CHAPTER Ab
Vital Statistics

Editor: Michael R. Haines

VITAL STATISTICS

Michael R. Haines

Vital statistics are indicators of the two major dynamic biological forces that shape population change: births and deaths. Sociological factors that can have a profound effect on population dynamics, such as marriages and divorces, are often considered vital statistics, but in this work they are treated separately in Chapter Ae, on family and household composition. Economic forces regulating in-migration and out-migration are treated in Chapter Ad, on international migration. Political forces that result in annexation of new territory that has a resident population or that dictate a partition of a national territory can also affect the population of a nation. In the American experience, boundary changes have almost always resulted in the acquisition of new territory (see Chapters Cf, on geography and the environment, and Ef, on outlying areas).

The official census taken every ten years provides a snapshot of the American population. The Census Bureau literally attempts to count every resident within U.S. territory. Mandated by the Constitution, these static population counts are the responsibility of the federal government. The collection of vital statistics on births and deaths, however, was left to state and local governments. In consequence, vital registration was instituted unevenly. The legal requirement for registering births and deaths and issuing birth and death certificates did not apply to the entire United States until 1933. Although some cities (for example, Boston, New York, Philadelphia, Baltimore, New Orleans) began vital registration early in the nineteenth century, the first state in the United States to do so was Massachusetts in 1842. An official Death Registration Area consisting of ten states and the District of Columbia was established in 1900 and gradually enlarged until it was completed in 1933. A parallel Birth Registration Area was instituted in 1915 and also reached complete coverage in 1933. The coverage of the official registration areas is given in Table Ab31–37, and the dates of entry into the two areas are given in Table Ab38–39.

The U.S. Census enumerators attempted to collect mortality information in the Censuses of 1850–1900 by questioning each household about the deaths of any of its members during the preceding year. The responses in principle could also be combined with the census of individuals under 1 year of age at the time of the census to calculate information on the number of births. Nevertheless, there were significant problems regarding completeness.

Many households failed to report deaths, and many people died without leaving behind a household that could report. This census information on adult mortality did improve over time, and after 1880 it was merged with state registration data. A similar process was not undertaken for birth counts because the census of infant mortality was considered too weak and the number of states with birth registration systems too few.

A criterion for admission to the official U.S. Death Registration Area after 1900 and the Birth Registration Area after 1915 was that registration be 90 percent complete. As late as 1935, it was estimated that birth registration was about 91 percent complete overall but only 80 percent complete for the nonwhite population. No comprehensive study of death registration completeness has been done, but the data appear to have been less than fully inclusive even in the most compliant states of the Death Registration Area in 1900.

One consequence of the lack of vital registration data before the early part of the twentieth century has been a resort to special estimation techniques and indirect measures of fertility and mortality to gain insight into the demographic transition of the nineteenth century. Most of the statistics presented are simple tabulations or standard demographic rates. However, a number of the newer findings arise from application of rather sophisticated techniques (Preston and Haines 1991; Haines and Preston 1997; Haines 1998a). Estimation of demographic information is important for research in social, demographic, and economic history. Basic demographic structures and events, reflected in birth and death rates, population size and structure, growth rates, the composition and growth of the labor force, marriage rates and patterns, household composition, the levels and nature of migration flows, causes of death, urbanization, and spatial population distribution, determine the "human capital" of society. Demographic events are important both as indicators of social and economic change and as integral components of modern economic growth.

Many of the measures presented here are termed "rates," such as birth and death rates (for example, see Table Ab1–10). Rates are calculated as the number of events (deaths) that took place in a region during a set period (usually a year) divided by the size of the relevant population. Generally, demographers prefer that the relevant population, often called the "base population," be the number of people who could experience the event. In the case of death rates, the base population would be the number of people alive at the beginning of the period plus those who were born during the period. In the case of births, the fertility rate (or general fertility rate) is calculated by dividing the number of births by the number of women of childbearing age, conventionally defined as all

women ages 15–44 years. Sometimes, because of data limitations or other considerations, a different denominator is chosen. Thus, a "crude birth rate" is the number of births in a given year divided by the total midyear population of the region. However defined, once calculated a demographic rate is usually expressed per 1,000 in the base population. A general fertility rate of 70 means that 7 percent of all women ages 15–44 years gave birth during a given year.

One problem that arises in interpreting simple demographic rates is that their value can depend on the underlying age composition of the base population. Thus, a country with many old persons is likely to have a higher crude death rate than one with a preponderance of young adults simply because the probability of an adult dying increases with age. For this reason demographers often calculate age-specific measures, such as the probability of dying at the specific age of 36 years. Age-specific rates can be measured at every single year of age and arranged into what is commonly called a "life table." Life tables are usually presented for a cross section of the population at various ages measured at a single point in time. Calculated in this way it is called a "period life table," in contrast with a "cohort life table," which would follow an actual group of people born at the same time from birth through the death of the last survivor.

Some of the statistics discussed in this essay arise from age-specific measures that have been summarized to be useful and intuitively interpretable. The life table, for example, takes age-specific death rates and converts them into the expectation of life at any age, that is, the average number of years of life remaining if people of that age experienced the age-specific mortality rates embodied in the life table. For example, Table Ab644–655 presents the expectation of life at birth for the white and black populations from 1850 onward. Table Ab656–703 gives the expectation of life at various ages around census dates from 1850 to 1990. Another life table measure is the probability of an infant surviving from birth to the first birthday (exact age of 1 year), which is presented here as the infant mortality rate (infant deaths per 1,000 live births per annum). Table Ab912–927 has the infant mortality rate for the United States, along with the fetal death ratio (the number of stillbirths per 1,000 live births) and neonatal mortality rate (infant deaths in the first twenty-seven days of life per 1,000 live births). The infant mortality rate on an annual basis for Massachusetts for 1850–1998 is found in Table Ab928. Massachusetts is the state with the longest reliable series of annual birth and infant death data.

Similarly, age-specific fertility rates can be summarized. One instance, provided in Table Ab52–117, is the total fertility rate (TFR), which is the sum of age-specific births for all women ages 15–49 years. This can be seen as the average number of births a woman would have if she remained alive throughout her entire reproductive life and if she personally experienced rates of childbearing at each age of life displayed by the cross-sectional age-specific data. Extensions of the TFR, and the gross and net reproduction rates, may be found in Table Ab306–314. The gross reproduction rate is the TFR for female births only. The net reproduction rate (NRR) is the gross reproduction rate adjusted for the mortality of women between birth and their age at the birth of their daughters. An NRR of 1.0 indicates that over the long term a society is reproducing just a replacement number of individuals. A rate above 1.0 indicates positive long-term population growth, and a rate below 1.0 points to a long-term population decline from the natural processes of births and deaths.

Table Ab315–346 provides yet another measure of fertility known as the child/woman ratio, which is the number of surviving children ages 0–4 years per 1,000 women ages 20–44 years. It is a wholly census-based fertility rate, requiring no vital statistics. It is, in fact, the main direct source of information on fertility in the nineteenth century for the United States and is the basis for the early estimates of crude birth rate and TFR also given in Table Ab40–51. The child/woman ratio does have some serious drawbacks because it deals with surviving children at the time of census and not actual births in the preceding five years. It also suffers from relative differences in underenumeration of young children and adult women. It has the advantage of being available back to 1800 for the white population and to 1830 for the black population. Table Ab315–346 provides this measure for geographic divisions and rural–urban residence from 1800 to 1990.

Another measure of fertility available historically is based on a census question that asked each woman how many live births she had experienced. Data on children ever born are set forth in Table Ab536–625 for the whole United States from the Censuses of 1900, 1910, and 1940–1990, and for two state censuses that happened to ask such questions – New York State in 1865 and Massachusetts in 1885 (see Tables Ab626–643).

The Long-Term Decline of Fertility

The young republic was notable for its large families and early marriages. The TFR in Table Ab1–10 indicates an average number of births per woman of approximately seven in 1800, followed by a sustained decline in birth rates up until the late 1940s (see Figure Ab-A). The unusual aspect of the American experience is that the decline began before the nation had developed a substantially urban or industrial character. Both rural and urban birth rates declined in parallel, although rural fertility remained higher throughout the period considered here. Fertility decreased across regions, but the South lagged behind the Northeast and Midwest with regard to the timing and speed of the decline (Table Ab315–346). Even the fertility of the antebellum slave population showed signs of decline just prior to 1860, though family sizes of blacks were, on average, significantly larger than those of whites (Sutch 1975; Steckel 1982).
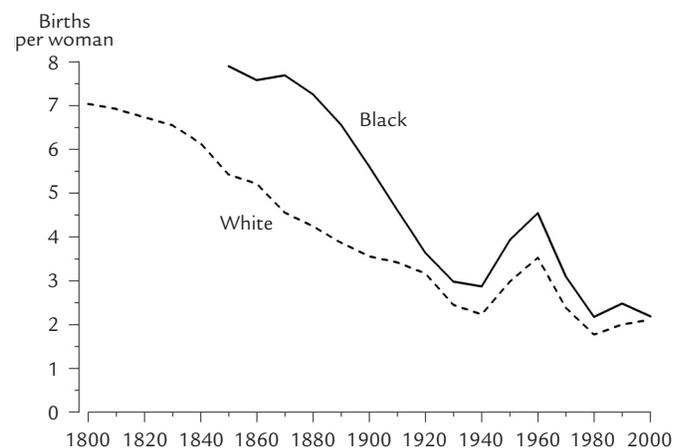

FIGURE Ab-A   Total fertility rate, by race: 1800–2000

Source
Series Ab5–6.

The fertility of the female population will be determined by the typical age at first sexual experience. As women come of child-bearing age in their midteens, postponement of sexual activity until older ages will reduce the number of years during which a woman may become pregnant. Statisticians have very little historical evidence about this extremely private aspect of most people's lives, but demographers suspect that the state of matrimony is highly correlated with sexual activity. Thus, the fraction of women who are married – "nuptiality" – is of interest in the study of the general fertility of a population. Either because of a cultural taboo against premarital sex or because a woman would often marry soon after discovering her first pregnancy, the age at first marriage is also an important demographic variable.

We know relatively little about marriage in the early nineteenth century and what we do know is rather speculative. Around 1800, the age of women at first marriage was probably rather young, perhaps below twenty years. Men on average were several years older at marriage, and all but a relatively small proportion of both sexes eventually married (Haines 1996). The federal census did not include a question on marital status until 1880 and did not begin reporting results until 1890. Several state censuses did, however, ask these questions earlier. A sample of seven New York state counties from the manuscripts of the Census of 1865, for example, reveals an estimated age at first marriage of 23.8 years for women and 26.6 years for men. Percentages of individuals who never married by the ages of 45–54 years were 7.4 percent for women and 5.9 percent for men, pointing to quite low levels of lifetime unmarried status. Although the typical age at marriage was probably higher in New York than in the nation as a whole and although it had very likely risen by 1865, nuptiality was still rather extensive by European standards. The average age at first marriage for women was 25.4 years in England and Wales in 1861 and 26.3 years in Germany in 1871 (with German men having had an average age at marriage as late as 28.8 years).

In 1880, when the U.S. Census first asked a question on marital status, the average age for women at first marriage was 23.0 years, while that for men was 26.5 years. The proportions of individuals who never married by middle age were still relatively low, at 7 percent for both men and women. Age at marriage rose slightly up until 1890 and 1900 and thereafter began a long-term decline up to the 1950s (see Table Ae489–506 and Figure Ab-B). Recent work with the Integrated Public Use Microdata Samples (IPUMS) of the federal census has provided estimates of the median age at marriage for the white population back to 1850 by using imputed marital status for 1850–1870.[1] The median age at marriage is the age at which half of all ever-married women in a cohort are younger and half are older. Those results indicate that the median age at marriage was roughly stable over that period, at about 25 years for men and 21–22 years for women (Table Ae481–488). This is in sharp contrast to the late twentieth century, when the age at marriage rose to its highest point (26–28 years for men and 25–29 years for women). However, in the late twentieth century the link between marriage and childbearing was beginning to weaken. In 1998, about one third of all births were to unmarried women.

A decomposition of the fertility decline into the contributions of nuptiality and marital fertility found that, up to approximately 1850, half of the decline could be attributed to adjustments in marriage age and marriage incidence. Thereafter, most of the decline

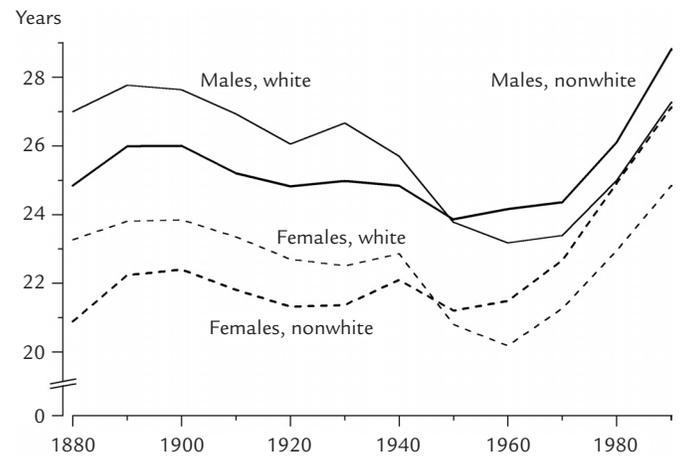[1] See the Guide to the Millennial Edition for information on IPUMS.

FIGURE Ab-B Singulate mean age at first marriage, by sex and race: 1880–1990

Sources
Ae491–492 and Ae501–502.

originated in lower fertility within marriage (Sanderson 1979, pp. 339–58).

Such evidence as we have concerning fertility differentials between native-born and foreign-born women points to relatively small differences at midcentury but generally higher fertility for the foreign-born thereafter. The fertility of native white women continued to decline, while the successive cohorts of incoming migrants continued to produce large families. Birth rates of native-born women of foreign-born parentage were intermediate between those of native white women of native parentage and foreign-born white women, suggesting a form of convergence to native white demographic patterns. Data on children ever born (parity) from a sample of seven New York counties in 1865 revealed few differences between native-born and foreign-born women born near the beginning of the nineteenth century (Table Ab626–636). However, published data from the Massachusetts Census of 1885 showed substantially more births per ever-married foreign-born woman relative to the native-born (Table Ab637–643). Such differentials also appeared in the data from the federal censuses of 1900 and 1910 (Table Ab536–625). Much of the difference was due to the lower age at marriage and lower percentages of individuals remaining single among the foreign-born. However, fertility within marriage was also greater for foreign-born women in the late nineteenth and early twentieth centuries. Relatively few of them, for instance, remained permanently childless. The federal Census of 1910 reported that native white women ages 55–64 years (that is, those born in the years 1846–1855) had an average number of children ever born of 4.4 (4.8 for ever-married women). More than 17 percent of all native white women (and 9 percent of those who married) remained childless. Among the foreign-born enumerated at the same census, the average number of children was 5.5 for all women and 5.8 for ever-married women, with only 12 percent of all women, and 7 percent of ever-married women, remaining childless. Such differentials between native-born and foreign-born women had largely disappeared for those born at the end of the nineteenth century and enumerated in 1940.

The inexorable decline of American birth rates continued apace after the Civil War. However, most of the decline originated in adjustments in fertility within marriage. Recent work with data from

the 1900, 1910, and 1940 federal censuses shows rapid reductions in marital fertility, especially among white urban women. In 1910, for example, more than half of native white urban women ages 45–49 years were estimated to have been effectively controlling fertility within marriage, and about a quarter of rural farm and nonfarm women were doing the same. Among younger women (ages 15–34 years) the proportions were much higher, rising to more than 70 percent for native white urban women and more than half for native white farm women. It could be said that the "two-child norm" was being established in the United States in this era. Some fascinating supporting evidence is furnished by a survey of the wives of professional and white-collar men over the period 1892–1920. Although the sample was small and nonrepresentative, it revealed extensive use of contraceptives and contraceptive practices and active strategies of family limitation. This was a preview of the rapid adoption of such behaviors in the twentieth century (David and Sanderson 1986, 1987).

The usual pattern of fertility transition is that women "stop" having children by lowering their age at last birth. By the 1890s, however, American women were also "spacing," that is, lengthening the intervals between births to reduce completed family size further. New estimates of age-specific fertility rates for the United States around the turn of the century point to low marital fertility at young ages, quite unlike the situation in Europe at the same time (Haines 1990).

The period after 1865 was also marked by reductions in fertility according to residence and race (Table Ab315–346). Rural fertility remained higher than urban fertility, but absolute differences diminished as both populations progressively limited family size. The absolute gap dropped from 474 more children ages 0–4 years per 1,000 women of childbearing age in rural than in urban areas in 1800 to 273 in 1920. As Table Ab1–10 shows, fertility differences by race also tended to converge after the middle of the nineteenth century. Whereas the TFR for blacks was 48 percent higher than that for whites in the 1850s, it was only 15 percent higher in 1920. Differentials in birth rates by race have persisted up to the present and actually widened somewhat after 1920 (Tables Ab40–117). Birth rates also varied across regions both before and after the Civil War, with the South and West having been areas of higher fertility relative to the Northeast and Midwest (Table Ab315–346).

Finally, although we know rather less about the fertility of different socioeconomic status groups, the evidence points to smaller families among higher socioeconomic status groups, such as professionals, proprietors, clerks, and other white-collar workers. This was true from at least the middle of the nineteenth century onward. Among proprietors, however, an exception was farmers who owned their own farms, who, throughout the century, typically had larger families than other groups. Fertility among unskilled workers tended to be closer to that of farmers, while skilled and semiskilled manual workers and craftsmen occupied an intermediate position. These socioeconomic fertility differences may have widened over the course of the nineteenth century before they eventually narrowed (Haines 1993).

One consequence of declining fertility has been the aging of the population. As Table Aa599–613 indicates, the median age of the American population rose from 16 in 1800 to over 20 in 1870, to over 25 in 1920, and finally to almost 33 in 1990. The reason is that births add people of age zero to the population but mortality, in contrast, affects all ages. As fertility declines, so does the proportion of children. The average age of the population rises.

The implications of this are great, changing the society from one oriented toward children to one centered on adults and the elderly.

Explaining the decline in fertility among Americans poses a series of difficult issues. Conventional demographic transition theory has placed great reliance on the changes in child costs and benefits associated with structural changes accompanying modern economic growth, such as urbanization, industrialization, the rise in literacy and education, and increased employment of women outside the home. The classic statement of the theory was made by Frank Notestein in 1953.

> The new ideal of the small family arose typically in the urban industrial society. It is impossible to be precise about the various causal factors, but apparently many were important. Urban life stripped the family of many functions in production, consumption, recreation, and education. In factory employment the individual stood on his own accomplishments. The new mobility of young people and the anonymity of city life reduced the pressure toward traditional behavior exerted by the family and community. In a period of rapidly developing technology, new skills were needed, and new opportunities for individual advancement arose. Education and a rational point of view became increasingly important. As a consequence the cost of child-rearing grew and the possibilities for economic contributions by children declined. Falling death-rates at once increased the size of the family to be supported and lowered the inducements to have many births. Women, moreover, found new independence from household obligations and new economic roles less compatible with childbearing.

But, of course, the fertility transition began in the United States well before many of these structural changes became important.

An early alternative theory to explain the American fertility decline for the antebellum period was the land availability hypothesis (Yasuba 1962; Forster and Tucker 1972). It was noted that in the period before the Civil War, an inverse relationship existed between population density and child/woman ratio. If high population density meant a scarcity of unoccupied agricultural land and high land prices, then the cost to farm families of endowing their children with a suitable means of earning a living in an agricultural society would be high. This may then be presumed to reduce the attractiveness of large families.

An intriguing alternative to the land availability–child bequest hypothesis has been proposed by Sundstrom and David (1988). They argue that smaller families resulted more from the development of nearby nonagricultural labor market opportunities than from the march of the frontier and the disappearance of inexpensive homesteads. With an active urban labor market, larger material inducements were necessary to keep children "down on the farm" once jobs were readily available within easy distance. A related model, that of Ransom and Sutch, emphasizes the westward migration of children who then "defaulted" on their implicit contracts to care for their parents in old age (Sutch 1990). In response, parents began accumulating real and financial assets as a substitute for offspring as retirement insurance, leading to smaller families. In general, however, the land availability, urban labor markets, and child default models are all consistent with the patterns of fertility decline across space and time. Such data do not allow us to discriminate between the models, which are, indeed, not mutually exclusive. Richard Steckel, using published state-level data from the 1850 and 1860 federal censuses, ran some tests on the

competing hypotheses. Although he found modest support for the land availability view, the strongest predictors of marital fertility differentials just prior to the Civil War were the presence of financial intermediaries (banks) and labor force structure (that is, the ratio of nonagricultural to agricultural labor force) (Steckel 1992a). These findings tend to support the Ransom and Sutch and Sundstrom and David theories, respectively.

Recent work by D. S. Smith (1996) with the rich data from the 1915 Census of Iowa provides support for the importance of education. Women with more years of schooling and exposure to grammar school, as opposed simply to the common school with no grades, were more effective in controlling family size. Education would facilitate the transmission of contraceptive knowledge, which may also help to explain the appearance of spacing from early in childbearing that was characteristic of the American experience around the end of the nineteenth century (Haines 1990).

Most of the hypotheses about the American fertility transition can also be fit into the more general model offered by Caldwell (1982). His model focuses attention on the net flow of resources between parents and children over the entire life course. Family limitation sets in when this net flow starts to tilt away from parents and toward children – in other words, when parents typically are transferring more resources to their children than they are receiving from them. This signifies a rise in the net cost of childrearing (that is, benefits minus costs) and is accelerated by factors such as the introduction of mass education (implying more years in school and lower child labor contributions to the family), child labor laws, and a more positive view of the value of transmitting human capital, rather than financial assets or physical capital, from parents to their children.

Before the Civil War, fertility among blacks was heavily influenced by their condition as slaves. Despite the higher infant and child mortality rates among blacks, child/woman ratios for blacks were higher than those for whites, pointing to even larger differential fertility for blacks. Further, the regional pattern was the opposite of that for the white population, with higher black child/woman ratios in the Southeast and lower ratios in the Southwest. Selective movement of adult unmarried slaves to the West and the emphasis on slave reproduction in the Old South likely played a role, as did the quite harsh work regimen on the larger plantations of the New South specializing in cotton and sugar. After the Civil War, the decline in black fertility was more similar in nature to the white fertility transition, influenced by urbanization, industrial development, the growing shortage of good farmland, and changes in family norms. Differential fertility in urban areas actually reversed, with fertility for blacks lower than that for whites in the late nineteenth and early twentieth centuries.

The Baby Boom

Fertility decline continued in the twentieth century, but it was punctuated by one of the most interesting demographic phenomena of modern times – the post–World War II "baby boom." As seen in Figure Ab-A and Tables Ab40–117, birth rates reached a low point in the late 1930s, remained low during World War II, and thereafter rose dramatically until they reached a peak in the early 1960s. Thereafter, birth rates fell to a low point in the middle of the 1970s, followed by a gradual rise until they reached a plateau at about long-term replacement (that is, a TFR of about 2.1). The white population is just about replacing itself, while the nonwhite population has fertility slightly above replacement levels.

The leading explanation for the baby boom and subsequent "baby bust" remains that of Richard Easterlin (1980). It involves an interaction of the small cohort size of young adults in the 1945–1962 period (caused by low birth rates in the 1920s and 1930s) and the prosperous post–World War II economy. These young adults in the peak childbearing years experienced high wages and incomes and low unemployment. They chose both higher consumption, especially housing, automobiles, home furnishings, and appliances, as well as more children. This was in part the result of tastes formed in the 1930s and early 1940s, when consumer goods were less available owing to Depression-era poverty and then war rationing. After about 1962 the process operated in reverse. Couples began experiencing relatively less favorable labor force outcomes and then chose to have fewer children.

The implications of the baby boom and subsequent baby bust are enormous. The large birth cohorts of 1946–1962 have influenced consumer spending, demand for housing, needs for schools and higher education, savings behavior, voting patterns, and many other aspects of the society and economy. The smaller cohort born in the late 1960s and the 1970s posed new challenges. The smaller size of the bust relative to the boom generation has thrown into question long-standing formulas defining Social Security benefits and Medicare. When the "boomers" retire, the smaller succeeding cohorts must assume the burden of paying for these benefits.

Since 1962 there has been what some have called a "second demographic transition" in terms of the family patterns described in Chapter Ae, on family and household composition. There has been a dramatic increase in cohabitation, along with a significant rise in divorce. The proportion of the population living alone has also risen notably, particularly among older individuals.

The Long-Term Decline of Mortality

We know less about the American mortality transition of the nineteenth century than we do about that for fertility. There are no ready census-based mortality measures such as the child/woman ratio, and vital statistics were absent or incomplete for most areas up until the early twentieth century. Demographers have tended to rely either on mortality data from states that had implemented death registration in the nineteenth century, such as Massachusetts, or on samples of genealogical data. Both techniques have disadvantages. Massachusetts began statewide civil vital statistics registration in 1842, but Massachusetts was not typical of the nation in the nineteenth century. It was more urban and industrial, was populated with a larger number of immigrants, and had lower fertility rates. The Massachusetts data reach reasonable quality by about 1860 and are reproduced here in Tables Ab928 and Ab1048–1058. Recent work with the genealogical data has concluded that adult mortality was relatively stable after about 1800 and then rose in the 1840s and 1850s before commencing long and slow improvement after the Civil War. This finding is surprising because we have evidence of rising real income per capita and of significant economic growth during the 1840–1860 period. However, income distribution may have worsened, and urbanization and immigration may have had more deleterious effects than hitherto believed. Further, the disease environment may have shifted in an unfavorable direction (Fogel 1986; Pope 1992; Haines, Craig, and Weiss 2003). It

should be noted, however, that the genealogical records exist only for individuals whose ancestors compiled a family tree. They may be subject to biases that stem from inadequate records, systematic errors introduced when linking records, and the selectivity that determines whose family tree is thought worth researching and which families are not traced in this manner.

We have better information for the post–Civil War period. Rural mortality probably began its decline in the 1870s because of improvements in diet, nutrition, housing, and other quality-of-life aspects on the farm. There would have been little role for public health systems before the twentieth century in rural areas. Urban mortality probably did not begin to decline prior to 1880, but thereafter urban public health measures – especially construction of central water distribution systems to deliver pure water and sanitary sewers – were important in producing a rapid decline of infectious diseases and mortality in the cities that installed these improvements (Melosi 2000). There is no doubt that mortality declined dramatically in both rural and urban areas after about 1900 (Preston and Haines 1991).

Tables Ab644–703 provide data on the expectation of life at birth and the infant mortality rate (deaths in the first year of life per 1,000 live births) for the white population from 1850 onward. No information is given for the years prior to 1850 because of the difficulty of finding comprehensive, comparable, and reliable mortality estimates. Both the expectation of life at birth and the infant mortality rate show sustained improvement in mortality (that is, rising expectation of life or falling infant mortality or crude death rates) only from about the 1870s onward.

What is apparent is that serious fluctuations in mortality were less likely after the 1870s, and that this was integral in the process of the mortality transition. This also confirms one unusual aspect of the American demographic transition: the decline of fertility commenced substantially earlier than that of mortality. Although levels of mortality in the United States in the middle of the nineteenth century were comparable with those in western and northern Europe, significant mortality fluctuations were still occurring right up to the twentieth century. Consistent control of mortality in terms of both sustained decline and damping of mortality peaks started only after the 1870s.

The evidence of rising mortality in the antebellum period from genealogical data is bolstered by data on the stature of adults. Males were shorter in the birth cohorts from about 1830 to about 1870 than those who were born earlier or later (Steckel 1992b; Haines 1998b). This evidence strongly suggests a poorer disease environment in that period as well as deteriorating nutrition. Following the Civil War, the disease environment improved in what might be called an "epidemiological transition." A variety of factors can affect mortality; these may be conveniently grouped into ecobiological, public health, medical, and socioeconomic categories. These divisions are not mutually exclusive because, for example, economic growth can make resources available for public health projects, and advances in medical science can inform the effectiveness of public health. Ecobiological factors, those that change the virulence of a disease course or its transmission mechanism, were not likely to have been significant. The remaining factors – socioeconomic, medical, and public health – are often difficult to disentangle. For example, if the germ theory of disease (a medical/scientific advance of the later nineteenth century) contributed to better techniques of water filtration and purification in public health projects, then how should the roles of medicine versus public health be apportioned? Thomas McKeown (1976) has proposed that, prior to the twentieth century, medical science contributed little to reduced mortality in Europe and elsewhere. This conclusion was based particularly on the experience of England and Wales, where much of the mortality decline between the 1840s and the 1930s was attributable to reductions in deaths from respiratory tuberculosis, other respiratory infections (for example, bronchitis), and nonspecific gastrointestinal diseases (for example, diarrhea, gastroenteritis), when as yet no effective medical therapies were available for these infections until well into the twentieth century. If ecobiological, medical, and public health factors are eliminated, the mortality decline must have been due to socioeconomic factors, especially in McKeown's view – better diet and nutrition, and improved clothing and shelter.

It is true that medical science did have a rather limited direct role before the twentieth century. In terms of specific therapies, smallpox vaccination was known by the late eighteenth century, and diphtheria and tetanus antitoxin and rabies therapy by the 1890s. The germ theory of disease, advanced by Pasteur in the 1860s and greatly advanced by the work of Koch and others in the 1870s and 1880s, was only slowly accepted by what was at the time a very conservative medical profession. Even after Robert Koch conclusively identified the tuberculosis bacillus and the cholera vibrio in 1882 and 1883, various theories of miasmas and anticontagionist views were common among physicians in the United States and elsewhere. Hospitals, having originated as pesthouses and almshouses, were (correctly) perceived as generally unhealthy places to be. Surgery was also very dangerous before the advances made by William Halsted in the 1880s and 1890s. Major thoracic surgery was rarely risked, and if attempted, patients had a high probability of dying from infection, shock, or both. The best practice in amputations was to do them quickly to minimize risks. Although anesthesia had been introduced in America in the 1840s and the use of antisepsis in the operating theater had been advocated by the British surgeon Joseph Lister in the 1860s, surgery cannot be considered as reasonably safe until the twentieth century.

Although the direct impact of medicine on mortality in the United States over this period is questionable, public health did play an important role and thereby indirectly allowed medicine a part. After John Snow identified a polluted water source as the origin of a cholera outbreak in London in 1854, pure water and sewage disposal became important issues for municipal authorities. New York City constructed its forty-mile-long Croton Aqueduct in 1844, and Boston was also tapping various outside water sources by aqueduct before the Civil War. Chicago, which drew on Lake Michigan for its water, also had to cope with sewage disposal directly into its water supply from the Chicago River. At an early date, all buildings in the entire downtown area were raised by one story to facilitate gravity sewage flow. Water intakes were moved farther offshore in the 1860s, requiring tunnels several miles long driven through solid rock. But this was only a temporary solution. Finally, the city reversed the flow of the Chicago River, using locks, and sent the effluent down to the Illinois River and away from the water intakes in Lake Michigan. The project took eight years (1892–1900) and was called one of the "engineering wonders of the modern world." The bond issue to fund it was approved by an overwhelming vote of 70,958 to 242.

A pattern was emerging in the late nineteenth century – massive public works projects in larger metropolitan areas to provide clean water and proper sewage disposal. But progress was uneven.

Baltimore and New Orleans, for example, were rather late in constructing adequate sanitary sewage systems. Filtration and chlorination were added to remove or neutralize particulate matter and microorganisms as a consequence of advances in the new science of bacteriology. Public health officials were often more eager to embrace the findings from bacteriology than were physicians, who sometimes saw public health officials as a professional threat. Most public works and public health projects were locally funded, with the consequence of uneven and intermittent progress toward water and sewer systems, public health departments, and so forth. Indeed, one reason for the better mortality showing of the ten largest cities in 1900 as compared with remaining cities with populations greater than 25,000 was the capacity of the largest cities to secure the necessary financial resources (Preston and Haines 1991; Melosi 2000; see also Chapter Dh, on services and utilities).

Progress in public health was not confined to water and sewer systems, though they were among the most effective weapons in the fight to prolong and enhance human life. Other areas of public health activity from the late nineteenth century onward included vaccination against smallpox; use of diphtheria and tetanus antitoxins (from the 1890s); more extensive use of quarantine (as more diseases were identified as contagious); cleaning urban streets and public areas; physical examinations for schoolchildren; health education; improved child labor and workplace health and safety laws; legislation and enforcement efforts to reduce food adulteration and especially to obtain pure milk; measures to eliminate ineffective or dangerous medications (the Pure Food and Drug Act of 1906); increased knowledge of and education concerning nutrition; stricter licensing of physicians, nurses, and midwives; more rigorous medical education; building codes to improve heat, plumbing, and ventilation in housing; measures to alleviate air pollution in urban settings; and the creation of state and local boards of health to oversee and administer these programs.

Public health measures proceeded on a broad front, but not without delays and considerable unevenness in enforcement and effectiveness. The issue of milk purity is a case in point. It became apparent that pasteurization (heating the milk to a temperature below boiling for a period of time), known since the 1860s, was the only effective means of ensuring a bacteria-free product. Pasteurization was, however, resisted by milk sellers, and it came into practice only relatively late. In 1911, only 15 percent of the milk in New York City was pasteurized. In 1908 only 20 percent of Chicago's milk was so treated. Pasteurization did not become compulsory in Chicago until 1908 and in New York City until 1912.

Long-Term Changes in the Causes of Death

Reductions in death from infectious and parasitic diseases, both of the respiratory (usually air-borne) and gastrointestinal (usually water-borne) types, explain much of the decrease in mortality in the late nineteenth and early twentieth centuries. In a study of Philadelphia over the period 1870–1930, about two thirds of the decline in death rates came from reductions in deaths caused by various infectious diseases. Most impressive is a 22 percent decline in deaths from respiratory tuberculosis. Among children there were significant reductions in mortality from diphtheria and croup, scarlet fever, smallpox, and respiratory tuberculosis. Diphtheria antitoxin, water filtration, and quarantine helped, but an improved standard of living was also important, especially in controlling tuberculosis,

for which no specific therapy was available until the 1940s. Well-nourished individuals were better able to resist tuberculosis. As the population became increasingly well fed, the disease became a less frequent cause of death.

Reliable cause-of-death information for large areas of the nation became available in 1900 with the initiation of the Death Registration Area. The crude death rate declined (at least for the Death Registration Area) by 25 percent between 1900 and 1920 (see Tables Ab929–987). Of this decline, 70 percent was accounted for by a reduction in death from infectious and parasitic diseases. Of that reduction from infectious disease, 24 percent was attributable to reductions in mortality from respiratory tuberculosis. Over the longer period of 1900–1960, the crude death rate declined by 45 percent, while mortality from all infectious and parasitic diseases was reduced by 90 percent. Deaths from infectious and parasitic diseases declined from 57 percent of all deaths to only 8.8 percent. In addition, diagnosis and reporting were becoming more accurate, as deaths in the category "other and unknown" declined from 21 percent of all deaths to 9.3 percent. The decline in mortality from infectious disease actually exceeded that from all causes combined because mortality from chronic, degenerative diseases (cancer, cardiovascular disease) increased.

One of the great events in human history has been the prolongation of life and reduction in mortality in the modern era, attributable chiefly to great declines in death from epidemic and endemic infectious disease. Americans and most inhabitants of the developed world no longer live with the kind of fear and fatalism that characterized a world in which sudden and pervasive death from disease was a fact of life. For the United States, most of this improvement took place in the twentieth century.

Mortality Differentials

During the nineteenth century, both prior to and during the mortality transition commencing in the 1870s, significant differentials in mortality existed – by sex, rural–urban residence, race, region, origin (native versus foreign-born), and socioeconomic status. Male mortality usually exceeds female mortality at all ages. This was generally true in the United States in the nineteenth century. The relative differences were often smaller than those in the mid to late twentieth century, as a consequence of the hazards of childbearing, pervasive exposure to disease-causing organisms, and the effects of differential health behaviors. Sex differentials in mortality increased in the twentieth century as sex differentials in smoking rose.

It is clear that, before about 1920, urban mortality was much in excess of rural mortality and, in general, the larger the city, the higher the death rate (Haines 2001). A variety of circumstances contributed to the greater mortality in cities: greater density and crowding, leading to the more rapid spread of infection; a higher likelihood of contaminated air, water, and food; garbage, horse droppings, and carrion in streets; larger inflows of foreign migrants, both new foci of infection and new victims; and migrants from the countryside who had not been exposed to the harsher urban disease environment. Cities were the home of the poor.

According to the Death Registration Area life tables for 1900–1902, the expectation of life at birth was 48.2 years for white males overall – 44 years in urban areas and 54 years in rural places. The comparable results for females were similar: 51.1 years overall,

48 years in urban areas, and 55 years in rural areas. For the seven states with reasonable registration data in both 1890 and 1900, the ratio of urban to rural crude death rates reported in the 1890 Census was 1.27, and that in 1900 was 1.18. For young children (ages 1–4 years) the ratios were much higher, with urban mortality being 107 percent higher in 1890 and 97 percent higher in 1900. For infants the excess urban mortality was 63 percent in 1890 and 49 percent in 1900. Residence in cities – with poorer water quality, lack of refrigeration to keep food and milk fresh, and close proximity to a variety of pathogens – was very hazardous to the youngest inhabitants.

The higher urban mortality rates began to diminish after the turn of the century, especially as public health measures and improved diet, shelter, and general living standards took effect. The excess in expectation of life at birth for rural white males over that for urban white males was 10 years in 1900. This fell to 7.7 years in 1910, 5.4 years in 1930, and 2.6 years by 1940.

The black population of the United States certainly experienced higher death rates, both as slaves and then as a free population, than did whites. Table Ab1–10 provides a breakdown of the expectation of life at birth and the infant mortality rate by race (see Figure Ab-C). For the 1890s, based on estimates using the 1900 Census public use sample, the infant mortality rate was 111 infant deaths per 1,000 live births for the white population and 170 for the black population. The implied expectations of life at birth were 51.8 years for whites and 41.8 years for blacks (see Tables Ab644–911 and Ab952–987). The differential clearly had not disappeared by 1920, when the absolute difference in expectation of life at birth by race was 10.4 years, and the infant mortality rate for blacks was 60 percent higher than that for whites, despite the fact that blacks still lived in predominantly rural areas. Even in 1980, although some convergence had occurred, the difference in life expectancy was 6.3 years, and black infant mortality was 90 percent higher than white. The absolute difference had narrowed, but the relative
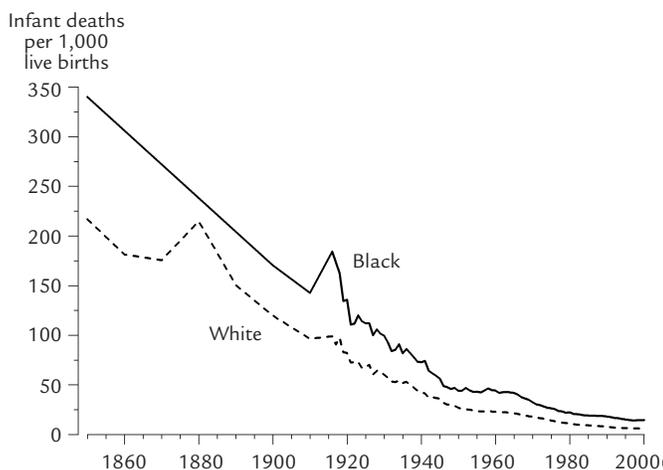
difference in infant survival had actually worsened. The disadvantaged status of the black population is apparent, as mortality is a sensitive indicator of socioeconomic well-being.

The mortality and health of the antebellum slave population has more recently been studied using plantation records and coastal shipping manifests that gave the heights of transported slaves. It has revealed very high mortality and very stunted stature among slave infants and young children, pointing to poor health conditions. For example, the infant mortality rate for slaves is estimated to have been as high as 350 infant deaths per 1,000 live births, in comparison to 197 for the entire American population in 1860. Death rates among slave children ages 1–4 years were also very high. A hypothesis for the high mortality and short stature of slave children is that they were not given much animal protein in their diets until about the age of ten years. In addition, pregnant and lactating women were often kept hard at field work, leading to lower birth weights, less breast-feeding, and early weaning (Steckel 1986a, 1986b).

Information on mortality differences between the native- and the foreign-born populations is ambiguous. In Massachusetts, for example, the crude death rate for the native population was higher (20.4 per 1,000 population) than that for the foreign-born (17.4) for the period 1888–1895. This difference disappears, however, once the results are adjusted for the younger age structure of the immigrant population. Using census samples to estimate the mortality of children of native- and foreign-born parents reveals the opposite: for seven New York counties in 1865, the probability of dying before age five years was 0.19 for children of native-born parents but 0.23 for children of foreign-born parents. The same calculation using the national sample of the 1900 Census gives a probability of death before age five years of 0.166 when both parents were native-born and 0.217 when both parents were immigrants. For the Death Registration Area life tables of 1900–1902, life expectancies at age 10 years were rather similar by origin: 51.6 years for native white males and 49.1 years for foreign-born white males. The results for 1909–1911 were 51.9 and 50.3 years, respectively. Differentials by origin were converging and had largely disappeared by the 1930s because the higher mortality of the foreign-born was largely attributable to lower socioeconomic status and a greater proportion living in large cities. As socioeconomic attainment narrowed between the groups and as the rural–urban mortality difference disappeared, the mortality penalty paid by the foreign-born also diminished.

Regional differences in mortality before the twentieth century are rather difficult to establish because of the incompleteness of geographic coverage of both vital statistics and local studies. In colonial times, especially in the seventeenth century, New England was the area with lowest mortality, while the region from the Chesapeake to the south had higher mortality. These differentials diminished in the eighteenth century, but the pattern continued into the first half of the nineteenth century, as is confirmed from estimates of adult mortality from genealogies for cohorts born in the late eighteenth and early nineteenth centuries. The Midwest also appeared as a relatively healthy region. For cohorts born in the middle of the century, however, these regional differences had dissipated. Indeed, the highest life expectation at age twenty years for white females born in the 1850s and 1860s was in the South Atlantic states. Regional differences, such as they were, converged into the twentieth century, but as late as 1950 the region with the lowest mortality was still the western portion of the Midwest, while the

FIGURE Ab-C   Infant mortality rate, by race: 1850–2000

Sources
Series Ab921 and Ab923.

Documentation
For 1850–1910, the data are for ten-year intervals for the white population. For the black population, there are no data points between 1850 and 1900, and for 1900 and 1910, the series is for ten-year intervals. Beginning in 1915, both series represent annual data. For 1915–1932, the series are for the official Birth Registration Area (see Tables Ab31–39 for the composition of the Birth Registration Area).

highest death rates were found in the Mountain states. Regional areas of poverty (for example, West Virginia, New Mexico) have led to significant variation across states.

Differences in survival probabilities also existed across socioeconomic groups, although here too the information is sketchy. Estimates of child mortality according to the occupation and socioeconomic status of the father from the 1900 and 1910 U.S. Census public use samples indicate that children of white-collar workers, professionals, proprietors, and farmers did better than average, while children of laborers, including agricultural laborers, had worse than average survival chances. The advantage to professionals, such as physicians, teachers, and clergy, was not great in 1900 but was becoming more pronounced by 1910. Data from registered births in the 1920s classified by occupation of father point to a widening of these socioeconomic status differentials as the century progressed. For instance, in 1900 children of laborers had mortality about 1.2–1.3 times greater than children of fathers with professional and technical occupations. This ratio had risen to about 1.4 in 1910 and over 2.0 by 1929.

Higher income and better educated groups more easily assimilated advice and improvements in child care, hygiene, and health practices and so were "leaders" in the American mortality decline of the early twentieth century. Public health improvements led to a reduction in the level of mortality but did not lead to a reduction in relative differentials across class and occupation groups. Rural–urban differences did converge into the early twentieth century, but both relative and absolute mortality differences by race did not. The role of personal and household health behavior has been inadequately emphasized in the debate on the origins of the mortality transition. It was very likely central, although the precise contribution to differential child mortality is not easy to assess. For adults, the mortality gradient observed at the turn of the century from high mortality among laborers to intermediate levels among skilled manual workers to the most favorable mortality among white-collar workers persisted up to the middle of the twentieth century. There is some evidence from earlier in the nineteenth century that socioeconomic variables, such as wealth or income, occupation, and literacy, were less important in predicting mortality differentials. For the 1850s, for instance, survival probabilities differed little between the children of the poor and the wealthy. Rural versus urban residence and region made more of a difference.

Overall, the mortality transition in the United States was a delayed event. Perhaps this is not surprising. Given the relatively low mortality in the United States in the late eighteenth century and given that the public health system was not yet capable of coping with the growing problems of urban life and contagion from population mobility, it would have been difficult to have had significant declines in mortality as soon as the early nineteenth century. Indeed, an increase in mortality was evident prior to the Civil War. The sustained decline commenced nationally only in the 1870s. A damping of year-to-year mortality fluctuations also took place after midcentury. It is not easy to assign credit to various causal factors in the mortality transition, but the principal proximate cause was the control of both epidemic and endemic infectious diseases. By the later nineteenth century, public health certainly contributed much, with improvements in diet, housing, and standard of living also being significant. The direct role of medical intervention was rather limited before the twentieth century but then increased as the germ theory of disease was accepted and better diagnosis and effective therapies were developed. Though difficult to assess,

changes in personal health behavior must be assigned importance, particularly after the turn of the twentieth century.

References

Caldwell, John C. 1982. *Theory of Fertility Decline*. Academic Press.

David, Paul, and Warren Sanderson. 1986. "Rudimentary Contraceptive Methods and the American Transition to Marital Fertility Control, 1855–1915." In Stanley L. Engerman and Robert E. Gallman, editors. *Long-Term Factors in American Economic Growth*. University of Chicago Press.

David, Paul, and Warren Sanderson. 1987. "The Emergence of a Two-Child Norm among American Birth Controllers." *Population and Development Review* 13 (1): 1–41.

Easterlin, Richard A. 1980. *Birth and Fortune: The Impact of Numbers on Personal Welfare*. Basic Books.

Fogel, Robert W. 1986. "Nutrition and the Decline in Mortality since 1700: Some Preliminary Findings." In Stanley L. Engerman and Robert E. Gallman, editors. *Long-Term Factors in American Economic Growth*. University of Chicago Press.

Forster, Colin, and G. S. L. Tucker. 1972. *Economic Opportunity and White American Fertility Ratios, 1800–1860*. Yale University Press.

Haines, Michael R. 1990. "Western Fertility in Mid-Transition: A Comparison of the United States and Selected Nations at the Turn of the Century." *Journal of Family History* 15 (1): 21–46.

Haines, Michael R. 1993. "Occupation and Social Class during Fertility Decline: Historical Perspectives." In John Gillis, David Levine, and Louis Tilly, editors. *The European Experience of Declining Fertility: 1850–1970*. Basil Blackwell.

Haines, Michael R. 1996. "Long Term Marriage Patterns in the United States from Colonial Times to the Present." *The History of the Family: An International Quarterly* 1 (1): 15–39.

Haines, Michael R. 1998a. "Estimated Life Tables for the United States, 1850–1910." *Historical Methods* 31 (4): 149–69.

Haines, Michael R. 1998b. "Health, Height, Nutrition, and Mortality: Evidence on the 'Antebellum Puzzle' from Union Army Recruits for New York State and the United States." In John Komlos and Joerg Baten, editors. *The Biological Standard of Living in Comparative Perspective*. Franz Steiner Verlag.

Haines, Michael R. 2001. "The Urban Mortality Transition in the United States, 1800 to 1940." *Annales de démographie historique* 1: 33–64.

Haines, Michael R., Lee A. Craig, and Thomas Weiss. 2003. "The Short and the Dead: Nutrition, Mortality, and the 'Antebellum Puzzle' in the United States." *Journal of Economic History* 63 (2): 385–416.

Haines, Michael R., and Samuel H. Preston. 1997. "The Use of the Census to Estimate Childhood Mortality: Comparisons from the 1900 and 1910 Public Use Samples." *Historical Methods* 30 (2): 77–96.

Haines, Michael R., and Richard H. Steckel, editors. 2000. *A Population History of North America*. Cambridge University Press.

McKeown, Thomas. 1976. *The Modern Rise of Population*. Academic Press.

Melosi, Martin V. 2000. *The Sanitary City: Urban Infrastructure in America from Colonial Times to the Present*. Johns Hopkins University Press.

Notestein, Frank W. 1953. "The Economics of Population and Food Supplies. I. The Economic Problems of Population Change." *Proceedings of the Eighth International Conference of Agricultural Economists*. Oxford University Press.

Pope, Clayne L. 1992. "Adult Mortality in America before 1900: A View from Family Histories." In Claudia Goldin and Hugh Rockoff, editors. *Strategic Factors in Nineteenth Century American Economic History: A Volume to Honor Robert W. Fogel*. University of Chicago Press.

Preston, Samuel H. 1970. "Older Male Mortality and Cigarette Smoking: A Demographic Analysis." Population Monograph No. 7. Institute of International Studies, University of California at Berkeley.

Preston, Samuel H., and Michael R. Haines. 1991. *Fatal Years: Child Mortality in Late Nineteenth Century America*. Princeton University Press.

Sanderson, Warren C. 1979. "Quantitative Aspects of Marriage, Fertility and Family Limitation in Nineteenth Century America: Another Application of the Coale Specifications." *Demography* 16 (3): 339–58.

Smith, Daniel Scott. 1996. "'The Number and Quality of Children': Education and Marital Fertility in Early Twentieth-Century Iowa." *Journal of Social History* 30 (2): 367–92.

Steckel, Richard. 1982. "The Fertility of American Slaves." *Research in Economic History* 7: 239–86.

Steckel, Richard H. 1986a. "A Peculiar Population: The Nutrition, Health, and Mortality of American Slaves from Childhood to Maturity." *Journal of Economic History* 46 (3): 721–41.

Steckel, Richard H. 1986b. "A Dreadful Childhood: Excess Mortality of American Slaves." *Social Science History* 10 (4): 427–65.

Steckel, Richard H. 1992a. "The Fertility Transition in the United States: Tests of Alternative Hypotheses." In Claudia Goldin and Hugh Rockoff, editors. *Strategic Factors in Nineteenth Century American Economic History*. University of Chicago Press.

Steckel, Richard H. 1992b. "Stature and Living Standards in the United States." In Robert E. Gallman and John Joseph Wallis, editors. *American Economic Growth and Standards of Living before the Civil War*. University of Chicago Press.

Sundstrom, William A., and Paul A. David. 1988. "Old-Age Security Motives, Labor Markets, and Farm Family Fertility in Antebellum America." *Explorations in Economic History* 25 (2): 164–97.

Sutch, Richard. 1975. "The Breeding of Slaves for Sale and the Westward Expansion of Slavery, 1850–1860." In Stanley L. Engerman and Eugene D. Genovese, editors. *Race and Slavery in the Western Hemisphere: Quantitative Studies*. Princeton University Press.

Sutch, Richard. 1990. "All Things Reconsidered: The Life-Cycle Perspective and the Third Task of Economic History." *Journal of Economic History* 51 (June): 1–18.

U.S. Bureau of the Census. 1975. *Historical Statistics of the United States from Colonial Times to 1970*. U.S. Government Printing Office.

Yasuba, Yasukichi. 1962. *Birth Rates of the White Population of the United States, 1800–1860: An Economic Analysis*. Johns Hopkins University Press.