



Revising Infant Mortality Rates for the Early Twentieth Century United States

Katherine Eriksson^{1,2,3} · Gregory T. Niemesh^{3,4} · Melissa Thomasson^{3,4}

Published online: 2 November 2018
© Population Association of America 2018

Abstract

Accurate vital statistics are required to understand the evolution of racial disparities in infant health and the causes of rapid secular decline in infant mortality during the early twentieth century. Unfortunately, U.S. infant mortality rates prior to 1950 suffer from an upward bias stemming from a severe underregistration of births. At one extreme, African American births in southern states went unregistered at the rate of 15 % to 25 %. In this study, we construct improved estimates of births and infant mortality in the United States for 1915–1940 using recently released complete count decennial census microdata combined with the counts of infant deaths from published sources. We check the veracity of our estimates with a major birth registration study completed in conjunction with the 1940 decennial census and find that the largest adjustments occur in states with less-complete birth registration systems. An additional advantage of our census-based estimation method is the extension backward of the birth and infant mortality series for years prior to published estimates of registered births, enabling previously impossible comparisons and estimations. Finally, we show that underregistration can bias effect estimates even in a panel setting with specifications that include location fixed effects and place-specific linear time trends.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s13524-018-0723-2>) contains supplementary material, which is available to authorized users.

✉ Gregory T. Niemesh
niemesgt@miamioh.edu

Katherine Eriksson
kaeriksson@ucdavis.edu

Melissa Thomasson
thomasma@miamioh.edu

¹ Department of Economics, National Bureau of Economic Research, 1144 SSH, Davis, CA 95616, USA

² LEAP, Department of Economics, Stellenbosch University, Matieland 7602, South Africa

³ National Bureau of Economic Research, Cambridge, MA 02138, USA

⁴ Department of Economics, Miami University, 800 East High Street, Oxford, OH 45056, USA

Keywords Infant mortality · Birth registration · Vital statistics · Underregistration

Introduction

Vital statistics form the foundation of our understanding of health trends for the United States and are regarded as indispensable when targeting effective public health programs and evaluating interventions. As early as the late nineteenth century, public health officials recognized the importance of statistics coming from the vital registration system as an important resource in the fight against infectious disease (Cassedy 1965). For modern researchers in economics, demography, and public health, vital statistics from the early twentieth century provide a rich data source to understand trends in mortality and longevity as well as socioeconomic correlates with health, and to estimate causal impacts of health interventions. A large strand of research has contributed to understanding overall trends in life expectancy and infant mortality in the twentieth century. In general, U.S. infant mortality followed a strong, downward trend in the twentieth century, and the racial gap in white–nonwhite infant mortality rates has persisted (although it has varied in magnitude). Researchers have focused on documenting and explaining these trends as well as on measuring the impact of public health interventions on both the level movements in these trends and the racial gaps within them.¹

Unfortunately, estimates of live births, infant mortality rates (IMRs), and maternal mortality rates prior to 1950 suffer from an upward bias stemming from a severe underregistration of births. Not only are rates incorrect, but the measurement error varies over races and locations in ways that are potentially correlated with variables of interest. Using newly released census microdata, we can now construct improved estimates of live births, infant mortality, and maternal mortality for the United States. In this study, we present our methods and estimates and demonstrate the potential implications of the revisions on our understanding of trends and racial differences in infant mortality.

To obtain the new estimates, we revise the number of births while leaving the published counts of infant deaths unchanged. Thus, differences between published and revised rates arise from using different estimates of live births. In addition to improving on published estimates, our method enables us to extend the existing series backward in time. Although current state-level infant mortality rates begin only after a state enters the birth registration area (BRA), we are now able to construct a series based on when a state entered the death registration area (DRA), which generally occurred prior to a state's entrance into the BRA.² As a result, our series allows for previously impossible comparisons of fertility and infant mortality across groups and analyses of earlier interventions.

We focus on infant mortality and compare our revised measure with existing series to demonstrate the importance of using the new estimates. Infant mortality rates (IMR) are computed by dividing registered deaths of infants by the number of registered live births occurring during a calendar year. Bias can enter the calculation through an

¹ Examples of recent studies that used underregistration-biased births estimates include Bhalotra and Venkataramani (2011), Clay et al. (2013), Collins and Thomasson (2004), Cutler and Miller (2005), Eriksson and Niemesh (2016), Hansen (2014), Jayachandran et al. (2010), Moehling and Thomasson (2014), and Thomasson and Treber (2008).

² States entered the DRA as early as 1880, but the BRA did not begin until 1915. Table A3 in the online appendix lists the entry dates for each state into the BRA and the DRA.

incorrect estimate of infant deaths (the numerator) or an incorrect estimate of births (the denominator). Contemporary evidence suggests that severe underregistration of births biased IMR estimates at least until 1940, with the bias varying by region and race (Grove 1943). Bias in the numerator from unregistered deaths was believed to be a minor issue. Thus, IMR estimates using registered events will vary inversely with the completeness of birth registration.³

To account for this severe underregistration of births prior to 1940, we construct revised annual, two-year, and five-year adjusted estimates of births, IMRs, and maternal mortality rates by state and race as well as at the national level.⁴ To create the new estimates, we calculate births as equal to the sum of the enumeration of live children in the census, the number of infant deaths, and the number of noninfant deaths. We begin with the enumeration of children in the decennial census for each *state of birth* \times *year of birth* \times *race* cell, using newly released complete census microdata for the 1920, 1930, and 1940 decennial censuses from IPUMS (Ruggles et al. 2017), which we then adjust by estimates of underenumeration (Hacker 2013; Land et al. 1984; Preston et al. 2003). Infant deaths are allocated to the state and year of occurrence. Deaths of children after infancy but prior to the subsequent decennial census enumeration are allocated to year and state of birth.

At the national level, the revised estimates suggest a lower black IMR relative to those in published sources, with larger differences prior to 1925: 12.6 percentage points in 1915 versus 18.1 in the published data. The lower initial level in 1915 also implies slower *progress* in black infant health. The IMR declined by 10.8 percentage points between 1915 and 1940 in the published data but only by 6.8 percentage points in the revised estimates. Because underregistration of births was not as severe for whites, revised estimates of the native-born white IMR do not deviate from published estimates as much. The largest difference occurs in 1915 and is only 1.0 percentage point. The core finding of lower IMR estimates stem from two factors: (1) accounting for the severe underregistration of black births, and (2) the extension of the IMR series to include primarily rural states (the South), which experienced lower IMRs than the northern states included in the published series.

As we show in this article, the large variation across states in the quality of birth registration data leads to significant revisions of the relative rankings of states based on infant mortality, which has important implications for regional differences and subsequent convergence. The South initially had a mortality advantage over the North for black infants, but rates converged as the urban penalty gradually declined over the course of the early twentieth century. When the revised estimates are used instead of the published rates, the southern mortality advantage widens as the adjustment method primarily lowers the black IMR in the South. Second, starting from a lower initial IMR

³ Researchers in the early twentieth century understood the biases present in infant mortality rates. For example, in 1934, a former president of the Population Association America wrote, "If birth registration is equally deficient in various states, only absolute values for birth rates and infant mortality are affected. However, if there are large differences in completeness between states, the comparative standing of states in these respects will vary correspondingly when they are ranked on an adjusted instead of an unadjusted basis" (Whelpton 1934:125).

⁴ City-level infant mortality estimates are also biased from birth underregistration. However, we are loath to use our procedure to adjust rates at the city level. The census records only the state of birth, not the city. Any allocation of census enumerations to a city of birth would be plagued by bias from cross-city migration. Because state of birth is known, this migration bias does not affect our state-based revised rates.

in the South implies a faster convergence rate between the regions. Finally, the level shift downward in the southern IMR delays the North overtaking South until the late 1930s, if at all before 1940.

Development of the Birth Registration Area and Evidence of Completeness

The Massachusetts legislature adopted the first registration law for vital events in 1842, with six other states enacting similar legislation by 1851. These early systems operated in only a few localities and suffered from lax enforcement (Lunde 1980). Despite the known flaws in the system, public health professionals realized the importance of vital statistics reporting in their efforts to combat and eradicate infectious disease in the latter half of the nineteenth century. The federalism of the time slowed the growth of the registration system, imposing a piecemeal state-by-state approach that eventually created nationally representative statistics.⁵ The death registration area (DRA) began in 1880 with two states, the District of Columbia, and several large cities. In 1900, the Census Bureau established a national DRA that initially included 10 states, mainly from the Northeast and Midwest. The DRA was completed in 1933 with the entrance of Texas.

It took longer to establish the birth registration area (BRA). Public health officials viewed mortality data as being more helpful for preventive medicine than birth data, and registrars believed enforcement of birth registration to be more difficult than for deaths (Cassedy 1965). However, after starting in 1915 with 10 states and the District of Columbia, the BRA was completed relatively quickly over a period of 18 years. Again, states in the Northeast, Middle Atlantic, and Midwest joined first, with most of the remainder of the country entering in the 1920s. Southern states lagged the others, and the BRA was not completed until 1933 with the entrance of Texas. A list of entrance dates for each state can be found in the online appendix, Table A3.

States seeking entrance to the BRA had to overcome two hurdles. First, the state legislature needed to enact and enforce registration laws in a manner deemed sufficient by the Census Bureau. The more difficult second hurdle was to show evidence that registrations were at least 90 % complete (Lunde 1980; Moriyama 1990). All tests of registration completeness proceeded by first obtaining a list of children born during a fixed period and then determining whether birth certificates had been filed for those children. The Census Bureau used various methods to obtain the list of names over the course of the early twentieth century. At the advent of the BRA, the test was conducted under the direction of the Census Bureau and consisted of comparing birth registrations against collected lists of births from postmasters, newspapers, death registers, and church records. Contemporaries acknowledged early on that the tests used to enter the BRA were woefully inadequate (Whelpton 1934). Cressy Wilbur, Chief Statistician for Vital Statistics of the United States for 1906–1914, believed the use of lists of births

⁵ Vital registration systems were and remain the responsibility of several states. The federal government's role is limited to the promotion of state registration systems and to working with the states to produce national-level statistics. An act of Congress in 1902 put the national system on a firm footing by making the Census Bureau a permanent agency and providing the authority to collect information on births and deaths.

collected by postmasters to be a highly biased sample for a test (Wilbur 1916). Deacon (1937) related the story of how after finding a 100 % registration rate from names provided by a postmaster, he came to find that the postmaster received the list directly from the local registrar. Later evidence showed the sources used to create the list of children—death registrations, hospital births, and newspaper announcements—were likely a highly selected sample of births: children born to urban, educated, and wealthier parents were more likely to appear in these sources and were also more likely than the population to register a birth (Moriyama 1990). The selected sample caused the tests to overestimate the completeness of the registration system. Nevertheless, entrance to the BRA was granted after a positive test result.

In the mid-1920s, the Census Bureau switched to a testing procedure based on postal cards, which were sent out in mass mailings to every known household. Residents were asked to list the occurrence of any deaths or births that occurred during the prior 12 months, with returned cards checked against birth registers. Although believed to be an improvement over collected lists, the postal card method suffered from its own biases. Errors entered the lists from memory lapses inherent in any recall method. More importantly, households with unregistered events were less likely to return the cards, as were households with low education and incomes (Moriyama 1990). Tests in Georgia and Maryland in 1934 used the postal card method and compared it with the results from a canvas of enumerators. The tests revealed that (1) registrations were more complete for white, urban households with higher incomes and education, as well as for hospital births; (2) the postal test card method led to overstatements of completeness because mail carriers were more likely to deliver the cards to households receiving other mail, which were those with higher income and education levels; and (3) households with higher incomes and greater levels of education were more likely to return the cards (Hedrich et al. 1939). Postal card tests, generally thought of as an improved method of testing for entrance into the BRA, grossly overstated the completeness of birth registrations. By the 1930s, officials at the Census Bureau recognized the need for a nationwide test built on proper sampling procedures.

In addition to biased samples, public health officials worried about the subsequent quality of registrations after the entrance test (Wilbur 1916). The early policy called for periodic retests using the collected lists methodology to ensure the 90 % cutoff continued to be met (Davis 1925). However, retests were infrequent—once in 16 years in the case of Michigan—and poor results rarely led to a state exiting the BRA (Deacon 1937). Despite evidence that a number of states were well under the 90 % cutoff, only two states were ever expelled: Rhode Island in 1919 (reentering in 1921) and South Carolina in 1925 (reentering in 1928) (Wilcox 1933). By the mid-1930s, the Census Bureau's policy was that retests were for the sole purpose of helping to improve the registration systems of underperforming states, not to threaten removal from the BRA (Lenhart 1943).

1940 Test of Birth Registration Completeness

The opportunity arose with the 1940 decennial census to develop a nationwide test that would greatly improve knowledge about the accuracy of the birth registration system (Grove 1943). Officials believed that census enumerators could provide a more representative list of children born during a sample period than previous methods.

Enumerators completed a special infant card for any child born during the four months prior to the census date.⁶ The Census Bureau then matched each infant card and recorded death of an infant to birth certificates filed in state registrar offices. The completeness of registrations was then estimated as the proportion of infant cards and registered deaths for which a birth certificate had been filed.

For the nation as a whole, 92.5 % of births were found to be registered, but large differences existed between races (94.0 % completeness for whites vs. 82.0 % for blacks), cities and rural areas (96.9 % completeness in cities with more than 10,000 in population vs. 88.0 % in small cities and rural areas combined), and hospital versus home births (98.5 % in hospitals vs. 86.1 % outside hospitals) (Grove 1943; Moriyama 1946).⁷ The card test suggested that underregistration in some states was quite severe in total, and particularly poor for blacks, including an upward bias in reported infant mortality for the South. For example, only 77.6 % of births were registered in South Carolina versus 99.4 % in Connecticut. The geographic variation in birth registration completeness for all races combined is shown in Fig. 1. In general, the South had the highest level of underregistration, with regional differences attributed to differences in urbanization and rates of hospital births (Moriyama 1946).⁸

Improvement in Birth Registration Completeness

Continued urbanization and increases in the proportion of births delivered in hospitals eventually reduced the number of unregistered births. Additionally, the value to the individual of holding a birth certificate rose because proof of age was increasingly required for receipt of government benefits, school attendance, and other privileges, such as a driver's license. Subsequent tests for registration completeness were conducted at a national scale in conjunction with the 1950 census and in the late 1960s using household surveys, such as the Current Population Survey and the Health Information Survey (Shapiro and Schachter 1952; U.S. Census Bureau 1973). The results of the tests between 1940 and 1950 suggest large improvements in birth registration at the national level: from 92.5 % in 1940 to 97.8 % in 1950. The national average, however, belied large regional differences for minorities.⁹ Completeness for southern nonwhites increased only to 92 % by 1950. For states in the Mountain census region with large Native American populations, the nonwhite completeness rate lagged at 78 %.¹⁰ By at least 1968, after the integration of hospitals in the South, the proportion

⁶ The recall period for the infant postal cards was limited to four months to reduce bias from memory lapses.

⁷ Table A1 in the online appendix reports results from the test by region and whether delivery occurred in a hospital.

⁸ A subsequent test conducted in 1950 showed major improvements over the decade, with 97.8 % registered for the nation as a whole, although some states lagged behind (Shapiro and Schachter 1952). The likely explanation for this rapid improvement is that registration completeness is highly correlated with the percentage of births delivered in a hospital. Completeness eventually reached close to 100 % by the mid- to late 1960s, when hospital deliveries approached 100 % of all births after the integration of hospitals in the South (U.S. Census Bureau 1973).

⁹ Figure A1 in the online appendix plots the proportion of all births registered from the 1950 test against that from the 1940 test. The figure clearly shows that all states increased the quality of their published birth data. However, some improved more than others. One plausible reason for this variability is that low rates of out-of-hospital births persisted for nonwhites in the South and West.

¹⁰ The Mountain census region includes the following states: Arizona, New Mexico, Nevada, Utah, Colorado, Wyoming, Idaho, and Montana.

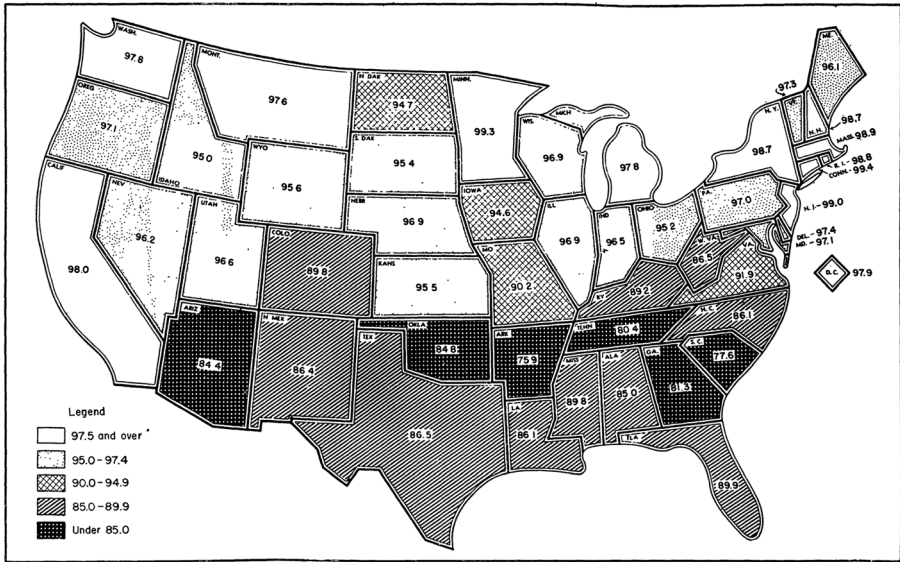


Fig. 1 Percentage completeness of birth registration by state, December 1, 1939–March 31, 1940. *Source:* Reprinted from Shapiro (1950)

of births delivered in a hospital converged to almost 99 % nationwide for all races combined, and the birth registration system covered close to the entire universe of all births: 99.4 % for whites and 98.0 % for nonwhites.

Why Does Underregistration Matter?

In general, IMR differences and treatment effect estimates will be biased when underregistration is correlated with the intervention or group attribute. Answering the question of why underregistration matters is simplified if we consider three scenarios. First, sometimes researchers would like to know the true IMR for a given place and time without making any comparisons. In this simple scenario, *any* underregistration of births will bias the estimate of IMR.

Second, researchers frequently make comparisons across locations, groups, or time. IMR differences arising from a cross-sectional comparison partially reduce the bias as long as the extent of underregistration remains constant across the groups being compared. However, underregistration appears to vary in important ways across groups and locations (e.g., higher bias in the IMR for blacks and in southern states). Later, we provide two applications of cross-sectional comparisons in which this bias can dramatically change results. The first shows the impact on the pace and timing of regional convergence in the North-South difference in black IMRs from Eriksson and Niemesh (2016). We also revisit Collins and Thomasson (2004) to conduct an Oaxaca-Blinder decomposition of the national black–white IMR gap using measures of socioeconomic status as explanatory variables.

In the third scenario, researchers use panel data with observations for each location taken over multiple points in time. Location fixed effects and location-specific trends

potentially account for any mismeasurement of IMR from differential completeness of the birth registration system. To explore this possibility, we estimate a series of regressions to determine the ability of state fixed effects and state-specific linear time trends to explain the gap between the published and adjusted infant mortality estimates. We use three measures for the gap that correspond to three specifications for IMR commonly used in the literature: the difference ($IMR^{PUB} - IMR^{ADJ}$), the ratio (IMR^{PUB} / IMR^{ADJ}), and the natural log of the ratio ($\ln(IMR^{PUB} / IMR^{ADJ})$). Additionally, we split the sample into black, native-born white, and total.¹¹ Regardless of how the gap is specified or on which sample the regression is run, between 16 % and 26 % of the variation in the gap remains when state fixed effects are included. After state-specific linear time trends are included, the remaining variation in the gap ranges from 7 % to 12 % across all samples. The standard deviation of the residuals from specifications that include linear trends ranges between 3 % and 7 % of the level of IMR, depending on the sample and how the gap is measured. In summary, the use of a panel setting to difference out unobservable characteristics or allowing for differential trends in unobservable characteristics does not fully remove the bias from causal estimates in the presence of birth underregistration.

Adjusting Infant Mortality Rates

In this section, we outline the method and data sources used to revise IMRs and birth estimates to account for the underregistration of births. We then graphically present the adjusted rates for different subcategories and discuss differences with the published vital statistics. The results of the exercise consist of a set of tables of IMRs by subcategory for one-, two-, and five-year averages for use by researchers. A full set of machine-readable tables is published as Eriksson et al. (2018).¹² In the end, we provide two additional estimates of infant mortality in addition to those in the published vital statistics: one that uses the census-based adjustment method, and a second series in which births are scaled by the extent of underregistration in the 1940 test in Grove (1943).

Published IMRs are constructed from registered deaths before age 1 and registered births using the following formula:

$$IMR_{s,r,t}^{PUBLISHED} = \frac{\text{Published Deaths}_{s,r,t}}{\text{Published Births}_{s,r,t}},$$

where s denotes state of occurrence, r denotes race, and t denotes calendar year. IMR is often reported as deaths per thousand live births, but we report in percentage points for simplicity. We know from contemporary evidence that ($\text{Published Births}_{s,r,t}$) is biased downward in a way that leads to an upward bias in IMRs for blacks and southern states.

To revise these rates, we rely on newly available complete count census microdata for 1920, 1930, and 1940 as the main source of information on the

¹¹ Table A4 in the online appendix reports results from this exercise.

¹² The white rates are for native-born whites only, and the nonwhite rates are for blacks. Foreign-born whites and nonblack nonwhites are excluded from the current estimates.

number of children who remained alive, published age-specific deaths for each state and race to account for noninfant deaths, and deaths of infants from published sources. In all estimates, the numerator of the IMR calculation—infant deaths—is held constant and comes from the published counts of registered deaths. Thus, any differences from the published mortality rates arise from an alternate estimate of live births. Our method provides a distinct improvement for understanding infant mortality during the early twentieth century United States.

Our adjusted IMRs can be expressed as follows:

$$IMR_{s,r,t}^{ADJUSTED} = \frac{\text{Published Deaths}_{s,r,t}}{\text{Adjusted Births}_{s,r,t}},$$

so that any difference with the published rates are entirely driven by differences in birth estimates. Our adjustment uses the complete count census data sets from IPUMS to estimate the number of live children by race, birth state, and birth year (Ruggles et al. 2017).¹³ Census counts suffer from underenumeration, which we adjust by the estimates of underenumeration contained in Land et al. (1984), Preston et al. (2003), and Hacker (2013). To this, we then add the number of infant and noninfant deaths during intervening years between the birth year and the census year, both of which come from published tables. The data appendix contains a lengthy discussion of the data sources used and additional detail on the construction of estimates.

Figure 2 plots the bias in the published rates (published minus adjusted rates) against the extent of underregistration in the 1940 test from Grove (1943). States with higher levels of underregistration do in fact see larger reductions in IMR using our census-based method, just as we would expect. Over time, the size of the adjustment falls, and the relationship between extent underregistration and the bias in published rates weakens. We interpret this set of facts as evidence of gradual improvement in the birth registration system over time.

One concern in our estimates is the potential for children to migrate outside their state of birth. Our estimate of live children includes those born in state *s* regardless of the state of residence at the time of the census. The potential migration of children outside their state of birth does not bias our estimates of births downward as long as they remain alive until the next census. The problem arises when children die between censuses. In the absence of a nationwide death index, we do not have complete information on children who died outside their state of birth. Bias enters our estimate when states had differential net migration rates or differential mortality rates. In all cases, infant deaths are allocated to the state of occurrence regardless of the child's birth state because we have no

¹³ The 1930 and 1940 censuses were conducted on April 1, so the census counts do not align with the vital statistics data reported by calendar year. Our estimates of births for a given year are underestimated when actual births are declining and are overestimated when actual births are increasing. In practice, reallocating "births" so that the census counts follow the calendar year does not meaningfully change IMR estimates. Moreover, year and state fixed effects in a panel setting account for any of the differences. An explanation of the allocation procedure and full set of results are available upon request from the authors. The 1920 census questions referred to January 1, and thus the 1920 census counts do not suffer from this problem.

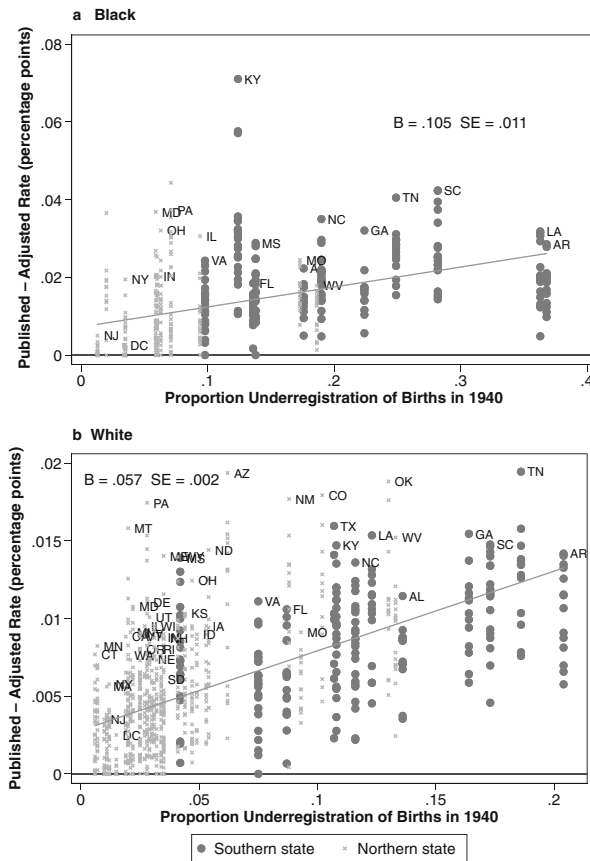


Fig. 2 Bias in published IMR relative to percentage underregistration: single-year IMR. Each single-year difference between adjusted and published IMRs is plotted against the percentage of underregistration from the 1940 test for each race and state. Underregistration is measured only once for each state, and thus all observations for a state are on the same vertical line. The slope coefficient and standard error from the regression line are estimated with controls for year of birth. See the online data appendix for a discussion of authors' calculations and sources used

information on state of birth for deaths in this age group. The bias from this source is limited because the out-of-state migration rate for infants was small (less than 1 % in 1940), most infant deaths occurred in the first 30 days of life, and the likelihood of migration with a sick infant was relatively small.

Deaths of noninfant children may pose a larger concern given that both the cumulative likelihood of migration and the hazard rate increase with age, implying an increased potential for noninfant children to die outside their state of birth. Working in the opposite direction, however, is the fact that mortality rates decrease rapidly after the first year of life, as do cross-state differences in age-specific mortality. In practice, bias from migrant deaths is small. Figure 3 plots adjusted infant mortality when noninfant deaths are allocated to state of birth versus adjusted infant mortality when noninfant deaths are allocated to state of

occurrence.¹⁴ The methods have a tight, almost one-for-one, relationship. Differences do arise, however, from the high rates of out-migration from southern states with large black populations during the Great Migration. Nevertheless, these differences are small. As such, we choose to allocate noninfant deaths to states of birth in the revised estimates, but we emphasize the limited importance of migration in this context.

Finally, a downward bias can enter through the numerator from nonregistered infant deaths. When a parent decides against registering a death, no record of the event exists, and thus no direct means to assess the size of death underregistration is available (Greville 1947). To our knowledge, no contemporary evidence exists for the special case of the extent of underregistration for infant deaths. Contemporaries clearly believed that the issue was less severe than for birth registration (Whelpton 1934; Wilbur 1916). Supporting this view, incentives were in place for death registration that were absent for birth registration. A cemetery burial, with the family or in churchyard, required a burial permit, which was issued only after a death was registered and a certificate was created. In the absence of a direct assessment of the potential bias from death underregistration, our revised rates provide a lower bound on infant mortality in the presence of death underregistration, whereas published rates provide an upper bound.¹⁵

As an additional robustness check, and to help illuminate the sources of potential bias across estimation methods, we present a second adjusted series in which registered births in every year are scaled by the extent of underregistration from the 1940 test reported in Grove (1943). The adjusted IMR by scaling births can be expressed as follows:¹⁶

$$IMR_{s,r,t}^{SCALED} = \frac{Published\ Deaths_{s,r,t}}{Adjusted_{SCALED}\ Births_{s,r,t}}.$$

Biases in scaled rates stem from changes over time in the extent of completeness of birth registration.¹⁷ The 1940 estimate of underregistration provides an increasingly

¹⁴ Rates with noninfant deaths allocated to state of birth are the preferred revised rates and correspond to Adjustment 4 in the online appendix. The procedure allocates the number of age-specific reported noninfant deaths in each *state of occurrence* to *states of birth* using the age-birth-state breakdown in the complete count censuses. For example, if 10 % of black 8-year-olds living in Illinois in the 1940 census were born in Mississippi, then 10 % of black noninfant deaths in Illinois are apportioned to black births in Mississippi for the 1932 birth year. Rates with noninfant deaths recorded in the state of occurrence corresponds to Adjustment 2 in the online appendix. Figure A2 in the online appendix plots the relationship for noninfant deaths.

¹⁵ In the machine-readable files, we provide an additional adjusted rate that does not scale up census counts by the extent of underenumeration, as in our preferred estimate. In some sense, underenumeration that enters the denominator provides a balance against error in the numerator from unregistered infant deaths. However, the differences in the two series may not be important in some contexts. Scaling by the extent of underenumeration causes a level shift down in IMR but does not affect the overall trend (see Fig. A4 in the online appendix). Moreover, the scaling does not affect results in a panel setting with year and state fixed effects.

¹⁶ This scaled rate corresponds to Adjustment 5 in the online appendix.

¹⁷ For births during the six months prior to the April census date in 1940, the estimates of registration completeness contained in Grove (1943) provide an accurate measure of the bias in infant mortality calculations. As such, we are confident in their use to make adjustments at the state level for 1939 and 1940. See the discussion of Adjustment 3 in the online appendix for more information.

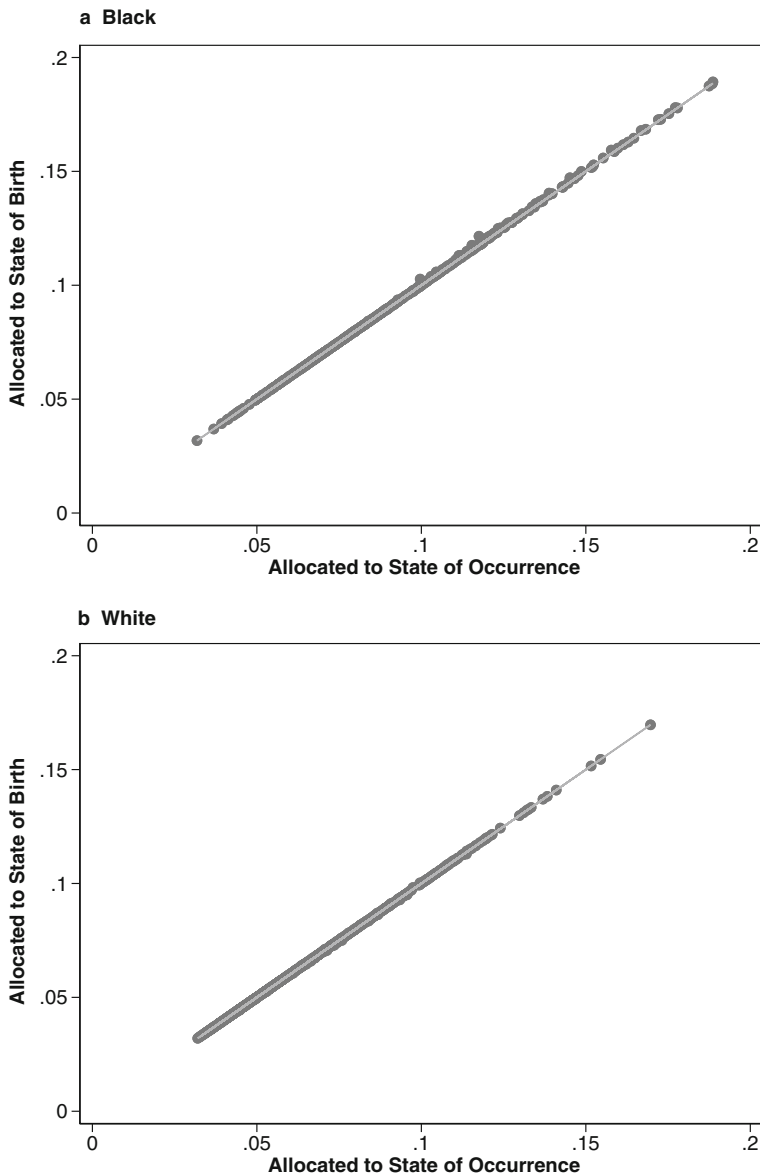


Fig. 3 Relationship between infant mortality allocating noninfant deaths to state of birth versus state of occurrence. Deaths of noninfants reported by age, race, and state of occurrence in the published tables are allocated to states of birth using the proportion of residents in each state of birth in the complete count census microdata for 1920, 1930, and 1940 by race and age. Observations are limited to states with at least 1,000 births for the figure showing black rates (panel b). See the online data appendix for a discussion of authors' calculations and sources used

uncertain or inaccurate method of adjustment the more distant the year of birth is from 1940. The processes that lead to registration—such as states placing importance on birth registration, and the proportion of births in hospitals or attended by a physician—

evolve gradually over time.¹⁸ Underregistration, then, likely followed a downward trend, introducing some bias into the scaled IMR estimates.

How should a researcher choose between the revised and published estimates? Comparing the potential sources of bias and how they vary across time and place is helpful to distinguish the proper estimate. A downward bias from unregistered infant deaths enters the numerators of both $IMR^{ADJUSTED}$ and IMR^{SCALED} , whereas bias from time-varying registration completeness affects only IMR^{SCALED} . In the end, we suggest using both IMR estimates as well as the original published rates to check any results for robustness. The bias present in any one of the three suggested estimates behaves differently in the cross-section and over time. Researchers who demonstrate that estimates are robust to the choice of series provide convincing evidence of a true effect. Additionally, the various rates can be used to provide a range of values for trends or group differences.

Finally, we want to emphasize that a major contribution of our work is to produce IMRs for states prior to entering the BRA. Most states entered the DRA before meeting the requirements to enter the BRA. We use the reported infant death counts in the mortality statistics volumes and our own estimates of births to construct infant mortality estimates for states prior to their entrance to the BRA.¹⁹ The additional data allow researchers to extend analysis further into the past.

Implications

We close by discussing a number of implications that arise from using revised IMRs in place of the published estimates. We first graphically show national trends in IMR by race and the black-white gap. The most important changes from using revised estimates are on cross-sectional comparisons, such as the pace and timing of regional convergence in the North–South difference. We then revisit Collins and Thomasson (2004) to conduct an Oaxaca-Blinder decomposition of the national black–white IMR gap using state-level socioeconomic status measures as explanatory variables.

Implications for National-Level IMR

Figure 4 plots three IMR series—published, restricted sample adjusted, and full sample adjusted—separately for blacks (panel a) and whites (panel b). The restricted sample adjusted series limits the sample to state-year observations that are also in the published series (i.e., in the BRA). The full sample adjusted series lifts that restriction and includes state-year observations for which our method fills a hole in the published series (i.e., the state is part of the DRA but not the BRA). Differences between the published and restricted sample adjusted series arise solely from differences in birth

¹⁸ The proportion of births registered clearly varies over time within a state. A simple way to argue the point is to notice the large differences in registration rates by whether the birth occurred in a hospital as well as the rapid increase in the proportion of hospital births over time. The 1940 test showed that 98.5 % of all hospital births were registered, and 86.1 % of births were outside hospitals. Moriyama (1946) estimated that only 36.9 % of births were hospital deliveries in 1935 but that this figure increased to 55.8 % in 1940 and to 75.6 % in 1944.

¹⁹ Table A2 in the online appendix lists the years and states for which new estimates are available.

estimates, not from changes in the composition of states. Differences between the published and full sample adjusted series arise from both changes in the composition of states and birth estimates.

Holding the sample of states constant between series, panel a of Fig. 4 suggests that adjustments to black rates lead to a level shift in IMR but not to any meaningful change in the trend. Prior to 1925, this meant primarily states in the Northeast and Midwest,

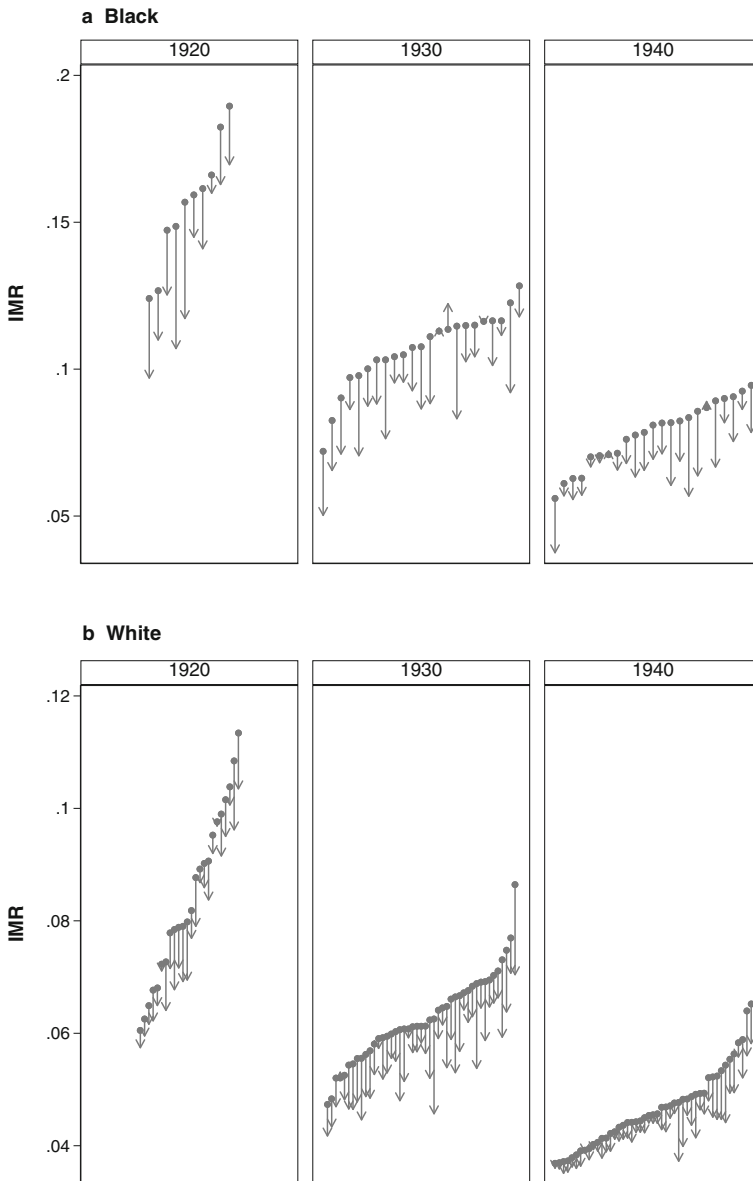


Fig. 4 Change in five-year average IMR by race and period from published to adjusted rates for states in the BRA. Each dot denotes a published IMR, and the placement of an arrow denotes an adjusted IMR. IMR is the average over the five years ending in the census year. States enter the sample when they enter the BRA

where blacks experienced elevated rates of mortality compared with the southern states that were not yet included. However, adding the low-IMR southern states, as in the full sample adjusted series, reduces IMR substantially in early years: by 30.2 % in 1915. As more southern states enter the BRA, the “Full Adj.” and “Restricted Adj.” series converge and become identical when the entrance of Texas completes the BRA in 1933. The evidence suggests that black health was not as poor as contemporaries thought, but it also implies that progress in black health proceeded at a slower rate: a fall in IMR of 6.8 percentage points from 1915 to 1940 compared with 10.9 percentage points in the published data.

Because black births were much more likely than white births to go unregistered, adjustments clearly reduce IMRs for blacks relative to whites at the national level, as shown in Fig. 5. The figures make clear that adjustments lead to a shift in the level of both the absolute and relative black–white gap in IMR but not to a revision in the trend. Thus, we find that the gap started from a smaller initial level but fell at roughly the same rate in terms of percentage points. Our understanding of national trends in the IMR gap does not seem to be greatly changed.

Implications for Cross-State Comparisons of IMR

The large variation across states in the quality of birth registration data, however, leads to significant revisions of cross-sectional comparisons. Figure 6 illustrates the magnitude of changes in IMRs from the adjustment procedure. The bias in the published rates was larger for blacks, for southern states, and in earlier decades. Figure 7 illustrates the number and magnitude of rank changes between the published and revised rates, capturing the impact on cross-sectional comparisons. The left *y*-axis ranks states by published IMR; the right *y*-axis ranks states by revised IMR, with the values for a state connected by a line. A downward slope in the line implies an improvement in rank. Panel a of Fig. 7 shows several rank changes, many of a large magnitude. In general, the southern states for which the revision lowered the IMR show improvements in rank at the expense of states in the Northeast and Midwest.

The effects of rank changes extend to regional differences and any subsequent convergence. In 1915, the South initially had a mortality advantage over the North for black infants, as shown in Fig. 8.²⁰ Much of the gap is explained by the existence of a black urban–rural penalty combined with the fact that blacks in the North lived in cities but were primarily rural in the South.²¹ IMRs in the North converged with those in the South as the urban penalty gradually declined over the course of the early twentieth century. In the published data, the North overtook the South by the early 1930s in terms of black infant health.

Three main implications emerge from using the revised estimates. First, the southern mortality advantage widens as the adjustment method primarily lowers black IMR in the South. Second, starting from a lower initial IMR in the South implies a faster

²⁰ We do not observe a similar regional convergence for whites as the urban penalty for infant whites had disappeared by 1920.

²¹ According to the 1940 decennial census, 89 % of blacks lived in urban areas in the North census region, whereas 34 % were urban dwellers in the South census region. Data underlying these calculations come from the full count 1940 census microdata from IPUMS.

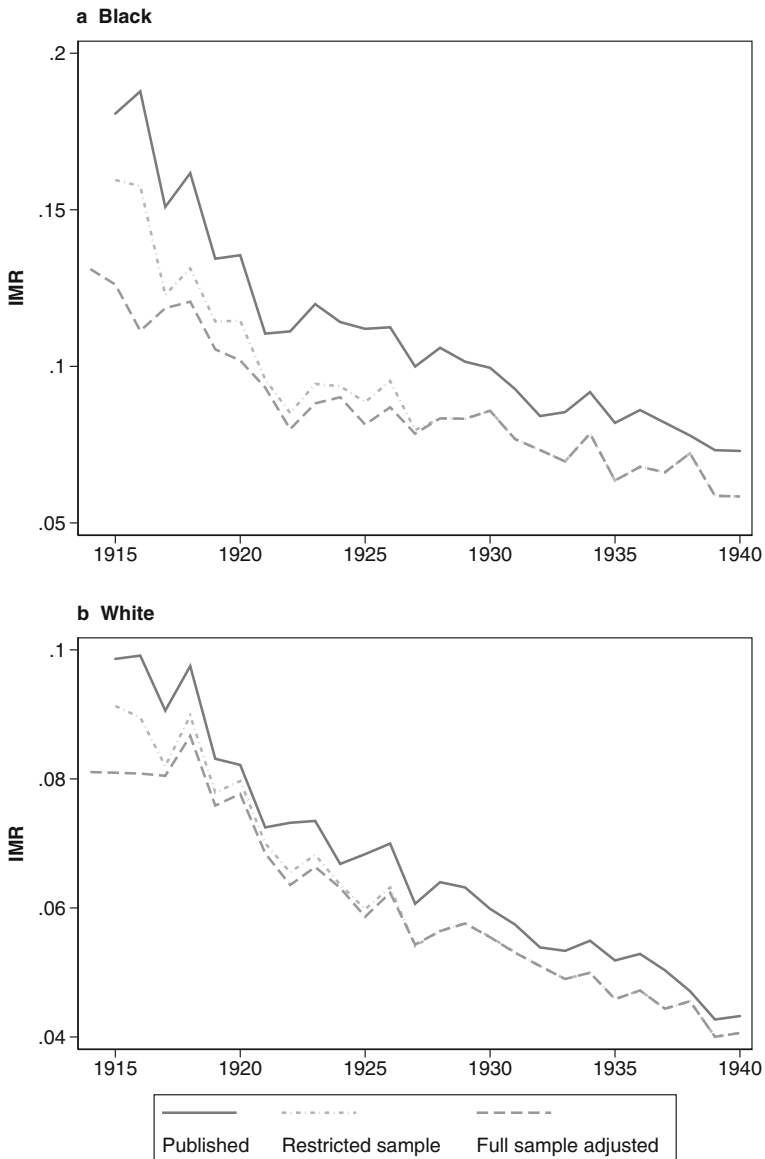


Fig. 5 Published versus adjusted national-level rates. Published rates include states in the BRA. The restricted sample adjusted series consists of the same set of states used in the calculation of the published series. The full sample adjusted series includes all states for which new rates exist. See the online data appendix for a discussion of authors' calculations and sources used

convergence rate between the regions. Finally, the level shift downward in southern IMR delays the North overtaking South until the late 1930s, if at all before 1940.

To illustrate the importance of our adjustments to cross-region comparisons, we reprint IMR comparisons from Eriksson and Niemesh (2016), who estimated the effect

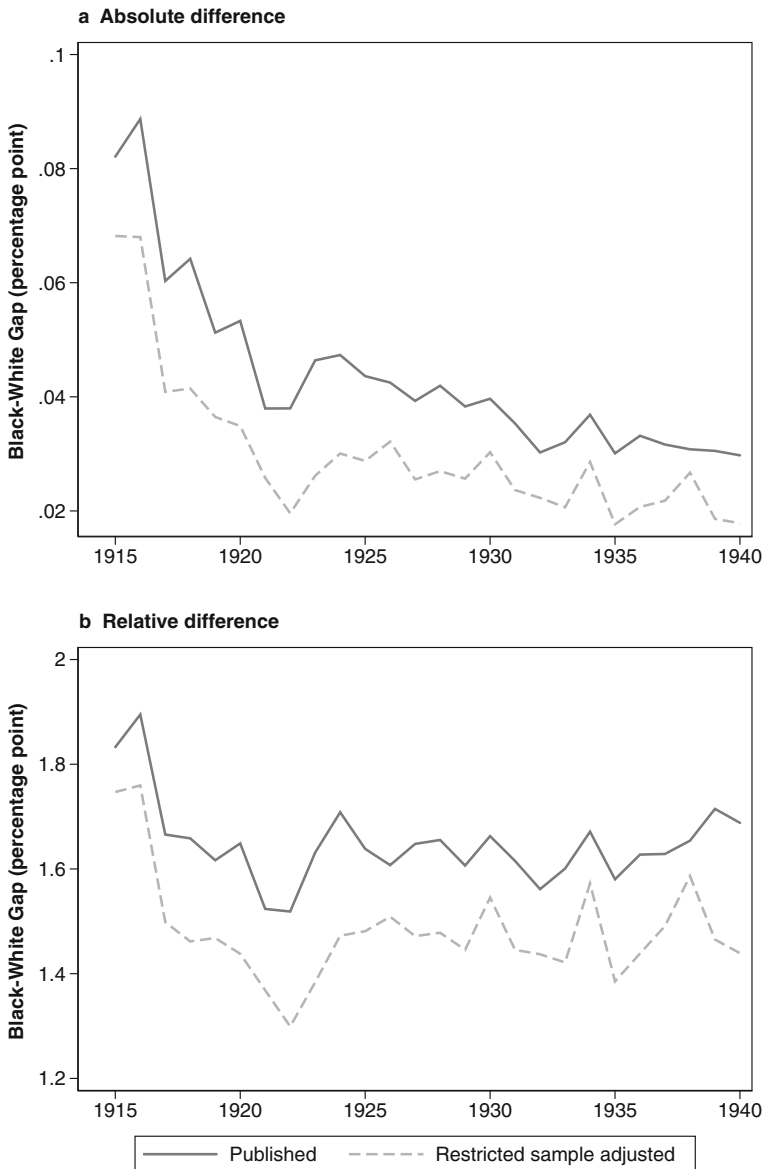
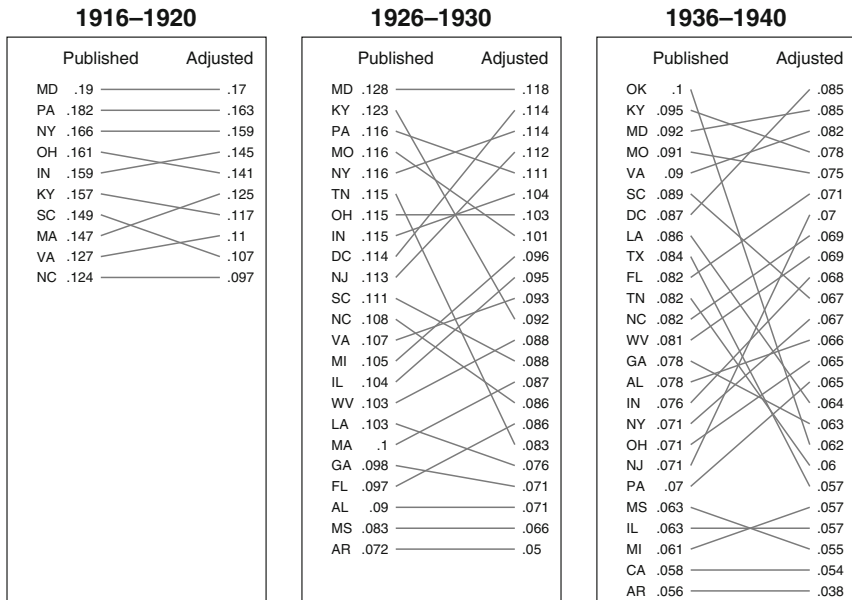


Fig. 6 Black–white gap in IMR: published versus adjusted rates. The same set of states is used in the adjusted rates as in the published rates. States enter the sample as they enter the BRA. See the online data appendix for a discussion of authors’ calculations and sources used

on the subsequent birth outcomes of infants to southern-born black parents after moving to the North during the first half of the Great Migration. Here, we are concerned solely with the observed differences in black IMR across regions as an indicator of the health environments from which blacks left and in which they settled. Table 1 reports regional comparisons with published estimates and revised estimates. The change in inference induced by the bias from underregistration of births is clear. In

a Black



b White

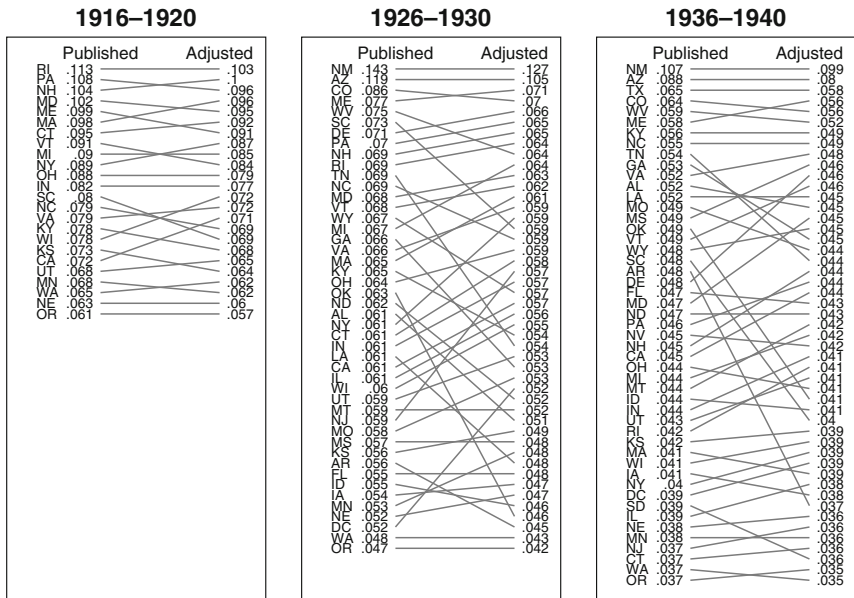


Fig. 7 Change in five-year average IMR ranking from published to adjusted rates. For each state in the BRA for a given period and with at least 5,000 births for each year over the five years, the chart ranks each state by published IMR on the left and adjusted IMR on the right. See the online appendix for a discussion of authors' calculations and sources used

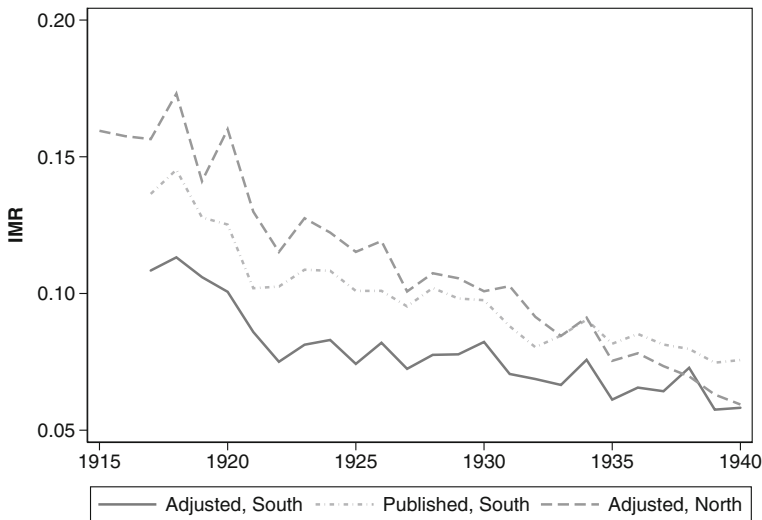


Fig. 8 Regional convergence of black IMR between southern and northern states. States are included in calculations as they enter the BRA. Published rates for the North are not shown because they are almost identical to the adjusted rates

the published data, black infant mortality was initially 33 % higher (4.4 percentage points) in the North, with the southern mortality advantage declining to only 10 % (1.1 percentage points) by the late 1920s and disappearing completely in the 1930s. The revised data widen the initial gap such that the IMR in the North is 52 % higher than in the South and increases the southern mortality advantage in all decades (rows labeled *Diff* in Table 1). Additionally, we find that IMRs were almost identical in 1940 rather than that the North overtook the South, as in the published data. Finally, the last row of Table 1 shows the bias in the regional comparison, calculated as the regional difference in the published data minus the regional difference in the revised data. The magnitude of the negative bias in each period is large: 23 %, 127 %, and 118 % of the published regional IMR difference. Clearly, accounting for underregistration bias with our revised rates dramatically changes the interpretation of the differential health risks faced by black infants across the two regions.

Replication of Collins and Thomasson (2004)

Finally, we use the revised state-level infant mortality rates to revisit findings of Collins and Thomasson (2004), who decomposed explanatory factors of the racial gap in infant mortality for the period 1920–1970. One of their main findings was that measures of income, urbanization, women’s education, and physicians per capita (broadly interpreted as socioeconomic status) explained a large portion of the black–white IMR gap prior to 1945 but a vanishingly small portion afterward. We show that after the underregistration of births is accounted for in the revised IMR estimates, the interpretation of the decomposition dramatically changes.²²

²² The authors were well aware of the underregistration of births and discussed how potential bias might enter their estimates. However, at the time, no direct way of accounting for the bias was available.

Table 1 Regional comparison of black IMR (percentage points)

	1916–1920	1926–1930	1936–1940
Published			
North	17.7	12.2	7.9
South	13.2	11.1	9.0
Difference	4.4	1.1	–1.1
Adjusted			
North	15.7	11.1	7.3
South	10.3	8.6	7.1
Diff	5.4	2.5	0.2
Bias	–1.0	–1.4	–1.3

Notes: The difference is North minus South IMR. Bias is calculated as the published rate minus the adjusted rate. States included in the sample are Illinois, Ohio, and Pennsylvania for the North; and North Carolina, South Carolina, and Tennessee for the South. The sample is limited to these states because of data limitations discussed in Eriksson and Niemesh (2016).

Collins and Thomasson (2004) ran an Oaxaca-Blinder decomposition of the black–white IMR gap in the period 1920–1970. Using observations taken every five years at the state and race level, they first regressed the natural log of IMR on physicians per capita and race-specific measures for income, women’s education, and urban status, and a set of year fixed effects. The β s were averaged over race for the decomposition. Table 2 juxtaposes the results of the published Collins and Thomasson decomposition and our revised IMRs. In the published IMR estimates, the explained gap makes up between 75 % and 96 % of raw difference prior to 1945, with SES (income and education) providing the majority of explanatory power.

Three major differences in the findings emerge when we conduct an identical decomposition procedure on revised rates.

1. A smaller raw black–white IMR gap emerges, not surprisingly, because the adjustment procedure lowers IMR relatively more for blacks than for whites.
2. The percentage “explained” by controls is significantly reduced, by up to 40 % after 1940, because of a change in estimated β s. By reducing infant mortality for blacks in the South—the low-income region for blacks—the strong correlation between income and IMR found in the original data is weakened. The change in explanatory power varies prior to 1945: a 17 % reduction in 1930 to a 6 % increase in 1940.
3. The contribution of racial income differences to the IMR gap is reduced by close to a factor of 10. Education, on the other hand, is only slightly reduced and remains the most important explanatory factor. Physicians per capita doubles in importance.

In summary, the use of corrected IMRs can change conclusions in meaningful ways in empirical exercises originally conducted with published vital statistics that include bias from the underregistration of births.

Table 2 Black–white decomposition results from Collins and Thomasson (2004) and using revised rates

	1920	1925	1930	1935	1940	1945	1950	1955	1960	1965	1970
A. Collins and Thomasson (2004)											
Published Rates											
Total log IMR gap	0.455	0.462	0.493	0.447	0.521	0.454	0.496	0.583	0.626	0.630	0.559
Gap “explained” by:											
Income	0.110	0.106	0.134	0.131	0.132	0.116	0.111	0.102	0.097	0.091	0.082
Education	0.337	0.293	0.343	0.329	0.307	0.265	0.223	0.166	0.121	0.093	0.069
Urban	-0.106	-0.079	-0.082	-0.069	-0.056	-0.037	-0.014	0.000	0.011	0.019	0.029
Physicians	0.034	0.027	0.034	0.038	0.042	0.037	0.029	0.022	0.016	0.012	0.007
Total “explained” gap	0.375	0.347	0.429	0.429	0.425	0.381	0.349	0.290	0.245	0.215	0.187
Total “unexplained” gap	0.079	0.115	0.064	0.017	0.097	0.073	0.147	0.293	0.381	0.414	0.372
Percentage “explained”	0.83	0.75	0.87	0.96	0.81	0.84	0.70	0.50	0.39	0.34	0.33
B. Revised Rates											
Total log IMR gap	0.341	0.356	0.429	0.312	0.366	0.454	0.496	0.583	0.626	0.63	0.559
Gap “explained” by:											
Income	0.007	0.006	0.007	0.007	0.007	0.006	0.006	0.005	0.005	0.005	0.004
Education	0.341	0.292	0.331	0.314	0.299	0.247	0.208	0.155	0.113	0.087	0.064
Urban	-0.136	-0.096	-0.086	-0.073	-0.061	-0.038	-0.015	0	0.011	0.019	0.03
Physicians	0.074	0.054	0.056	0.062	0.069	0.059	0.046	0.035	0.025	0.019	0.01
Total “explained” gap	0.286	0.255	0.308	0.31	0.314	0.274	0.246	0.195	0.154	0.13	0.108
Total “unexplained” gap	0.055	0.100	0.122	0.002	0.051	0.180	0.250	0.388	0.472	0.500	0.450
Percentage “explained”	0.84	0.72	0.72	0.99	0.86	0.60	0.50	0.33	0.25	0.21	0.19

Notes: The “Total log IMR gap” is the difference between the average (weighted by population) log white and nonwhite IMRs in each year. Each component of the “Gap explained by” section is the product of the difference in the variables average value for whites and nonwhites (in that year) and the relevant coefficients from a race-specific regression of infant mortality for the full 1920–1970 sample period. See Collins and Thomasson (2004: table 2).

Sources: See the text for construction of revised IMR for 1920–1940; see Collins and Thomasson (2004) for a description of the variables used in the decomposition and source of IMR for 1945–1970.

Conclusions

Researchers who study long-run trends and racial gaps in infant mortality have long relied on public vital statistics records, which play an important role when targeting, evaluating, and executing public health interventions. Unfortunately, known biases from underregistration of births have hindered our understanding of public health crises, trends, and the evolution of racial health disparities. Using newly released census microdata, we construct revised infant mortality series using a method based on the census enumeration of live children to obtain improved estimates of the number of births. To resolve the bias from underenumeration, when the census undercounts the number of children alive at the census date, we scale the count of children in the census by estimates of the extent of underenumeration (Hacker 2013; Land et al. 1984; Preston et al. 2003).

Using the revised series, we are able, for the first time, to get a sense of the magnitude of the biases caused by underregistration and their implications for research on the trends and determinants of infant mortality. We find that correcting for the underregistration of births, which was particularly problematic for blacks, lowers the IMR for blacks relative to native-born whites. Moreover, this shift downward in the black IMR implies a faster convergence rate between black and white infant mortality before 1940. Revisiting Eriksson and Niemesh (2016) and Collins and Thomasson (2004), we show that using the revised rates does affect their findings. For Eriksson and Niemesh (2016), accounting for the underregistration bias changes the interpretation of the differential health risks faced by black infants in the North versus the South. For Collins and Thomasson (2004), the percentage of the racial gap “explained” by the covariates is reduced, and physicians per capita play a greater role in explaining the gap.

How can and should scholars use these series in their research? Each series contains biases that behave differently in the cross-section and over time. The published series suffers from an undercount of births from an incomplete birth registration. Albeit a better estimate of births than the published series, the adjusted series undercounts births to the extent that underenumeration in the census is not fully accounted for in our procedure. Finally, error enters both the published and adjusted rates through the numerator from a miscount of the number of infant deaths, for which, to our knowledge, there is no estimate of the magnitude. As a result, we suggest using both the revised and published IMR series to check results for robustness. Researchers who demonstrate that estimates are not sensitive to the choice of series provide convincing evidence of a true effect. Additionally, the various rates can be used to provide a range of values for trends or group differences.

An additional benefit of the revised series is that we are able to extend the U.S. IMR series backward, to as early as 1910 in many states. For some states, such as Missouri, this enables researchers to look back 16 years earlier than the published estimates. Even states in the Northeast, such as Massachusetts, now have data that enable analysis to extend five years earlier than previously possible. Given that U.S. public health transitioned rapidly in the early twentieth century, the revised estimates will enable scholars to augment the large body of literature on public health and the mortality transition in the United States, including state-level programs, such as the Sheppard-Towner public health program (Moehling and Thomasson 2014), occupational

licensing in the health professions (Anderson et al. 2016), women's suffrage (Miller 2008), or in studies using state-level data (Hansen 2014; Jayachandran et al. 2010).

The analysis in our study could be extended in at least two ways. First, the published state-level infant mortality series contains a systematic upward bias throughout the postwar period until the 1970s, when underregistration of births ceased (U.S. Census Bureau 1973). As complete microdata are released for the 1950 and 1960 decennial censuses, our adjustment method can be used to correct the state-level infant mortality series for the 1940s and 1950s. Second, the 1940 test of birth registration completeness showed wide variation across counties within a state in the percentage of births registered, suggesting that local-level published infant mortality requires a correction. Extending our adjustment to the local level is a priority given that much of the research on the U.S. mortality transition uses IMR data and interventions at the county- or city-level: for example, water and sewage (Cutler and Miller 2005), milk safety laws (Komisarow 2017), lead water pipes (Clay et al. 2013; Troesken 2008), rural electrification (Lewis 2018), and access to hospital care (Thomasson and Treber 2008), among many others.

Our findings also have implications for developing and evaluating policy in less-developed countries. We show that mismeasured birth registration can bias IMR and distort policy analysis. We encourage all researchers using IMR data to become familiar with the level of birth registration underlying the estimates and to recognize how potential underregistration affects outcomes of interest.

Acknowledgments A set of machine-readable files of revised births and infant mortality rates is available online through the Inter-university Consortium for Political and Social Research (ICPSR 37076, version 1). We are particularly grateful to William Even, Analisa Packham, the editors of this journal, and three anonymous referees for helpful suggestions. We also thank Jeremy Attack, William J. Collins, Dora Costa, Gordon Hanson, Adriana Lleras-Muney, Seth Sanders, Marianne Wanamaker, and Sven Wilson for comments when portions of this work were included in the paper "Death In the Promised Land: The Great Migration and Black Infant Mortality." Brian Lee and Man-Ting Chang provided excellent research assistance.

References

- Anderson, D. M., Brown, R., Charles, K. K., & Rees, D. I. (2016). *The effect of occupational licensing on consumer welfare: Early midwifery laws and maternal mortality* (NBER Working Paper No. 22456). Cambridge, MA: National Bureau of Economic Research.
- Bhalotra, S., & Venkataramani, A. (2011). *The captain of the men of death and his shadow: Long-run impacts of early life pneumonia exposure* (IZA Discussion Paper No. 6041). Bonn, Germany: Institute for the Study of Labor.
- Cassedy, J. H. (1965). The registration area and American vital statistics: Development of a health research resource, 1885–1915. *Bulletin of the History of Medicine*, 39, 221–231.
- Clay, K., Troesken, W., & Haines, M. R. (2013). Lead and mortality. *Review of Economics and Statistics*, 96, 458–470.
- Collins, W. J., & Thomasson, M. A. (2004). The declining contribution of socioeconomic disparities to the racial gap in infant mortality rates, 1920–1970. *Southern Economic Journal*, 70, 746–776.
- Cutler, D., & Miller, G. (2005). The role of public health improvements in health advances: The twentieth-century United States. *Demography*, 42, 1–22.
- Davis, W. H. (1925). Necessity for completing the registration area by 1930. *American Journal of Public Health*, 15, 399–404.
- Deacon, W. J. V. (1937). Tests and promotion of registration of births and deaths. *American Journal of Public Health*, 27, 492–498.
- Eriksson, K., & Niemesh, G. T. (2016). *Death in the promised land: The Great Migration and black infant mortality*. Retrieved from <https://ssrn.com/abstract=3071053>

- Eriksson, K., Niemesh, G. T., & Thomasson, M. (2018). *Revised infant mortality rates and births for the United States 1915–1940* (ICPSR 37076, version 1) [Computer file]. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor]. <https://doi.org/10.3886/ICPSR37076.v1>
- Greville, T. N. E. (1947). *United States life tables and actuarial tables, 1939–1941*. Washington, DC: National Office of Vital Statistics.
- Grove, R. (1943). *Studies in the completeness of birth registration, part I. Completeness of birth registration in the United States, December 1, 1939 to March 31, 1940* (Vital Statistics—Special Reports Series, Vol. 17, No. 18). Washington, DC: National Office of Vital Statistics.
- Hacker, J. D. (2013). New estimates of census coverage in the United States, 1850–1930. *Social Science History*, 37, 71–101.
- Hansen, C. W. (2014). Cause of death and development in the US. *Journal of Development Economics*, 109, 143–153.
- Hedrich, A. W., Collinson, J., & Rhoads, F. D. (1939). *Comparison of birth tests by several methods in Georgia and Maryland* (Vital Statistics—Special Reports Series, Vol. 7, No. 60). Washington, DC: National Office of Vital Statistics.
- Jayachandran, S., Lleras-Muney, A., & Smith, K. V. (2010). Modern medicine and the twentieth century decline in mortality: Evidence on the impact of sulfa drugs. *American Economic Journal: Applied Economics*, 2(2), 118–146.
- Komisarow, S. (2017). Public health regulation and mortality: Evidence from early 20th century milk laws. *Journal of Health Economics*, 56, 126–144.
- Land, K. C., Hough, G. C., Jr., & McMillan, M. M. (1984). New midyear age-sex-color-specific estimates of the U.S. population for the 1940s and 1950s: Including a revision of coverage estimates for the 1940 and 1950 censuses. *Demography*, 21, 623–645.
- Lenhart, R. F. (1943). Completeness of birth registration in the United States in 1940. *American Journal of Public Health*, 33, 685–690.
- Lewis, J. (2018). Infant health, women’s fertility, and rural electrification in the United States, 1930–1960. *Journal of Economic History*, 78, 118–154.
- Lunde, A. S. (1980). *The organization of the civil registration system of the United States* (Technical Report No. 8). Bethesda, MD: International Institute for Vital Registration and Statistics.
- Miller, G. (2008). Women’s suffrage, political responsiveness, and child survival in American history. *Quarterly Journal of Economics*, 123, 1287–1327.
- Moehling, C., & Thomasson, M. (2014). Saving babies: The impact of public health education programs on infant mortality. *Demography*, 51, 367–386.
- Moriyama, I. (1946). *Estimated completeness of birth registration: United States, 1935 to 1944* (Vital Statistics—Special Reports Series, Vol. 23, No. 10). Washington, DC: National Office of Vital Statistics.
- Moriyama, I. (1990). *Measurement of birth and death registration completeness* (Technical Report No. 43). Bethesda, MD: International Institute for Vital Registration and Statistics.
- Preston, S. H., Elo, I. T., Hill, M. E., & Rosenwaike, I. (2003). *The deemography of African Americans 1930–1990*. Dordrecht, the Netherlands: Kluwer Academic Publishers.
- Ruggles, S., Genadek, K., Goeken, R., Grover, J., & Sobek, M. (2017). *Integrated public use microdata series: Version 7.0* [Machine-readable database]. Minneapolis: University of Minnesota. <https://doi.org/10.18128/D010.V7.0>.
- Shapiro, S. (1950). Estimating birth registration completeness. *Journal of the American Statistical Association*, 45, 261–264.
- Shapiro, S., & Schachter, J. (1952). Birth registration completeness, United States, 1950. *Public Health Reports (1896–1970)*, 67, 513–524.
- Thomasson, M., & Treber, J. (2008). From home to hospital: The evolution of childbirth in the United States, 1928–1940. *Explorations in Economic History*, 45, 76–99.
- Troesken, W. (2008). Lead water pipes and infant mortality in turn-of-the-century Massachusetts. *Journal of Human Resources*, 43, 553–575.
- U.S. Census Bureau. (1973). *Test of birth registration completeness 1964 to 1968* (Evaluation and Research Program Series PHC(E)-2). Washington, DC: U.S. Government Printing Office.
- Whelpton, P. K. (1934). The completeness of birth registration in the United States. *Journal of the American Statistical Association*, 29, 125–136.
- Wilbur, C. (1916). *The federal registration service of the United States: Its development, problems, and defects*. Washington, DC: U.S. Government Printing Office.
- Wilcox, W. F. (1933). *Introduction to the vital statistics of the United States, 1900 to 1930*. Washington, DC: U.S. Department of Commerce, U.S. Census Bureau.