876.

Statistics Norway (1910), Norges folkemængde fordelt paa de enkelte aldersaar 1846-1901. Norges Officielle Statistik V. 113. Kristiania (Oslo), H. Aschehoug & Co.

Backer, Julie 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 (accessed 20211005).

South Africa (1918, 1921) IMR 1910-1921: International Historical Statistics 1750-2010: Africa, Asia and Oceania (Volume 1), Series A7 Africa, p. 1 (DOI 10.1057/9781137305688). U5 populations by sex 1911, 1918: Official Yearbook of the Union ... 1922, pp. 158-59; 1921: Official Yearbook of the Union ... 1925, p. 868.

Various Links for above sources

Australia u5 https://www.abs.gov.au/statistics/people/population/historical-population/2016/3105065001ds0002_2019.xls accessed 2021 June 27. https://socialsciences.org.au/library/historical-statistics-chapter-2/

Australia imr https://www.abs.gov.au/statistics/people/population/historical-population/2016/3105065001ds0005_2019.xls accessed 2021 June 27 https://socialsciences.org.au/library/historical-statistics-chapter-3/

HMD:

Australia, (1954-1971); Belgium, (1846-1970); Denmark, (1840-1970); England and Wales, (1851-1970); Finland (1885-1970); France (1901-1968); Netherlands (1859-1970); New Zealand (1951-1971); Norway (1855-1970); Scotland (1861-1970); Spain (1920-1970); Sweden (1755-1970); USA (1940-1970).