Volume Title: Health and Labor Force Participation over the Life Cycle: Evidence from the Past

Volume Author/Editor: Dora L. Costa, editor

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-11618-2

Volume URL: http://www.nber.org/books/cost03-1

Conference Date: February 2-3, 2001

Publication Date: January 2003

Title: The Rich and the Dead. Socioeconomic Status and Mortality in the United States, 1850–1860

Author: Joseph P. Ferrie

URL: http://www.nber.org/chapters/c9627

2

The Rich and the Dead
Socioeconomic Status and Mortality in the United States, 1850–1860

Joseph P. Ferrie

2.1 Introduction

Research on the link between socioeconomic status and mortality in the late-twentieth-century United States has demonstrated that those lower in status die at earlier ages and suffer from more sickness and disease throughout their lives (Williams 1998; Lantz et al. 1998). Although a great deal of attention has now been devoted to explaining why those lower in status have worse outcomes, and the possibility that the causal link between health and status runs in both directions (with poor health leading to low status), such investigations lack a long-run perspective (Smith 1999). For example, although wide disparities in mortality rates by status were observed as early as the 1960s (Kitagawa and Hauser 1973), we do not know whether the disparities observed over the last four decades are large or small by historical standards.

Perhaps these disparities are merely the continuation of poor outcomes for poor people that generations have failed to erase—the result of poor nutrition, inadequate housing, or harsh working conditions. Or perhaps disparate health outcomes by status are a product of developments in medicine and technology in the late twentieth century that have given a new advantage to those with the incomes to purchase them. Knowing how health out-

Joseph P. Ferrie is associate professor of economics and faculty fellow at the Institute for Policy Research, both at Northwestern University, and a research associate of the National Bureau of Economic Research.

The author thanks participants at Northwestern University's Economic History Workshop and Institute for Policy Research faculty seminar, the 2001 American Economic Association meetings, the NBER conference participants, (particularly discussants Richard Steckel and Irwin Rosenberg), and Deirdre McCloskey and Stan Engerman for extremely helpful comments. Financial support was provided by the National Science Foundation.

comes differed by economic status in an earlier era (e.g., at a time when medical knowledge was rudimentary at best) can help distinguish between these explanations.

This study introduces new evidence on the individual-level correlates of mortality, particularly socioeconomic status measured by occupation and family wealth, created by merging the mortality and population schedules of the 1850 and 1860 federal population censuses. The experiences of several populations that have been overlooked in previous analyses of mortality in the middle of the nineteenth century are explored. For example, although the mortality of young children has been studied, it has been impossible to examine the mortality of older children and most young adults at the individual level. Although studies of the mortality of Union Army veterans have provided insights into the mortality of older adults, this work has of necessity ignored the experiences of Southerners, women, and children.

2.2 What We Know about Nineteenth-Century Socioeconomic Status and Mortality

There is a consensus today that low status is associated with increased risk for a variety of diseases, as well as a substantially increased risk of premature mortality. Attention has now largely turned to discovering the mechanisms that produce these disparate outcomes. An understanding of the long-run progress made in narrowing disparities in health outcomes by status, however, has been more difficult to attain. There are few sources of data on mortality with information on status available before the Second World War. In fact, no nationally uniform system of reporting deaths was in place until the completion of the Death Registration Area in 1933. Before that time, those interested in the link between status and mortality were forced to rely on data less representative of the national experience. Three published studies and one ongoing research project have attempted to assess the link between status and mortality for the second half of the nineteenth century.

The first of these estimated crude death rates of taxpayers and nontaxpayers for 1865 in Providence, Rhode Island (Chapin 1924). The annual crude death rate for taxpayers was 11 per thousand, while the corresponding rate for nontaxpayers was 25 per thousand. Although this suggests a substantial gap in crude death rates by status, it is less than satisfying in a number of respects. The first is the year examined: 1865 was the last year of the U.S. Civil War. Given the disruptions to commerce, industry, and agriculture, as well as the large number of Rhode Island's inhabitants who enlisted, this is unlikely to have been a year representative of the mid-nineteenth-century mortality experience. The second difficulty is the narrow geographic coverage of the study: It examines a significant urban center, but in 1860 only 21 percent of the U.S. population lived in places of 2,500 or more inhabitants. An additional shortcoming is that the study is unable to distinguish among different causes of death, although we know today that not all causes are equally susceptible to the influence of status. Finally, the experience of a single city for a single year tells us little about trends in the link between status and mortality over the late nineteenth century; data from several years are necessary to establish a pattern of increase or decline in the relationship between status and mortality.

The second study to examine the relationship between status and mortality for the late nineteenth century used data from the 1900 U.S. Census of Population, which for only the second time contained a question on "children ever born" (Preston and Haines 1991). The authors used this information, together with the composition of the household actually observed in the 1900 population schedules, to infer infant and child mortality for each household. There was no significant relationship between higher status and lower infant and child mortality, when status was measured by the occupation of the husband's occupation or imputed income (Preston and Haines, 154–56). Although there was higher mortality among those in households headed by unskilled laborers than among those in households headed by other workers, there were no substantial differences in mortality by occupation among households headed by individuals who were not unskilled laborers. They did find, however, that property ownership was associated with lower infant and child mortality than renting (Preston and Haines, 157–58).

Although this study is useful for its broad geographic coverage and the representativeness of the population it examines, it also has some important limitations. The first is the inability to assess the mortality experience of adults: Mortality was inferred from the question on "children ever born" and the observed household composition in 1900, so it was not possible to say whether individuals at older ages who were absent from the households where their mothers were enumerated had died or simply moved out. This study is also somewhat limited in the components of socioeconomic status that it can examine: Although the household head's occupation was recorded, there was no information collected in the 1900 census on the value of the household's wealth.[1] Such information was included in the 1850–70 population censuses, and can thus be used in the sample that will be constructed in the present project. Another difficulty with the Preston and Haines (1991) study is that, like the 1865 Providence, Rhode Island, study,

1. Although the census asked whether the family's residence was owned or rented, it did not inquire as to the value of the property, or the value of any other assets held by the family. If there are differences in the impact of different types of wealth on mortality, even the data on home ownership would then present an incomplete picture of the link between the family's socioeconomic status and its mortality experience.

it provides information at only one date (1900). Although deaths that occurred prior to 1900 can be inferred, it is impossible to say much about deaths that occurred much prior to 1885, nor to say with much precision when the deaths that can be inferred actually occurred. This may substantially attenuate any underlying link between observed household socioeconomic status (measured in 1900) and the household's infant and child mortality experience over the preceding years. It is also impossible with these data to examine causes of death and uncover links between status and specific mortality risks.

Finally, one study has examined the link between status and mortality with a sample that covers the entire United States and includes the information on wealth provided in the 1850 and 1860 federal population censuses (Steckel 1988). The project used 1,600 households linked from the 1850 census population schedules to the 1860 population schedules. Mortality within the household was inferred by comparing the household's composition in 1850 and in 1860. Like Preston and Haines (1991), Steckel found no relationship between status (measured by real estate wealth, literacy, and father's occupation) and infant and child mortality. Like the other studies described above, however, this project was unable to disaggregate by cause of death and provides information on status and mortality at but a single point in time.

The University of Chicago's Center for Population Economics (CPE) is using information from Union Army pension records to assess the link between socioeconomic status (among other factors) and later disability and premature mortality. Although this work is able to provide tremendously detailed information on diseases and causes of death as documented by health science professionals, it covers a relatively narrow population: veterans of the Union Army who survived late enough into the nineteenth century to obtain a federal pension. It says nothing about mortality among infants, children, women, or younger men. Furthermore, it is limited to the Northern population. The present study complements this work: Although the data on causes of death are less precise, they cover the populations and regions missed in the Union Army veterans project.

A recent unpublished study (Haines, Craig, and Weiss 2000) examined county-level crude death rates for 1850 (calculated from the mortality schedules used here) and found that wealthier counties actually had higher crude death rates. The authors conclude that this surprising finding "is consistent with the view that wealthier areas were those with more urbanization and greater levels of commercialization and better transport connections" (Haines, Weiss, and Craig, 8). Although their methodology makes it possible to say how aggregate wealth in a county affected aggregate mortality levels, their findings cannot tell us how status at the individual level affected individual-level mortality—and it is at the individual level that the link between status and mortality is probably strongest, if it exists.

Table 2.1 **Sample Records from Mortality Schedules (1850) for Perry County, Illinois**

Name	Age	Occupation	Sex	Month of Death	Cause of Death	Birthplace
Cunningham, Margaret J.	25	None	F	December	Fever	South Carolina
Curlee, James	22	Student	M	February	Fever	Tennessee
Dermon, Jane	60	None	F	May	Bowel inflammation	Ireland
Dunn, James	33	Farmer	M	August	Fever	South Carolina

2.3 The Data

As part of the regular decennial federal censuses of 1850 through 1880, census marshals asked each household how many members had died in the twelve preceding months. Although published totals from these inquiries were included in the 1850 through 1880 census volumes (and these figures form the basis for many mid-nineteenth-century U.S. life tables; e.g., Haines 1998), the data have never been examined at the individual level.[2] Several difficulties have prevented their full exploitation.[3]

The greatest difficulty is the inaccessibility of the original manuscript schedules. After the census office's tabulations were completed, the schedules were returned to archives in the states where the data had been gathered. Records from a few states have not survived, some have not been microfilmed, and none were available in machine-readable form until recently. Entries for over 400,000 decedents from the 1850 through 1880 mortality schedules have now been either transcribed and published (Volkel 1972, 1979; Hahn 1983, 1987) or computerized (Jackson 1999). Table 2.1 shows several records from the 1850 mortality schedules from Perry County, Illinois, to illustrate the range of information available from this source.

There are four likely sources of bias in these data. The first is that, based on model life tables and the published totals, it appears that mortality at very young and very old ages is underreported, and that overall mortality is underestimated by as much as 40 percent. The second bias is that surviving households are probably more likely to report deaths that occurred closer in time to the date of the census enumeration. The third bias is the underenumeration of deaths in households where all members died and thus left no survivors to report their deaths to the census enumerator. The final bias results from the reporting of the cause of death by household members rather than by health care professionals. This no doubt leads to common mistakes (like reporting "typhus" when the cause of death was "typhoid"),

2. Among those who have made use of the published totals, in addition to Haines (1998) are Fogel and Engerman (1974, 101) to calculate slave death rates, and Jacobson (1957).
3. These difficulties are summarized in Condran and Crimmins (1979).

but can be remedied to some extent by grouping diseases into broad categories, reflecting either easily identified physical symptoms or the likely susceptibility to the influence of socioeconomic status. For the present study, which will examine mortality rates by comparing the mortality schedules to the population schedules for a set of identical counties, these biases are substantial problems only if underreporting or misreporting varies by status differently in the mortality and population schedules. If an undercount of deaths in low-status families results from such families' being missed entirely by the census, then both the survivors and decedents will be absent from the combined data, leaving the mortality rate unaffected.[4]

Although it is not possible to test whether reporting of the *number* of deaths varied by status, it is possible to assess whether the reported *timing* of the deaths that were reported varied by status. If low-status families were as likely as high-status families to report deaths more distant in time from the census date, we can have somewhat greater confidence in the reliability of the reporting of deaths by status.[5] After decedents from the 1860 mortality schedules were matched to their surviving families in the 1860 population schedules (as described below), the distribution of the months in which deaths occurred was calculated for high- (total wealth > 0) and low- (total wealth $= 0$) status families. Figure 2.1 shows that the distributions are

4. An example can help assess the possible magnitude of the bias from underenumeration (failure of an entire family to appear in the population schedules, and the lack of information reported by these families in the mortality schedules) or underreporting (failure of families reported in the population schedules to inform the census marshal that a death had occurred that should have been included in the mortality schedules) in the estimated effect of wealth on mortality. Imagine a population containing 100,000 individuals, half in families with zero wealth and half in families with positive wealth. The mortality rate among those in families with zero wealth is 30 per thousand; it is 10 per thousand among those in families with positive wealth. The possession of positive wealth reduces mortality by 0.020. Suppose now that 40 percent of those in families with zero wealth (both survivors and decedents) were missed by the census marshals (and none of those with positive wealth were missed). The difference in the mortality rate by wealth ownership for the remaining 80,000 observations is still 0.020. Suppose now that underenumeration was zero for both groups, but that 20 percent of the deaths in zero-wealth families were not reported. Wealth now appears to reduce mortality by 0.014. If the 20 percent underreporting rate was instead applied to the positive-wealth families, wealth appears to reduce mortality by 0.022. This suggests that (a) the failure of entire low-wealth families to appear in the census (in both the population and mortality schedules) leads to no bias in the effect of wealth on mortality; (b) the failure of low-wealth families to report deaths when the rest of the family was enumerated can bias the effect of wealth downward from its true value (by 30 percent in this case); and (c) the failure of high-wealth families to report deaths can bias the effect of wealth upward from its true value (by 10 percent in this case).

5. This abstracts from the possibility that the timing of deaths in the year prior to the census was systematically related to status (e.g., if poor nutrition or poor housing made deaths in the winter more likely for families of low status). If such was the case, we might observe months when low-status families reported a larger fraction of their deaths than high-status families. There would be no reason to expect, however, that the gap between the fractions reported by high- and low-status families would widen continuously as time from the census date increased if both high- and low-status families are able to remember accurately the months in which deaths occurred.

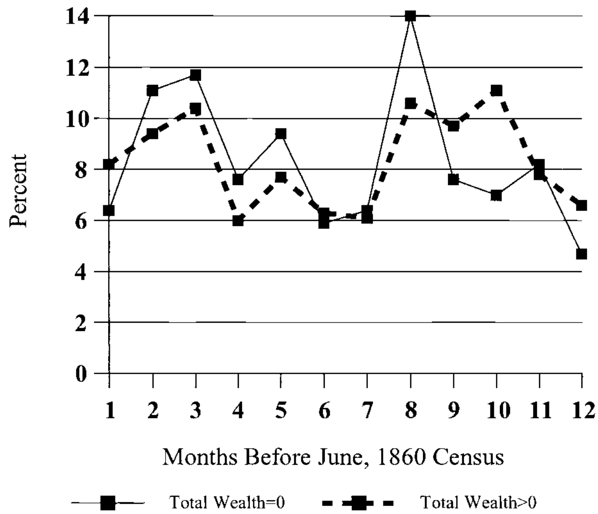

Fig. 2.1 Distribution of months of death by family total wealth in three Illinois counties and two Alabama counties, 1860

Notes: For "Total Wealth = 0," $N = 171$; for "Total Wealth > 0," $N = 587$. The chi-square statistic for the homogeneity of the two distributions is 7.5622 ($p = 0.8180$).

similar except at ten months prior to the census (August 1859). The overall distributions are statistically indistinguishable.

The advantages of using individual observations from the mortality schedules more than outweigh the shortcomings. For example, when combined with the information on each family's socioeconomic status in the population schedules, the mortality schedules provide the best and most broadly representative view we are likely to get of the socioeconomic correlates of mortality by cause of death. The range of places that can be examined makes it possible to assess the impact of a variety of environmental forces (such as climate and the presence of sanitation and public health systems) on the relationship between status and mortality.

By themselves, the data in the mortality schedules are an extremely valuable and heretofore unexploited source of information on the health of the nineteenth-century U.S. population. As table 2.1 shows, the mortality schedules themselves contain some information on status—each decedent's occupation at the time of death was reported. But a great deal more can be done after linking the mortality schedules to the population schedules collected at the same time. Table 2.2 shows the information relating to status than can be obtained from the 1850 and 1860 population schedules. Each piece of information is reported for each surviving member of the family.

Two data sets will be employed. In the first, individuals from a particular

Table 2.2 Variables in Population Schedules Related to Socioeconomic Status

1850	1860
Occupation	Occupation
Real estate wealth	Real estate wealth
	Personal wealth
Literacy	Literacy
School attendance	School attendance
Pauper	Pauper
Criminality	Criminality
Disability	Disability

county in the mortality schedules will be merged with individuals from the corresponding county in the population schedules. This will make possible an examination of the correlates of mortality at the individual level that controls for characteristics common to the two schedules (age, birthplace, occupation, and characteristics of the county). Since occupation will be the only measure of socioeconomic status available in this merged data set, attention will be confined to males over the age of twenty.[6] Fifty counties, shown in figure 2.2, were selected for which the 1850 population schedules have been entirely transcribed and for which decedents were included in the computerized mortality database.[7] The counties are concentrated in the Midwest, the upper South, and Alabama.[8] The linkage produced a sample of 927 adult male decedents and 82,246 adult male survivors in the fifty-county area.

6. An alternative strategy would be to combine decedents from all places in the computerized mortality schedules with individuals from the population schedules for the same places who appeared in the public-use sample of the 1850 census. The strategy employed here (focusing on places that have been completely transcribed) will allow for more detailed controls for location-specific effects at the level of minor civil divisions when, at a later date in the project, the manuscripts of the mortality schedules for these fifty counties are searched to determine the town or township of residence for the decedents. In recording deaths, census marshals often remarked upon the quality of the soil, the local climate, the prevalence of endemic diseases, and general economic conditions observed in the town or township. For example, James Searcy, the assistant census marshal who enumerated the southern division of Henry County, Alabama, in 1850, noted: "Bilious fevers or diseases are the most prevalent malady in my district caused by excessive drought and heat. Well watered and qualely light gray soil—generally oak, pine, hickory, cedar timber, lime and marl . . ." (Volkel 1972, 94).

7. The computerized mortality schedule transcriptions were obtained from Jackson (1999). The computerized population schedule transcriptions were obtained from the on-line archives of the USGenWeb project at http://www.usgenweb.org.

8. Alabama: Baldwin, Blount, Conecuh, Henry, Jackson, Jefferson, Lowndes, Madison, Marengo, Monroe, Shelby, Washington, and Wilcox; Illinois: Clark, Crawford, Gallatin, Grundy, Hamilton, McDonough, Perry, Saline, Sangamon, Schuyler, Scott, Stark, Washington, and Wayne; Indiana: Boone, Fayette, Kosciusko, and White; Iowa: Appanoose and Cedar; Kentucky: Simpson and Spencer; Michigan: Ionia and Lapeer; North Carolina: Northampton and Wake; Ohio: Henry, Pike, Sandusky, and Williams; Pennsylvania: Carbon, Sullivan, and Tioga; Texas: Galveston; Virginia: Charlotte, Fauquier, and Madison.

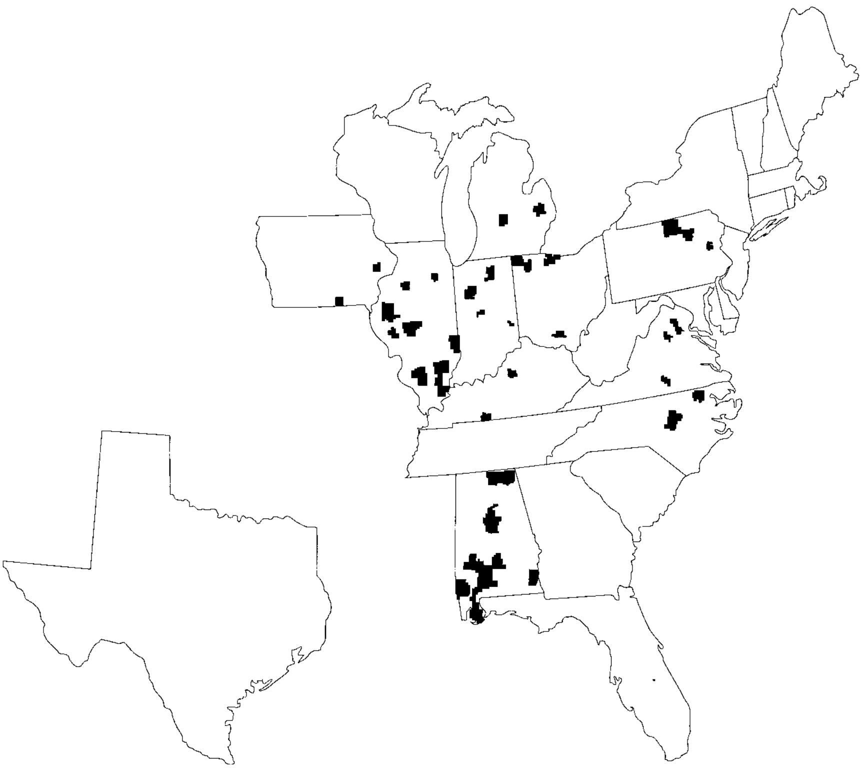

Fig. 2.2 Fifty counties used in 1850 analysis of mortality by occupation

For a smaller set of counties, it was possible to do better than this. A second data set was created for these places by merging decedents with the families in which they were living before their deaths. This was done by choosing locations for which the population and mortality schedules have been completely transcribed, and for which it was possible to sort both schedules in the order in which they were visited by the census marshal. Since the mortality schedule was filled out at the same time as the population schedule, it was possible to set the schedules side by side and locate the families in the population schedules who reported the deaths in the mortality schedules.[9]

9. It would have also been feasible here to use individuals from the mortality schedules and a sample of individuals from the public-use sample for 1850 or 1860. But because the entire population schedule had to be examined to locate the surviving households of decedents, the set of decedents had to be limited to those places with both mortality and population schedules sorted in their original order—simply examining a 1 percent sample from the population schedule would identify the households of very few decedents with certainty. This suggested using the completely transcribed population schedules for the set of survivors. This will also make possible the analysis of "neighborhood effects" later in the project. As was the case for

Table 2.3 shows an example of this linkage for part of Shelby County, Illinois in 1860. The mortality schedule listed the following individuals in order: Louisa Compton (age 38), Emma A. Compton (age 1), John W. Lanning (age 6), and Jane B. Graybill (infant). Using the names, ages, and birthplaces, these individuals were then inserted back into the population schedules in the families who reported their deaths.

For 1850, it was possible to merge decedents with their surviving families for five counties in Illinois (Morgan, Jackson, Union, Saline, and Washington) and one in Alabama (Shelby). For 1860, it was possible for three counties in Illinois (Perry, Shelby, and Vermilion) and two in Alabama (St. Clair and Tuscaloosa). These counties are shown in figures 2.3 and 2.4.[10] For each family member, the linked data contain the individual's age, sex, race, birthplace, family wealth (the sum of the wealth reported for each family member), the occupation of the family head, family size, whether the family member died in the twelve months before the census, and the cause of death for decedents. A family is defined here and throughout the analysis as a group of individuals with a common surname residing in the same household. A household is a group of individuals living in the same residence, regardless of their surnames.[11]

Table 2.4 shows the marginal effects from a probit regression in which the dependent variable is 1 if the individual was linked to the population schedules and zero otherwise. The overall linkage rate was 85 percent, but age reduced the probability of successful linkage, males were 3 to 4 percentage points more likely to be matched, and those born in the same state of their residence at death were 5 percentage points less likely to be linked. The effect of age reflects the increasing probability than an individual will be living away from the family as he or she becomes older: Unless the mother died in childbirth, an infant was survived by at least one person with the same surname; children at higher ages have a higher probability of having been orphaned and having taken up residence with another family; young adults may have left their families to work on nearby farms or in local businesses; and older adults are more likely to have seen their children move out and to have been a family's last surviving member. The negative effect of having been born locally probably reflects the difficulty in linking single persons

the linkage of adult males described above, this linkage could in theory be done for any locations with extant manuscripts of the population schedules and extant manuscripts of the mortality schedules. The counties chosen for analysis here were those with transcribed population schedules and transcribed mortality schedules, to limit the time spent in data transcription.

10. The transcribed mortality schedules were taken from Volkel (1972 and 1979) for Illinois and from Hahn (1983 and 1987) for Alabama. The computerized population schedule transcriptions were obtained from the on-line archives of the USGenWeb project at http://www.usgenweb.org.

11. For example, in table 2.3, Jane Gallagher and George Thompson in household number 590 are members of the same household as the three Kinsels, but Gallagher and Thompson are each counted as separate families in the analysis.

Table 2.3 **Example of Linked Population and Mortality Schedules, Shelby County, Illinois, Southern District (1860)**

Household Number	Name	Age	Sex	Occupation	Birthplace	Real Estate	Personal Estate	Cause of Death
589	Compton, Charles	45	M	Farmer	Virginia	$4,200	$1,880	
589[a]	Compton, Louisa	38	F		Ohio			Consumption
589	Compton, Jonathan	18	M	Laborer	Ohio			
589	Compton, Thomas	16	M	Laborer	Ohio			
589	Compton, Marion	14	M		Ohio			
589	Compton, Mary Jane	12	F		Ohio			
589	Compton, Charlie	10	M		Ohio			
589	Compton, Sarah E.	8	F		Illinois			
589	Compton, Louisa	5	F		Illinois			
589[a]	Compton, Emma A.	1	F		Illinois			Unknown
590	Kinsel, John	27	M	Farmer	Ohio		$1,195	
590	Kinsel, Cemantha	21	F		Ohio			
590	Kinsel, Simon J.	1	M		Illinois			
590	Gallagher, Jane	11	F		Illinois			
590	Thompson, George	24	M	Laborer	Ohio			
591	Lanning, William	39	M	Farmer	Pennsylvania	$1,200	$795	
591	Lanning, Mariah	39	F		Ohio			
591	Lanning, Delila J.	17	F		Ohio			
591	Lanning, Thomas M.	15	M	Laborer	Ohio			
591	Lanning, Derick D.	12	M		Ohio			
591[a]	Lanning, John W.	6	M		Ohio			Scarlet fever
591	Lanning, Daniel G.	1	M		Illinois			
591	Sinn, Rhoda	66	F		New Jersey			
592	Graybill, Samuel R.	37	M	Farmer	Ohio	$2,500	$1,300	

(*continued*)

Table 2.3 (continued)

Household Number	Name	Age	Sex	Occupation	Birthplace	Real Estate	Personal Estate	Cause of Death
592	Graybill, Sarah H.	36	F		Ohio			
592	Graybill, Thomas J.	13	M		Ohio			
592	Graybill, Isaac G.	12	M		Ohio			
592	Graybill, Henry C.	10	M		Ohio			
592	Graybill, Carlisle	9	M		Ohio			
592	Graybill, George	8	M		Ohio			
592	Graybill, Sarah Olive	3	F		Ohio			
592	Graybill, James B.	0	M		Illinois			
592	Graybill, Jane B.	0	F		Illinois			Brain congestion

[a]Entry inserted from mortality schedules.

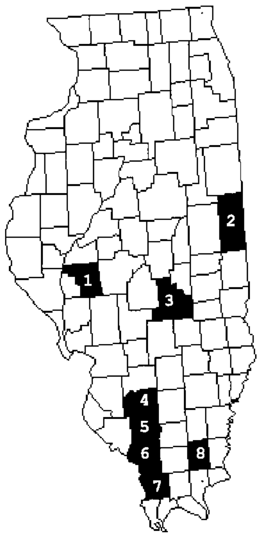

Fig. 2.3 Illinois counties in 1850 and 1860 analysis by wealth: (1) Morgan, (2) Vermilion, (3) Shelby, (4) Washington, (5) Perry, (6) Jackson, (7) Union, and (8) Saline

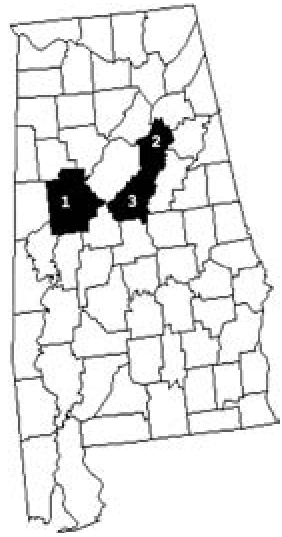

Fig. 2.4 Alabama counties in 1850 and 1860 analysis by wealth: (1) Tuscaloosa, (2) St. Clair, (3) Shelby

who resided with nonfamily members and whose deaths, although reported to the census marshal, would have left no persons in the population schedule with the same surname.

Three adjustments were made to the linked sample before it was analyzed. To reduce the influence of extreme outliers, 2,958 individuals in families with more than $10,000 in real estate wealth were deleted. Because individuals residing in households where their surnames were unique were not at risk to be successfully linked from the mortality schedules to the population schedules, although the deaths of such individuals would have been reported in the mortality schedules, the sample was further limited by discarding 1,382 additional individuals whose surnames appeared only once in the households where they were enumerated after the decedents were inserted into their surviving households.[12] Finally, because it was unclear how

12. If an individual was the only person with a surname in a household and that person died, his or her death would have been reported in the mortality schedule. But because there was no one else with the same surname in the household that reported the death, it would not be possible to identify which household reported the death when trying to link individuals from the mortality schedules back to the population schedules, so such deaths would remain unlinked. If a person with a surname unique within a household did not die, that person would appear in the population schedule. The linkage procedure would thus be biased toward survivors among those with surnames unique in their households of residence. To eliminate this bias, it is necessary to remove those with surnames unique in their households of residence from the

Table 2.4 Marginal Effects from Probit Regressions on Successful Linkage from Mortality to Population Schedules

Variable	(1)	(2)	(3)	(4)
Ages 1–4	–0.0827	–0.0824	–0.0827	–0.0820
	(2.21)**	(2.18)**	(2.22)**	(2.18)**
Ages 5–19	–0.1203	–0.1235	–0.1244	–0.1249
	(2.73)***	(2.75)***	(2.83)***	(2.79)***
Ages 20–44	–0.1531	–0.1633	–0.1511	–0.1595
	(3.43)***	(3.51)***	(3.41)***	(3.45)***
Ages 45+	–0.2253	–0.2339	–0.2248	–0.2309
	(3.78)***	(3.83)***	(3.80)***	(3.80)***
Male	0.0347	0.0366	0.0371	0.0388
	(1.63)	(1.71)*	(1.76)*	(1.83)*
Born in state of enumeration	–0.0464	–0.0449	–0.0493	–0.0480
	(1.54)	(1.48)	(1.66)*	(1.62)
Foreign-born	–0.0634	–0.0619	–0.0592	–0.0587
	(1.11)	(1.08)	(1.05)	(1.04)
Controls				
Year	Yes	Yes	No	No
County	Yes	Yes	Yes	Yes
Cause of death	No	Yes	No	Yes
Month of death	No	No	Yes	Yes
Predicted probability	0.8526	0.8529	0.8575	0.8580
Pseudo-R^2	0.0384	0.0399	0.0540	0.0566
Observations	1,193	1,193	1,193	1,193

Notes: Absolute value of z-statistics in parentheses. Omitted categories are "Infant," "Female," "Born outside state of enumeration," and "Native-born."

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

well household wealth would approximate the resources available to household members in group housing or in families with several boarders, 19,562 individuals in households of ten or more members were discarded. The sample that resulted from the linkage and from these adjustments contains 304 decedents and 38,996 survivors in 1850, and 511 decedents and 52,268 survivors in 1860.

In both the 1850 sample of adult male decedents merged with surviving adult males and the 1850 and 1860 samples of individual decedents merged with their surviving families, the counties for which the analysis can be per-

population schedules. In table 2.3, the individuals discarded were Jane Gallagher, George Thompson, and Rhoda Sinn. Although the fact that these individuals were living with families other than their own is potentially useful information on their economic status, it was thought prudent for now to keep analysis of their mortality experience separate from that of individuals living in multiple-person family units. For 1870 and 1880, such merging of decedents who had surnames unique in the household of residence back into the household of residence is feasible, as the mortality schedule reports the number of the households that reported each death.

formed were determined largely by whether genealogists had transcribed mortality and population schedules. These counties are uniformly rural. In 1850, only four places with populations over 3,000 are included: Mauch Chunk, Pennsylvania (pop. 5,203), Springfield, Illinois (pop. 4,533), Raleigh, North Carolina (pop. 4,518), and Galveston, Texas (pop. 4,177). In 1860, only two places with 2,000 or more inhabitants are included: Tuscaloosa, Alabama (pop. 3,989), and Elwood, Illinois (pop. 2,000). It was not possible to locate any counties in the Middle Atlantic or New England states for which linkage was possible.[13] Descriptive statistics for both samples as well as the most common causes of death are shown in tables 2A.1 through 2A.3 in the appendix to this chapter.

2.4 Analysis of Socioeconomic Status and Mortality among Adult Males in 1850

The first hypothesis to be tested is that in 1850 individuals in higher-income occupations had lower mortality rates than individuals in lower-income occupations. The exact mechanism through which this relationship operates will not be tested, but it seems reasonable to imagine that higher-status individuals may be able to purchase better nutrition (both more calories and a greater variety of calorie sources) and better housing (larger, better ventilated, farther from sanitary hazards, more thoroughly protected against rain and cold). The relationship between status and mortality will not be the same for all causes of death. It will be strongest for those causes of death most susceptible to living circumstances. Death from tuberculosis (best transmitted among individuals weakened by poor nutrition or exposure to other diseases and living in cramped, poorly ventilated places) will be more strongly associated with low status than death from drowning.

Occupations are grouped into four broad categories: white-collar (professional, managerial, clerical and sales, and government), craft, farmer, and laborer (including operatives and unskilled workers). Those with no reported occupation are listed as unknown. Farmers and white-collar workers had higher incomes than craft workers, who in turn had higher incomes than laborers, so if income differences are an important source of differences in mortality, laborers should have higher mortality than otherwise identical craftsmen, who should have higher mortality than otherwise identical white-collar workers.[14]

13. The 1860 mortality and population schedules for Albany, New York, have been linked by David Davenport, and the author has linked 1860 mortality and population schedules for several wards in Chicago. Results for these places will appear at a later point in the project.

14. Although we have no reliable estimates of farmers' income in 1850 or 1860, Margo (2000, 45) reports that in 1850, assuming twenty-six workdays per month, common laborers in the Midwest earned $20.80 per month, craftsmen earned $35.10 per month, and white-collar workers earned $47.12 per month. In the South Central region, common laborers earned $22.10 per month, craftsmen earned $47.06 per month, and white-collar workers earned $60.84 per month.

There are several likely influences on mortality that must be controlled for to isolate the role of status. The most important is obviously age. It is also possible to identify individuals born outside the state in which they resided at the time of the census. Migrants may have had lower mortality if they are more physically fit than nonmigrants, but their introduction to a new disease environment may have a countervailing effect on their mortality. There may also be differences in the physical or economic environment across counties or regions that influence mortality. Haines, Weiss, and Craig (2000) include a measure of the availability of transportation. The multivariate analysis includes this county-level variable as well as regional dummies for the West (Indiana, Michigan, Illinois, Iowa, and Texas) and South (North Carolina, Virginia, Alabama, and Kentucky).

Columns (1) and (3) of table 2.5 present baseline probit regressions with death (from any cause) as the dependent variable, and age and occupational categories as the only independent variables. Columns (2) and (4) introduce additional controls for individual and county characteristics. Separate regressions are shown for younger (aged twenty to forty-four) and older (aged forty-five and older) males to allow the effects of age and occupation to change as age increases.

As expected, there is a clear age pattern: The risk of death increases with age. But the pattern of increase is nonlinear, as age has a larger impact for those aged 45 and above. Death rates were higher among the foreign-born and lower in the South. The results for occupation provide little support for the hypothesis that those in lower-income occupations suffered from higher mortality rates: Among those with reported occupations, only white-collar workers in column (1) had a mortality rate different from laborers that was statistically significant and large in magnitude (with white-collar workers' mortality rate 3 per thousand lower than that of laborers, the omitted category). Even this difference is eliminated when the additional controls for individual and county characteristics are added in column (2).[15] The predicted probabilities in columns (2) and (4) in fact suggest that the mortality rate of unskilled laborers was roughly the same as that of white-collar workers, and perhaps somewhat below that of craftsmen, despite the low incomes earned by laborers.[16]

Table 2.6 presents a similar regression, with death from consumption (tuberculosis) as the dependent variable. The results in column (2) for males aged twenty to forty-four do indeed reveal substantial and statistically sig-

15. The much higher mortality of those with no reported occupation may be the result of the withdrawal from regular work of those debilitated by chronic health conditions in the time before their deaths. One in five males aged twenty and over in the mortality schedules used here had no reported occupation; for the corresponding population schedules, only one in ten had no reported occupation.

16. The predicted probabilities are calculated for a baseline individual (a native-born male born outside his state of enumeration, in a Northern county with no transportation access) at age thirty in columns (1) and (2) and at age fifty in columns (3) and (4).

Table 2.5 **Marginal Effects from Probit Regressions on Mortality From All Causes among Males Aged 20+, 1850**

	Ages 20–44 (1)	Ages 20–44 (2)	Ages 45+ (3)	Ages 45+ (4)
Age	0.0002	0.0001	0.0008	0.0008
	(3.51)***	(2.73)***	(5.06)***	(5.17)***
White-collar	–0.0030	–0.0007	–0.0033	0.0001
	(1.84)*	(0.40)	(0.64)	(0.02)
Farmer	–0.0014	0.0008	–0.0052	–0.0024
	(1.36)	(0.73)	(1.48)	(0.69)
Craftsman	0.0004	0.0016	–0.0007	0.0014
	(0.30)	(1.26)	(0.17)	(0.36)
Unknown	0.0247	0.0298	0.0302	0.0343
	(10.64)***	(11.63)***	(4.67)***	(5.07)***
Born in state of enumeration		0.0001		0.0024
		(0.13)		(0.85)
Foreign-born		0.0103		0.0139
		(7.86)***		(4.35)***
Transportation access		–0.0000		–0.0023
		(0.04)		(1.18)
West		0.0014		0.0025
		(1.54)		(1.00)
South		–0.0023		–0.0049
		(2.40)**		(1.93)*
Predicted probability				
White-collar	0.0054	0.0052	0.0137	0.0148
Farmer	0.0070	0.0064	0.0125	0.0127
Craftsman	0.0087	0.0071	0.0163	0.0161
Laborer	0.0084	0.0058	0.0171	0.0147
Unknown	0.0352	0.0322	0.0217	0.0208
Pseudo-R^2	0.0298	0.0452	0.0266	0.0374
Observations	64,241	64,241	18,932	18,932

Notes: Absolute value of z-statistics in parentheses. The figures shown are partial derivatives. Omitted categories for the categorical variables are "Laborer," "Born outside state of enumeration," "Native-born," "North," and "No access to rail or water transportation." Transportation access was taken from Craig, Palmquist, and Weiss (1998); the authors graciously provided their data in machine-readable format.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

nificant differences in mortality from consumption by occupation, controlling for other characteristics. The differences, however, are decidedly *not* those we would expect if income was all that occupational category indicated. Laborers have consumption mortality rates that are only half those of both white-collar workers and craftsmen.

This may reflect the importance of the workplace environment. Farmers and common laborers, most of whom in these rural counties would have

Table 2.6 Marginal Effects from Probit Regressions on Mortality From Consumption among Males Aged 20, 1850

	Ages 20–44 (1)	Ages 20–44 (2)	Ages 45+ (3)	Ages 45+ (4)
Age	–0.0000	–0.0000	0.0001	0.0001
	(0.37)	(0.20)	(1.66)*	(1.75)*
White-collar	0.0009	0.0018	–0.0013	–0.0008
	(1.11)	(1.98)**	(1.13)	(0.66)
Farmer	0.0002	0.0005	–0.0023	–0.0011
	(0.38)	(1.33)	(2.09)**	(1.22)
Craftsman	0.0011	0.0014	–0.0014	–0.0010
	(1.82)*	(2.35)**	(1.69)*	(1.19)
Unknown	0.0017	0.0027	0.0009	0.0020
	(2.06)**	(2.85)***	(0.66)	(1.29)
Born in state of enumeration		0.0005		–0.0007
		(1.52)		(0.93)
Foreign-born		0.0008		0.0006
		(1.76)*		(0.79)
Transportation access		0.0001		0.0004
		(0.25)		(0.69)
West		–0.0006		–0.0012
		(2.16)**		(1.96)**
South		–0.0012		–0.0014
		(3.86)***		(2.19)**
Predicted probability				
White-collar	0.0015	0.0026	0.0011	0.0021
Farmer	0.0010	0.0015	0.0014	0.0023
Craftsman	0.0018	0.0023	0.0011	0.0018
Laborer	0.0009	0.0009	0.0037	0.0042
Unknown	0.0023	0.0035	0.0023	0.0038
Pseudo-R^2	0.0073	0.0270	0.0252	0.0460
Observations	63,729	63,729	18,631	18,631

Notes: Absolute value of z-statistics in parentheses. The figures shown are partial derivatives. Omitted categories for the categorical variables are "Laborer," "Born outside state of enumeration," "Native-born," "North," and "No access to rail or water transportation." Transportation access was taken from Craig, Palmquist, and Weiss (1998); the authors graciously provided their data in machine-readable format.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

been farm laborers, generally worked outdoors and had fewer workplace opportunities to come into direct contact with other people than white-collar or craft workers. It may also reflect some self-selection into occupations consistent with health: Those debilitated by consumption may have taken up less physically demanding employment in workshops and offices rather than the strenuous manual work of farmers and common laborers. The results for consumption demonstrate the inadequacy of a simple measure of socioeconomic status (occupation) as a determinant of mortality, and leave open the possibility that socioeconomic status is at least in part determined by health status.

2.5 Analysis of Socioeconomic Status and among Families in 1850 and 1860

The sample of decedents merged with their surviving families in 1850 and 1860 presents two advantages over the data used in the preceding analysis of the link between occupation and mortality for adult males. The first is that it contains total family wealth (real estate wealth in 1850 and 1860, as well as personal wealth in 1860), which is likely to be a more meaningful measure of the economic resources at a family's disposal than the family head's occupational title (which may signify environmental conditions as well). The second is that it provides an opportunity to examine the mortality experienced by all family members, not just adult males. This is useful in itself, as the mortality of older children (aged five to nineteen) has been overlooked in previous studies of the socioeconomic correlates of mortality in the late nineteenth and early twentieth centuries. But the ability to look at several age categories is also useful in that it can eliminate the possibility that socioeconomic status itself is caused by health status, by focusing on the mortality of infants and children who were too young to provide income to the family.[17]

Table 2.7 presents marginal effects from probit regressions by age categories in which mortality is the dependent variable, with controls only for whether the family owned real estate and age in panel A and a full set of individual, family, and location controls in panel B.[18] In no case is the impact of possession of real estate by the family associated with a statistical or substantive reduction in mortality.[19] In six of the ten regressions, the sign of the coefficient on the possession of real estate is actually positive. Of the other controls, the most interesting results are for the occupation of the family head: For adults (aged twenty and over) in columns (4) and (5), residence in a family headed by either a farmer or an unskilled laborer was associated with lower mortality than residence in a family headed by a white-collar

17. This approach to overcoming the problem of reverse causation is suggested by Case, Lubotsky, and Paxson (2001).

18. Additional specifications for the wealth variable are employed in appendix tables 2A.4 through 2A.13—the natural log of wealth and dummies for various levels of wealth. None of the qualitative findings described here are altered by the use of these alternative specifications. The set of controls used in the following regressions differs slightly from that used in the regressions for males aged twenty and over in the previous section: family wealth and the family head's occupation replace the individual's own occupation as the measure of socioeconomic status (as children seldom had reported occupations), and dummies for state and year replace the controls for location (dummies for region and transportation access), which was preferable given the small number of locations and their relative homogeneity within states. Other family-level variables (e.g., parents' literacy and birthplaces) were introduced into the analysis, but did not have much influence on mortality and so were excluded.

19. Steckel (1988, 338–39) finds no relationship between family real estate holdings in 1850 and the survival of infants, children aged one to four, and female spouses between 1850 and 1860.

Table 2.7 **Marginal Effects from Probit Regressions on Mortality from All Causes, 1850 and 1860**

Variable	Infants (1)	Ages 1–4 (2)	Ages 5–19 (3)	Ages 20–44 (4)	Ages 45+ (5)
A.					
Real estate > 0	0.0091	0.0024	–0.0002	–0.0003	0.0042
	(1.29)	(1.26)	(0.29)	(0.38)	(1.60)
Age		–0.0076	–0.0000	0.0002	0.0004
		(8.06)***	(0.60)	(2.63)***	(3.67)***
Predicted probability	0.0473	0.0135	0.0035	0.0058	0.0131
Pseudo-R^2	0.0012	0.0331	0.0003	0.0031	0.0123
Observations	3,662	13,610	34,629	31,369	8,809
B.					
Real estate > 0	0.0085	0.0028	–0.0004	–0.0009	0.0033
	(1.15)	(1.43)	(0.53)	(1.06)	(1.41)
Age		–0.0074	–0.0000	0.0001	0.0005
		(8.09)***	(0.28)	(2.17)**	(4.34)***
Male	0.0128	0.0049	0.0013	0.0002	0.0082
	(1.82)*	(2.66)***	(2.00)**	(0.28)	(3.74)***
Family size	0.0018	0.0004	0.0002	0.0010	0.0019
	(0.96)	(0.68)	(0.93)	(4.72)***	(3.79)***
Born in state of enumeration	–0.0009	–0.0033	0.0006	0.0038	0.0125
	(0.05)	(1.05)	(0.88)	(3.57)***	(1.55)
Foreign-born		0.0090	–0.0013	0.0017	–0.0019
		(0.79)	(0.66)	(1.13)	(0.53)
Head was farmer	–0.0105	–0.0038	–0.0001	–0.0034	–0.0127
	(1.24)	(1.64)	(0.11)	(3.61)***	(4.61)***
Head was laborer	–0.0047	–0.0002	0.0004	–0.0027	–0.0074
	(0.35)	(0.05)	(0.28)	(1.90)*	(1.71)*
Illinois	0.0005	0.0022	–0.0001	0.0001	–0.0031
	(0.05)	(1.00)	(0.13)	(0.12)	(1.29)
1860	0.0045	0.0061	–0.0000	–0.0001	0.0047
	(0.60)	(3.12)***	(0.04)	(0.12)	(2.10)**
Predicted probability	0.4692	0.0128	0.0034	0.0052	0.0107
Pseudo-R^2	0.0057	0.0446	0.0046	0.0229	0.0568
Observations	3,647	13,610	34,629	31,369	8,809

Notes: Absolute value of *z*-statistics in parentheses. The figures shown are partial derivatives. The sample consists of all individuals in the population schedules and all individuals in the mortality schedules who were merged with families in the population schedules. Wealth is measured at the family level. Omitted categories for the categorical variables are "Real estate = 0," "Female," "Born outside state of enumeration," "Native-born," "Head was white-collar or craftsman," "Alabama," and "1850."

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

worker or craftsman. This is consistent with the finding for adult males in the previous section, although it applies to both males and females here.

Table 2.8 uses personal wealth rather than real estate wealth. The census did not begin to collect personal wealth data until 1860, so table 2.8 omits observations from 1850. The results are generally more favorable for the hypothesis that wealth was negatively associated with mortality. Either with or without the additional individual, family, and location controls, possession of personal wealth reduced mortality for infants, older children, and younger adults. For these groups, death was half as likely in the twelve months prior to the census in families that possessed personal wealth as it was in families that did not.

The finding that personal wealth has more impact on mortality than real estate wealth may reflect the greater liquidity of personal wealth, and the importance of the household's assets in smoothing consumption: When a negative shock to household income occurs, personal wealth can be liquidated more easily than real estate wealth to compensate for the shock. It would be easier for the household to sell some of its furniture or implements than it would be to sell some of its land—by their nature, moveable assets (personal estate) can be relocated to where there is a demand for them, while immoveable assets (real estate) must find a buyer at their fixed location. These effects are exacerbated if shocks to household income are correlated across the community (say, because of bad weather), since even fewer local buyers for the land a household wishes to liquidate will be available, while the option of transporting some personal property to a market center for liquidation remains.

There are noteworthy differences in the impact of wealth on mortality at different ages. The effect is greatest for infants but small in magnitude and statistically insignificant for children aged one to four. The effect of wealth is greater for infants than for older children aged five to nineteen.[20] In modern data, the effect of the family's economic circumstances on the health of its children increases as the age of the child increases (Case, Lubotsky, and Paxson 2001). This appears to be the case because children in low–socioeconomic status households receive a larger number of adverse shocks to their health as they age, rather than because they are less able to recover from a given shock (Currie and Stabile 2002). The absence of such a pattern for the period examined here may be the result of high infant death rates: Infants in low–socioeconomic status households, who would be at risk to die later in response to an adverse shock under modern conditions where infant deaths are rare, are in effect "weeded out" by high infant mortality.

For 1860, it is also possible to examine the simultaneous influence of real and personal wealth on mortality. Table 2.9 includes both. The finding of a

20. Condran and Crimmins (1979, 14) believe that the five-to-twenty age group was the most accurately reported in the mortality schedules. If they are correct, then the impact of wealth for this group's mortality shown in table 2.8 is perhaps the strongest evidence that wealth's effect is more than an artifact of inaccuracies in the mortality schedules.

Table 2.8 **Marginal Effects From Probit Regressions on Mortality from All Causes, 1860**

Variable	Infants (1)	Ages 1–4 (2)	Ages 5–19 (3)	Ages 20–44 (4)	Ages 45+ (5)
A.					
Personal estate > 0	–0.0438	–0.0011	–0.0027	–0.0048	0.0086
	(2.71)***	(0.25)	(1.90)*	(2.51)**	(1.50)
Age		–0.0107	–0.0001	0.0001	0.0006
		(8.00)***	(0.66)	(1.72)*	(3.61)***
Predicted probability	0.0486	0.0148	0.0034	0.0056	0.0152
Pseudo-R^2	0.0081	0.0513	0.0042	0.0066	0.0173
Observations	2,154	7,730	19,480	18,279	5,136
B.					
Personal estate > 0	–0.0386	–0.0005	–0.0025	–0.0050	0.0090
	(2.32)**	(0.12)	(1.66)*	(2.64)***	(1.76)*
Age		–0.0106	–0.0000	0.0001	0.0007
		(8.03)***	(0.23)	(1.38)	(4.09)***
Male	0.0175	0.0071	0.0004	–0.0014	0.0061
	(1.88)*	(2.77)***	(0.46)	(1.38)	(1.91)*
Family size	0.0000	0.0006	0.0002	0.0010	0.0022
	(0.01)	(0.85)	(0.96)	(3.94)***	(2.81)***
Born in state of enumeration	0.0076	–0.0009	0.0017	0.0031	0.0220
	(0.37)	(0.22)	(1.89)*	(2.45)**	(1.92)*
Foreign-born		0.0413		–0.0000	–0.0002
		(1.58)		(0.02)	(0.03)
Head was farmer	–0.0098	–0.0037	–0.0007	–0.0017	–0.0075
	(0.89)	(1.19)	(0.65)	(1.46)	(1.87)*
Head was laborer	–0.0047	0.0008	0.0009	–0.0021	–0.0061
	(0.31)	(0.17)	(0.50)	(1.21)	(0.89)
Illinois	0.0057	0.0031	–0.0000	–0.0000	–0.0076
	(0.58)	(1.11)	(0.06)	(0.03)	(2.22)**
Predicted probability	0.0479	0.0142	0.0033	0.0050	0.0136
Pseudo-R^2	0.0141	0.0615	0.0109	0.0274	0.0448
Observations	2,149	7,730	19,009	18,279	5,136

Notes: Absolute value of *z*-statistics in parentheses. The figures shown are partial derivatives. The sample consists of all individuals in the population schedules and all individuals in the mortality schedules who were merged with families in the population schedules. Wealth is measured at the family level. Omitted categories for the categorical variables are “Personal estate = 0,” “Female,” “Born outside state of enumeration,” “Native-born,” “Head was white-collar or craftsman,” and “Alabama.”

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Table 2.9 **Marginal Effects from Probit Regressions on Mortality from All Causes, 1860**

Variable	Infants (1)	Ages 1–4 (2)	Ages 5–19 (3)	Ages 20–44 (4)	Ages 45+ (5)
A.					
Real estate > 0	0.0119	0.0037	0.0001	0.0023	0.0049
	(1.19)	(1.32)	(0.09)	(1.87)	(1.20)
Personal estate > 0	–0.0535	–0.0034	0.0028	–0.0070	0.0067
	(2.94)***	(0.68)	(1.74)*	(3.04)**	(1.05)
Age		–0.0107	–0.0001	0.0001	0.0006
		(8.00)***	(0.66)	(1.53)	(3.57)***
Predicted probability	0.0484	0.0147	0.0034	0.0055	0.0151
Pseudo-R^2	0.0098	0.0525	0.0042	0.0094	0.0190
Observations	2,154	7,730	19,480	18,279	5,136
B.					
Real estate > 0	0.0134	0.0048	–0.0002	0.0013	0.0039
	(1.28)	(1.70)*	(0.18)	(1.14)	(1.02)
Personal estate > 0	–0.0479	–0.0031	–0.0023	–0.0063	0.0077
	(2.60)***	(0.64)	(1.44)	(2.86)***	(1.38)
Age		–0.0105	–0.0000	0.0001	0.0007
		(8.03)***	(0.22)	(1.31)	(4.06)***
Male	0.0174	0.0070	0.0004	–0.0014	0.0062
	(1.88)*	(2.76)***	(0.46)	(1.38)	(1.93)*
Family size	–0.0005	0.0004	0.0002	0.0010	0.0021
	(0.21)	(0.58)	(0.97)	(3.80)***	(2.78)***
Born in state of enumeration	0.0057	–0.0016	0.0018	0.0030	0.0223
	(0.27)	(0.40)	(1.89)*	(2.37)**	(1.94)*
Foreign-born		0.0401		–0.0001	–0.0003
		(1.56)		(0.05)	(0.05)
Head was farmer	–0.0100	–0.0041	–0.0006	–0.0019	–0.0078
	(0.90)	(1.30)	(0.63)	(1.55)	(1.95)*
Head was laborer	–0.0004	0.0024	0.0008	–0.0018	–0.0052
	(0.02)	(0.49)	(0.47)	(1.00)	(0.73)
Illinois	0.0059	0.0032	–0.0000	0.0000	–0.0072
	(0.60)	(1.15)	(0.06)	(0.03)	(2.12)**
Predicted probability	0.0476	0.0140	0.0033	0.0049	0.0135
Pseudo-R^2	0.0160	0.0635	0.0109	0.0284	0.0460
Observations	2,149	7,730	19,009	18,279	5,136

Notes: Absolute value of z-statistics in parentheses. The figures shown are partial derivatives. The sample consists of all individuals in the population schedules and all individuals in the mortality schedules who were merged with families in the population schedules. Wealth is measured at the family level. Omitted categories for the categorical variables are "Real estate = 0," "Personal estate = 0," "Female," "Born outside state of enumeration," "Native-born," "Head was white-collar or craftsman," and "Alabama."

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

strong negative relationship between personal wealth and mortality for several age groups (particularly infants and adults aged twenty to forty-four) remains. The results for real wealth in panel A now present a puzzle, however. For adults aged twenty to forty-four, possession of real estate increased mortality risk, controlling for personal wealth. This effect is both large in magnitude and statistically significant. For adults aged forty-five and over, mortality is also higher among those in families with real estate than among those in families without it, although not at conventional levels of statistical significance. Although it seems plausible that personal wealth would provide more protection against mortality than real wealth, it is unclear why real wealth would actually lead to increased mortality. Part of the answer lies in the absence of the full set of controls: When other personal, family, and location controls are added in panel B, the negative effect of real estate wealth is reduced.

In order to explore this anomaly further, however, a final specification was adopted that allows the effects of wealth and age to differ across locations. Table 2.10 examines mortality with controls for age and for the possession of real wealth and personal wealth, as well as interactions between these and residence in Illinois. The effect of wealth is not uniform across locations. The perverse positive relationship between possession of real estate wealth and mortality is observed only in Alabama; in Illinois, it is exactly offset by the interaction for ages twenty to forty-four and more than offset for ages forty-five and over. In the latter case, the possession of real estate wealth is now associated with unambiguously lower mortality in Illinois. The interactions between Illinois and both age and personal wealth are statistically and substantively insignificant, so these effects are similar in Illinois and Alabama. Two possible explanations for why real estate ownership is associated with higher mortality in Alabama but not in Illinois come to mind. The first follows from differences in physical geography. In the South, some of the most valuable land was alluvial property near rivers and streams, at low elevations. These places had soil and climate conditions particularly conducive to the cultivation of cotton and commanded high prices per acre. But such places may have been particularly unhealthy locations in which to live, compared to land at higher elevations. Families with high levels of real estate wealth may have been more likely to own land in these relatively less healthy locations.

The second obvious difference between the two states is the presence of slaves in Alabama. Although neither St. Clair County nor Tuscaloosa County had unusually high numbers of slaves per farm compared to other counties in Alabama or in the rest of the South, they nonetheless had more slaves than Perry, Shelby, and Vermilion Counties in Illinois.[21] Families in

21. St. Clair County had an average of only two slaves per farm, and Tuscaloosa had seven (which was the median for Alabama counties in 1860). Wilcox County had the state's highest ratio of slaves per farm with sixty-two. The mean for the entire state was nine.

Table 2.10 Marginal Effects from Probit Regressions on Mortality from All Causes, 1860

Variable	Infants (1)	Ages 1–4 (2)	Ages 5–19 (3)	Ages 20–44 (4)	Ages 45+ (5)
Real estate > 0	0.0173	0.0072	–0.0012	0.0067	0.0172
	(0.97)	(1.25)	(0.72)	(2.82)***	(2.57)**
(Illinois) · real estate > 0	–0.073	–0.0044	0.0019	–0.0066	–0.0268
	(0.34)	(0.66)	(0.92)	(2.35)**	(2.71)***
Personal estate > 0	–0.0528	–0.0095	–0.0022	–0.0079	0.0020
	(1.83)*	(1.01)	(1.01)	(1.98)**	(0.22)
(Illinois) · personal estate > 0	–0.0027	0.0057	–0.0003	0.0005	0.0132
	(0.09)	(0.62)	(0.11)	(0.14)	(0.86)
Age		–0.0101	–0.0000	0.0000	0.0006
		(3.94)***	(0.02)	(0.21)	(2.74)***
(Illinois) · age		–0.0006	–0.0001	0.0001	–0.0001
		(0.19)	(0.47)	(0.84)	(0.36)
Illinois	0.0129	0.0028	–0.0004	–0.0006	0.0087
	(0.55)	(0.29)	(0.14)	(0.10)	(0.37)
Predicted probability	0.0483	0.0146	0.0033	0.0053	0.0140
Pseudo-R^2	0.0105	0.0543	0.0060	0.0157	0.0341
Observations	2,154	7,730	19,480	18,279	5,136

Notes: Absolute value of z-statistics in parentheses. The figures shown are partial derivatives. The sample consists of all individuals in the population schedules and all individuals in the mortality schedules who were merged with families in the population schedules. Wealth is measured at the family level. Omitted categories for the categorical variables are "Real estate = 0," "Personal estate = 0," and "Alabama."

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Alabama, where greater real estate wealth was likely associated with the ownership of slaves, may have had more daily exposure to individuals lower in socioeconomic status, and therefore had a greater likelihood of contracting infectious diseases than families on isolated farms in Illinois whose only contact with non–family members may have been occasional trips to town or visits to neighbors whose socioeconomic status would not have differed markedly from their own.[22]

22. To assess these two hypotheses, future work on the link between wealth and mortality in the South will explore (a) the impact of the characteristics of the minor civil divisions in which families were located (by including information on such local attributes as elevation and soil type, available from the U.S. Geological Survey and from the mortality schedules themselves) and (b) the impact of the presence of slaves on individual farms (by merging families from the population schedules with their data from the slave schedules, which reported the age and number of slaves owned on each farm in the South). Both of the merged samples used here will also be linked to data on wages at the county level in the Census of Social Statistics described in Margo (2000).

2.6 Conclusions and Extensions

Socioeconomic status was an important force shaping the mortality rates experienced by Americans in the middle of the nineteenth century, at least in the sample of rural counties examined here. Although occupation was a poor proxy for status among adult males in 1850, the effect of personal wealth on mortality was quite large in magnitude. For example, using the coefficients in panel B of column (1) in table 2.8, the mortality rate for male infants born and residing in Illinois in five-person families headed by farmers was 97 per thousand in families that did not own any personal wealth; the mortality rate for otherwise identical infants was roughly half as great (53 per thousand) in families that possessed any personal wealth.[23] Using the coefficients in column (4), the mortality rate for thirty-year-old males born and residing in Illinois in five-person families headed by farmers was 11 per thousand in families that did not own any personal wealth; the mortality rate for otherwise identical thirty-year-old males was less than half as great (5 per thousand) in families that possessed any personal wealth.[24]

The analysis presented here suffers from two principal shortcomings. The first is the inability to say anything about the experience of urban dwellers. Data for Chicago and Albany will be added as the project progresses, but more information from the cities of the Northeast—inundated with immigrants and beset with crowding, poor sanitation, and substandard housing—will be essential to understand the full scope of mid-nineteenth-century America's mortality record. The second shortcoming is the only brief attention given to causes of death and their likely different relationships to socioeconomic status. Nonetheless, these findings suggest that when Americans moved into cities and towns and factories as the first half of the nineteenth century closed, they had already experienced substantial disparities in health outcomes in the rural, agricultural settings they left behind. Even on farms and in small towns, the more affluent experienced longer lives than their poorer neighbors.

This is not to say that the disparities in health outcomes by socioeconomic status observed today simply continued through the end of the nineteenth and into the twentieth century. It is possible that, as urbanization occurred and individuals were exposed to an increasing number of health hazards related to crowding and sanitation from which wealth might pro-

23. Slightly less dramatic (although still substantial) differences can be seen if the coefficients in table 2A.9 are employed: Using the same values for the other control variables, infants in families with no personal wealth faced a mortality rate of eighty-eight per thousand. In families with $250 in personal wealth (the 3rd decile of the family personal wealth distribution), they faced a rate of forty-two per thousand, while in families with $2,500 in personal wealth (the top decile), they faced a rate of thirty-three per thousand.

24. If the coefficients in table 2A.9 are employed, using the same values for the other control variables, thirty-year-old males in families with no personal wealth faced a mortality rate of eleven per thousand. In families with $250 in personal wealth, they faced a rate of five per thousand, while in families with $2,500 in personal wealth, they faced a rate of four per thousand.

vide little escape, the socioeconomic status–mortality gradient became less steep. With the eradication of many urban health hazards in the twentieth century, a significant role for high status in preventing disease and death may have reappeared.

The significance of the gap in mortality between high- and low-wealth households in the middle of the nineteenth century is instead its appearance at a time and place lacking many of the advantages thought to contribute to better health and lower mortality among the wealthy today: education and knowledge of sound health practices, and access to health care professionals, sophisticated diagnostic technologies, and efficacious treatments. Large differences in mortality between those with and those without wealth in rural communities in the mid-nineteenth-century United States demonstrate the important role played by general living standards and the material conditions of day-to-day life in shaping mortality patterns.

Appendix

Table 2A.1 Descriptive Statistics for Male Decedents and Survivors Aged 20 and Over in 50 Counties, 1850

Variable	Mean	Standard Deviation	Minimum	Maximum
Died in twelve months prior to census				
All causes	0.0111	0.1050	0	1
Consumption	0.0014	0.0371	0	1
Age				
20–44 years	0.7724	0.4193	0	1
45+ years	0.2276	0.4193	0	1
Birthplace				
State of enumeration	0.2972	0.4571	0	1
Foreign-born	0.1118	0.3151	0	1
Transportation access	0.5925	0.4914	0	1
Region				
North	0.2195	0.4139	0	1
West	0.4323	0.4954	0	1
South	0.3482	0.4764	0	1
Occupation				
White-collar	0.0564	0.2306	0	1
Farmer	0.5899	0.4919	0	1
Craftsman	0.1508	0.3579	0	1
Laborer	0.1430	0.3500	0	1
Unknown	0.0599	0.2374	0	1
Observations	83,173			

Note: Decedents were drawn from the mortality schedules of the 1850 U.S. Census of Population; survivors were drawn from the population schedules of the 1850 U.S. Census of Population.

Table 2A.2 **Descriptive Statistics for Decedents and Survivors in 11 Alabama and Illinois Counties, 1850 and 1860**

Variable	Mean	Standard Deviation	Minimum	Maximum
Died in twelve months prior to census				
All ages	0.0089	0.0937	0	1
Infant	0.0475	0.2128	0	1
Ages 1–4	0.0158	0.1247	0	1
Ages 5–19	0.0035	0.0588	0	1
Ages 20–44	0.0059	0.0764	0	1
Ages 45+	0.0139	0.1169	0	1
Age				
Infant	0.0398	0.1954	0	1
Ages 1–4	0.1478	0.3549	0	1
Ages 5–19	0.3761	0.4844	0	1
Ages 20–44	0.3407	0.4739	0	1
Ages 45+	0.0957	0.2941	0	1
Male	0.5152	0.4998	0	1
Family size	5.9572	1.9853	2	9
Birthplace				
State of enumeration	0.5311	0.4990	0	1
Foreign-born	0.0550	0.2281	0	1
County				
Alabama	0.2443	0.4297	0	1
Shelby (1850)	0.0593	0.2362	0	1
St. Clair (1860)	0.0796	0.2706	0	1
Tuscaloosa (1860)	0.1054	0.3071	0	1
Illinois	0.7557	0.4297	0	1
Jackson (1850)	0.0536	0.2253	0	1
Morgan (1850)	0.1359	0.3427	0	1
Saline (1850)	0.0488	0.2154	0	1
Union (1850)	0.0635	0.2439	0	1
Washington (1850)	0.0657	0.2477	0	1
Perry (1860)	0.0869	0.2817	0	1
Shelby (1860)	0.1281	0.3342	0	1
Vermilion (1860)	0.1732	0.3784	0	1
Family real estate = 0	0.3940	0.4886	0	1
Family real estate > 0	0.6060	0.4886	0	1
Under \$500	0.1741	0.3792	0	1
\$500–999	0.1182	0.3229	0	1
\$1,000–2,499	0.1924	0.3942	0	1
\$2,500–4,999	0.0841	0.2776	0	1
\$5,000+	0.0371	0.1891	0	1
Log(family real estate + \$1.00)	4.1209	3.4464	0	9.2103
Family personal estate = 0	0.1014	0.3019	0	1
Family personal estate > 0	0.8986	0.3019	0	1
Under \$500	0.4333	0.4955	0	1
\$500–999	0.2121	0.4088	0	1

Table 2A.2 (continued)

Variable	Mean	Standard Deviation	Minimum	Maximum
$1,000–2,499	0.1697	0.3754	0	1
$2,500–4,999	0.0381	0.1915	0	1
$5,000+	0.0453	0.2079	0	1
Log (family personal estate + $1.00)	5.6029	2.2371	0	13.1902
Occupation of Household Head				
Laborer	0.0626	0.2423	0	1
Farmer	0.6926	0.4614	0	1
Observations	92,079			

Note: Single-person families, families of ten or more, and households with real estate wealth over $10,000 are excluded.

Table 2A.3 **Causes of Death (percent)**

Cause	1850 Sample: Males Males in 50 Counties, Ages 20+	1850 and 1860 Samples: 11 Counties in Illinois and Alabama				
		Infant	Ages 1–4	Ages 5–19	Ages 20–44	Ages 45+
Consumption	12.2	0.6	0.5	3.3	19.8	10.7
Fevers[a]	17.6	9.2	16.3	35.8	29.1	22.1
Typhoid/typhus	2.7	—	0.5	0.8	1.1	0.8
Pneumonia	4.2	2.3	3.4	4.2	5.5	7.4
Diarrhea	3.0	2.2	2.3	0.8	—	0.8
Cholera[b]	8.1	4.0	1.9	0.8	5.0	0.8
Dropsy	3.8	—	—	0.8	3.9	9.0
Accident[c]	4.2	0.6	3.1	6.7	3.9	2.5
Apoplexy	1.7	—	—	—	—	1.6
Croup	—	14.9	15.8	6.7	1.1	—
Flux	1.3	1.2	—	1.7	—	—
Whooping cough	—	4.6	2.3	0.8	—	—
Dysentery	0.7	2.3	3.3	1.7	0.6	—
Brain inflammation	1.7	3.2	4.2	2.5	1.2	0.8
Old age	0.5	—	—	—	—	4.9
Unknown	6.4	16.1	8.8	6.7	3.3	4.9
Observations	927	154	215	120	182	122

[a]Includes yellow fever, lung fever, and brain fever.

[b]Includes cholera infantum.

[c]Includes murder, suicide, drowning, burning, scalding, and gunshot.

Table 2A.4 **Probit Regressions on Mortality, 1850 and 1860 (partial derivatives)**

Variable	Infants (1)	Ages 1–4 (2)	Ages 5–19 (3)	Ages 20–44 (4)	Ages 45+ (5)
A.					
Log(real estate + $1.00)	0.0012	0.0005	–0.0001	–0.0000	0.0006
	(1.17)	(1.83)*	(0.77)	(0.24)	(1.66)*
Age		–0.0075	–0.0000	0.0002	0.0004
		(8.06)***	(0.57)	(2.63)***	(3.69)***
Predicted probability	0.0474	0.0134	0.0035	0.0058	0.0131
Pseudo-R^2	0.0010	0.0396	0.0037	0.0030	0.0125
Observations	3,662	13,610	34,629	31,369	8,809
B.					
Real estate					
$100–499	0.0129	0.0031	0.0001	–0.0014	0.0020
	(1.29)	(1.16)	(0.15)	(1.19)	(0.49)
$500–999	–0.0032	–0.0034	0.0010	0.0002	0.0027
	(0.28)	(1.10)	(0.97)	(0.17)	(0.62)
$1,000–2,499	0.0179	0.0013	–0.0003	0.0009	0.0078
	(1.63)	(0.48)	(0.33)	(0.76)	(2.11)**
$2,500–4,999	0.0087	0.0063	–0.0013	–0.0008	0.0032
	(0.55)	(1.49)	(1.20)	(0.51)	(0.71)
$5,000+	–0.0050	0.0269	–0.0021	–0.0024	0.0076
	(0.20)	(3.55)***	(1.33)	(1.04)	(1.23)
Age		–0.0074	–0.0000	0.0002	0.0004
		(8.04)***	(0.46)	(2.56)**	(3.72)***
Predicted probability	0.0471	0.0131	0.0034	0.0057	0.0130
Pseudo-R^2	0.0031	0.0396	0.0037	0.0049	0.0142
Observations	3,662	13,610	34,629	31,369	8,809

Note: Absolute value of *z*-statistics in parentheses.
***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Table 2A.5 **Probit Regressions on Mortality, 1850 and 1860 (partial derivatives)**

Variable	Infants (1)	Ages 1–4 (2)	Ages 5–19 (3)	Ages 20–44 (4)	Ages 45+ (5)
Log(real estate + $1.00)	0.0011	0.0005	–0.0001	–0.0001	0.0004
	(0.97)	(1.74)*	(1.08)	(1.01)	(1.29)
Age		–0.0074	–0.0000	0.0001	0.0005
		(8.09)***	(0.21)	(2.16)**	(4.35)***
Male	0.0129	0.0049	0.0013	0.0002	0.0082
	(1.83)*	(2.66)***	(2.02)**	(0.29)	(3.73)***
Family size	0.0018	0.0003	0.0002	0.0010	0.0019
	(0.93)	(0.56)	(1.01)	(4.72)***	(3.76)***
Born in state of enumeration	–0.0007	–0.0035	0.0007	0.0038	0.0125
	(0.04)	(1.10)	(0.97)	(3.57)***	(1.55)
Foreign-born		0.0087	–0.0013	0.0017	–0.0019
		(0.77)	(0.65)	(1.13)	(0.51)
Head was farmer	–0.0104	–0.0038	–0.0000	–0.0035	–0.0126
	(1.23)	(1.67)*	(0.05)	(3.62)***	(4.59)***
Head was laborer	–0.0050	0.0001	0.0003	–0.0027	–0.0074
	(0.37)	(0.03)	(0.20)	(1.90)*	(1.71)*
Illinois	0.0000	0.0020	–0.0000	0.0002	–0.0033
	(0.00)	(0.91)	(0.06)	(0.17)	(1.39)
1860	0.0041	0.0057	0.0001	–0.0000	0.0045
	(0.53)	(2.92)***	(0.10)	(0.06)	(2.01)**
Predicted probability	0.0469	0.0128	0.0034	0.0052	0.0107
Pseudo-R^2	0.0054	0.0451	0.0051	0.0228	0.0566
Observations	3,647	13,610	34,629	31,369	8,809

Note: Absolute value of z-statistics in parentheses.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Table 2A.6 Probit Regressions on Mortality, 1850 and 1860 (partial derivatives)

Variable	Infants (1)	Ages 1–4 (2)	Ages 5–19 (3)	Ages 20–44 (4)	Ages 45+ (5)
Real estate					
$100–499	0.0135	0.0051	0.0001	–0.0016	0.0030
	(1.32)	(1.84)*	(0.07)	(1.45)	(0.83)
$500–999	–0.0043	–0.0025	0.0008	–0.0001	0.0018
	(0.37)	(0.79)	(0.79)	(0.12)	(0.48)
$1,000–2,499	0.0163	0.0012	–0.0005	0.0000	0.0061
	(1.45)	(0.44)	(0.66)	(0.04)	(1.87)*
$2,500–4,999	0.0068	0.0044	–0.0016	–0.0014	0.0020
	(0.42)	(1.07)	(1.50)	(0.95)	(0.51)
$5,000+	–0.0087	0.0234	–0.0023	–0.0028	0.0038
	(0.35)	(3.19)***	(1.55)	(1.47)	(0.76)
Age		–0.0073	–0.0000	0.0001	0.0005
		(8.08)***	(0.06)	(2.12)**	(4.35)***
Male	0.0128	0.0048	0.0012	0.0002	0.0082
	(1.82)*	(2.63)***	(2.03)**	(0.30)	(3.73)***
Family size	0.0020	0.0002	0.0002	0.0010	0.0019
	(1.03)	(0.44)	(1.15)	(4.72)***	(3.75)***
Born in state of enumeration	0.0001	–0.0035	0.0007	0.0037	0.0122
	(0.01)	(1.10)	(1.06)	(3.58)***	(1.52)
Foreign-born		0.0091	–0.0013	0.0016	–0.0020
		(0.80)	(0.67)	(1.06)	(0.55)
Head was farmer	–0.0108	–0.0037	–0.0001	–0.0034	–0.0128
	(1.28)	(1.62)	(0.11)	(3.62)***	(4.65)***
Head was laborer	–0.0053	–0.0002	0.0002	–0.0026	–0.0073
	(0.39)	(0.05)	(0.15)	(1.87)*	(1.69)*
Illinois	0.0006	0.0020	0.0003	0.0001	–0.0031
	(0.07)	(0.89)	(0.37)	(0.15)	(1.26)
1860	0.0049	0.0057	0.0004	–0.0001	0.0046
	(0.63)	(2.82)***	(0.58)	(0.13)	(1.99)**
Predicted probability	0.0466	0.0125	0.0033	0.0051	0.0106
Pseudo-R^2	0.0079	0.0501	0.0086	0.0247	0.0581
Observations	3,647	13,610	34,629	31,369	8,809

Note: Absolute value of z-statistics in parentheses.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Table 2A.7	Probit Regressions on Mortality, 1860 (partial derivatives)				
Variable	Infants (1)	Ages 1–4 (2)	Ages 5–19 (3)	Ages 20–44 (4)	Ages 45+ (5)
A.					
Log(personal estate + $1.00)	–0.0048	–0.0001	–0.0003	–0.0002	0.0016
	(2.44)**	(0.23)	(1.94)*	(1.05)	(2.00)**
Age		–0.0107	–0.0001	0.0001	0.0006
		(7.99)***	(0.61)	(1.74)*	(3.60)***
Predicted probability	0.0488	0.0148	0.0034	0.0057	0.0151
Pseudo-R^2	0.0068	0.0513	0.0045	0.0031	0.0192
Observations	2,154	7,730	19,480	18,279	5,136
B.					
Personal estate					
$100–499	–0.0431	–0.0004	–0.0018	–0.0045	0.119
	(3.10)***	(0.09)	(1.53)	(2.84)***	(1.43)
$500–999	–0.0139	–0.0050	–0.0015	–0.0035	0.0094
	(1.01)	(1.06)	(1.31)	(2.32)**	(1.03)
$1,000–2,499	–0.0283	–0.0003	–0.0027	–0.0019	0.0097
	(1.96)**	(0.05)	(2.32)**	(1.16)	(1.05)
$2,500–4,999	–0.0346	0.0071	–0.0013	–0.0034	0.0297
	(1.61)	(0.79)	(0.70)	(1.40)	(1.90)*
$5,000+	–0.0321	0.0015	–0.0016	0.0005	0.0260
	(1.40)	(0.18)	(0.98)	(0.18)	(1.99)**
Age		–0.0107	–0.0001	0.0001	0.0006
		(8.00)***	(0.63)	(1.67)*	(3.53)***
Predicted probability	0.0476	0.0146	0.0033	0.0055	0.0150
Pseudo-R^2	0.0152	0.0537	0.0068	0.0115	0.0216
Observations	2,154	7,730	19,480	18,279	5,136

Note: Absolute value of *z*-statistics in parentheses.
***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Table 2A.8 **Probit Regressions on Mortality, 1860 (partial derivatives)**

Variable	Infants (1)	Ages 1–4 (2)	Ages 5–19 (3)	Ages 20–44 (4)	Ages 45+ (5)
Log(personal estate + $1.00)	–0.0043	0.0001	–0.0003	–0.0004	0.0013
	(2.03)**	(0.17)	(1.96)*	(1.73)*	(1.79)*
Age		–0.0106	–0.0000	0.0001	0.0007
		(8.02)***	(0.16)	(1.42)	(4.09)***
Male	0.0174	0.0071	0.0004	–0.0014	0.0062
	(1.86)*	(2.77)***	(0.48)	(1.36)	(1.91)*
Family size	0.0005	0.0006	0.0003	0.0010	0.0021
	(0.19)	(0.79)	(1.08)	(3.91)***	(2.76)***
Born in state of enumeration	0.0078	–0.0010	0.0018	0.0033	0.0216
	(0.38)	(0.24)	(2.01)**	(2.57)**	(1.90)*
Foreign-born		0.0412		0.0000	–0.0001
		(1.58)		(0.00)	(0.02)
Head was farmer	–0.0108	–0.0039	–0.0005	–0.0020	–0.0077
	(0.98)	(1.24)	(0.55)	(1.61)	(1.91)*
Head was laborer	–0.0075	0.0008	0.0007	–0.0023	–0.0055
	(0.51)	(0.17)	(0.38)	(1.34)	(0.79)
Illinois	0.0046	0.0031	–0.0001	–0.0003	–0.0062
	(0.47)	(1.10)	(0.15)	(0.28)	(1.85)*
Predicted probability	0.0480	0.0141	0.0033	0.0050	0.0136
Pseudo-R^2	0.0128	0.0615	0.0121	0.0247	0.0445
Observations	2,149	7,730	19,009	18,279	5,136

Note: Absolute value of z-statistics in parentheses.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Table 2A.9 **Probit Regressions on Mortality, 1860 (partial derivatives)**

Variable	Infants (1)	Ages 1–4 (2)	Ages 5–19 (3)	Ages 20–44 (4)	Ages 45+ (5)
Personal estate					
$100–499	–0.0375	–0.0004	–0.0015	–0.0040	0.0140
	(2.62)***	(0.09)	(1.25)	(2.68)***	(1.73)*
$500–999	–0.0075	–0.0048	–0.0015	–0.0036	0.0122
	(0.49)	(1.00)	(1.16)	(2.48)**	(1.32)
$1,000–2,499	–0.0238	0.0002	–0.0028	–0.0025	0.0109
	(1.54)	(0.05)	(2.26)**	(1.64)	(1.18)
$2,500–4,999	–0.0324	0.0085	–0.0015	–0.0035	0.0263
	(1.45)	(0.92)	(0.83)	(1.76)*	(1.78)*
$5,000+	–0.0281	0.0053	–0.0019	–0.0010	0.0211
	(1.15)	(0.57)	(1.17)	(0.45)	(1.76)*
Age		–0.0105	–0.0000	0.0001	0.0007
		(8.02)***	(0.15)	(1.39)	(4.06)***
Male	0.0166	0.0071	0.0004	–0.0014	0.0062
	(1.81)*	(2.80)***	(0.51)	(1.43)	(1.92)*
Family size	–0.0005	0.0006	0.0003	0.0010	0.0022
	(0.20)	(0.77)	(1.10)	(3.75)***	(2.83)***
Born in state of enumeration	0.0053	–0.0007	0.0018	0.0029	0.0214
	(0.25)	(0.18)	(2.04)**	(2.33)**	(1.89)*
Foreign-born		0.0426		–0.0001	–0.0002
		(1.64)		(0.03)	(0.04)
Head was farmer	–0.0138	–0.0031	–0.0006	–0.0017	–0.0068
	(1.23)	(0.99)	(0.59)	(1.41)	(1.72)*
Head was laborer	–0.0024	0.0008	0.0006	–0.0019	–0.0062
	(0.15)	(0.18)	(0.35)	(1.10)	(0.91)
Illinois	0.0028	0.0041	–0.0000	0.0003	–0.0064
	(0.28)	(1.45)	(0.02)	(0.24)	(1.79)*
Predicted probability	0.0467	0.0139	0.0032	0.0049	0.0135
Pseudo-R^2	0.0217	0.0643	0.0143	0.0298	0.0465
Observations	2,149	7,730	19,009	18,279	5,136

Note: Absolute value of z-statistics in parentheses.
***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Table 2A.10 Probit Regressions on Mortality, 1860 (partial derivatives)

Variable	Infants (1)	Ages 1–4 (2)	Ages 5–19 (3)	Ages 20–44 (4)	Ages 45+ (5)
A.					
Log (real estate + \$1.00)	0.0024	0.0009	0.0000	0.0003	0.0004
	(1.54)	(2.04)**	(0.01)	(1.53)	(0.61)
Log (personal estate + \$1.00)	–0.0068	–0.0009	–0.0003	–0.0005	0.0013
	(2.88)***	(1.31)	(1.58)	(1.73)*	(1.43)
Age		–0.0106	–0.0001	0.0001	0.0006
		(7.99)***	(0.61)	(1.57)	(3.59)***
Predicted probability	0.0484	0.0146	0.0034	0.0057	0.0151
Pseudo-R^2	0.0096	0.0542	0.0045	0.0049	0.0197
Observations	2,154	7,730	19,480	18,279	5,136
B.					
Real estate					
\$100–499	0.0205	0.0065	0.0011	0.0014	0.0011
	(1.35)	(1.48)	(0.79)	(0.70)	(0.17)
\$500–999	–0.0190	–0.0017	0.0001	0.0038	0.0048
	(1.23)	(0.36)	(0.04)	(1.71)*	(0.74)
\$1,000–2,499	0.0179	0.0033	0.0005	0.0024	0.0117
	(1.15)	(0.79)	(0.39)	(1.31)	(1.96)*
\$2,500–4,999	0.0159	0.0118	–0.0005	0.0005	0.0001
	(0.77)	(1.90)*	(0.31)	(0.24)	(0.02)
\$5,000+	–0.0105	0.0338	–0.0012	–0.0013	0.0072
	(0.36)	(3.21)***	(0.60)	(0.46)	(0.87)
Personal estate					
\$100–499	0.0454	0.0016	0.0020	0.0052	0.0090
	(3.19)***	(0.35)	(1.63)	(3.15)***	(1.09)
\$500–999	–0.0192	–0.0073	–0.0016	–0.0043	0.0036
	(1.29)	(1.53)	(1.25)	(2.66)***	(0.41)
\$1,000–2,499	–0.0319	–0.0054	–0.0026	–0.0025	0.0035
	(2.03)**	(1.05)	(1.91)*	(1.37)	(0.39)
\$2,500–4,999	–0.0368	–0.0008	–0.0012	–0.0038	0.0186
	(1.75)*	(0.10)	(0.60)	(1.62)	(1.30)
\$5,000+	–0.0332	–0.0057	–0.0015	–0.0004	0.0153
	(1.43)	(0.80)	(0.81)	(0.13)	(1.26)
Age		–0.0104	–0.0001	0.0001	0.0006
		(7.98)***	(0.57)	(1.56)	(3.52)***
Predicted probability	0.0462	0.0141	0.0033	0.0054	0.0145
Pseudo-R^2	0.0232	0.0630	0.0086	0.0153	0.0285
Observations	2,154	7,730	19,480	18,279	5,136

Note: Absolute value of z-statistics in parentheses.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Table 2A.11 Probit Regressions on Mortality, 1860 (partial derivatives)

Variable	Infants (1)	Ages 1–4 (2)	Ages 5–19 (3)	Ages 20–44 (4)	Ages 45+ (5)
Log (real estate + \$1.00)	0.0024	0.0010	–0.0000	0.0002	0.0004
	(1.48)	(2.19)**	(0.26)	(0.99)	(0.74)
Log (personal estate + \$1.00)	–0.0060	–0.0006	–0.0003	–0.0005	0.0010
	(2.49)**	(0.88)	(1.51)	(2.00)**	(1.21)
Age		–0.0105	–0.0000	0.0001	0.0007
		(8.03)***	(0.14)	(1.35)	(4.08)***
Male	0.0173	0.0069	0.0004	–0.0014	0.0062
	(1.87)*	(2.75)***	(0.49)	(1.36)	(1.92)*
Family size	–0.0001	0.0003	0.0003	0.0010	0.0021
	(0.05)	(0.48)	(1.09)	(3.81)***	(2.73)***
Born in state of enumeration	0.0062	–0.0018	0.0019	0.0032	0.0217
	(0.29)	(0.44)	(2.03)**	(2.50)**	(1.91)*
Foreign-born		0.0397		–0.0000	–0.0003
		(1.56)		(0.02)	(0.04)
Head was farmer	–0.0105	–0.0042	–0.0005	–0.0020	–0.0078
	(0.96)	(1.32)	(0.54)	(1.66)*	(1.94)*
Head was laborer	–0.0031	0.0026	0.0006	–0.0021	–0.0049
	(0.20)	(0.54)	(0.34)	(1.18)	(0.67)
Illinois	0.0037	0.0026	–0.0001	–0.0004	–0.0065
	(0.37)	(0.93)	(0.12)	(0.34)	(1.91)*
Predicted probability	0.0477	0.0139	0.0033	0.0050	0.0136
Pseudo-R^2	0.0154	0.0648	0.0122	0.0255	0.0451
Observations	2,149	7,730	19,009	18,279	5,136

Note: Absolute value of z-statistics in parentheses.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Table 2A.12 Probit Regressions on Mortality, 1860 (partial derivatives)

Variable	Infants (1)	Ages 1–4 (2)	Ages 5–19 (3)	Ages 20–44 (4)	Ages 45+ (5)
Real estate					
$100–499	0.0218	0.0081	0.0010	0.0010	0.0003
	(1.40)	(1.78)*	(0.68)	(0.53)	(0.06)
$500–999	–0.0178	–0.0001	–0.0000	0.0032	0.0036
	(1.15)	(0.02)	(0.01)	(1.56)	(0.60)
$1,000–2,499	0.0198	0.0040	0.0002	0.0016	0.0113
	(1.24)	(0.94)	(0.18)	(0.94)	(1.99)**
$2,500–4,999	0.0201	0.0109	–0.0008	–0.0003	0.0015
	(0.93)	(1.75)*	(0.50)	(0.14)	(0.24)
$5,000	–0.0062	0.0328	–0.0015	–0.0018	0.0092
	(0.20)	(3.10)***	(0.77)	(0.74)	(1.10)
Personal estate					
$100–499	–0.0400	–0.0016	–0.0017	–0.0047	0.0118
	(2.74)***	(0.36)	(1.38)	(2.97)***	(1.45)
$500–999	–0.0136	–0.0067	–0.0015	–0.0041	0.0067
	(0.85)	(1.42)	(1.11)	(2.72)***	(0.75)
$1,000–2,499	–0.0285	–0.0045	–0.0026	–0.0027	0.0046
	(1.77)*	(0.86)	(1.86)*	(1.62)	(0.52)
$2,500–4,999	–0.0353	0.0004	–0.0012	–0.0036	0.0158
	(1.66)*	(0.05)	(0.62)	(1.79)*	(1.19)
$5,000+	–0.0305	–0.0030	–0.0016	–0.0010	0.0111
	(1.26)	(0.39)	(0.85)	(0.42)	(1.02)
Age		–0.0102	–0.0000	0.0001	0.0006
		(8.01)***	(0.07)	(1.31)	(4.04)***
Male	0.0162	0.0068	0.0004	–0.0014	0.0061
	(1.80)*	(2.73)***	(0.51)	(1.40)	(1.93)*
Family size	–0.0006	0.0003	0.0003	0.0010	0.0021
	(0.26)	(0.37)	(1.17)	(3.80)***	(2.76)***
Born in state of enumeration	0.0049	–0.0016	0.0019	0.0028	0.0207
	(0.24)	(0.39)	(2.06)**	(2.29)**	(1.86)*
Foreign-born		0.0408		–0.0003	–0.0008
		(1.61)		(0.14)	(0.12)
Head was farmer	–0.0143	–0.0034	–0.0006	–0.0017	–0.0072
	(1.29)	(1.13)	(0.60)	(1.43)	(1.82)*
Head was laborer	–0.0006	0.0018	0.0006	–0.0017	–0.0050
	(0.04)	(0.39)	(0.33)	(0.94)	(0.72)
Illinois	0.0020	0.0030	0.0003	0.0007	–0.0067
	(0.19)	(1.00)	(0.32)	(0.64)	(1.75)*
Predicted probability	0.0454	0.0134	0.0032	0.0048	0.0131
Pseudo-R^2	0.0300	0.0730	0.0163	0.0334	0.0534
Observations	2,149	7,730	19,009	18,279	5,136

Note: Absolute value of *z*-statistics in parentheses.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Table 2A.13 Profit Regressions on Mortality, 1860 (partial derivatives)

Variable	Infants (1)	Ages 1–4 (2)	Ages 5–19 (3)	Ages 20–44 (4)	Ages 45+ (5)
Log(real estate + $1.00)	0.0054	0.0015	–0.0002	0.0010	0.0040
	(1.65)*	(1.51)	(0.60)	(2.48)**	(3.18)***
(Illinois) · log(real estate + $1.00)	–0.0040	–0.0008	0.0002	–0.0009	–0.0049
	(1.08)	(0.70)	(0.70)	(1.94)*	(3.42)***
Log(personal estate + $1.00)	–0.0099	–0.0009	–0.0003	–0.0004	–0.0004
	(2.63)***	(0.75)	(1.03)	(1.07)	(0.36)
(Illinois) · log(personal estate + $1.00)	0.0049	–0.0001	0.0000	–0.0003	0.0019
	(1.01)	(0.09)	(0.07)	(0.51)	(1.08)
Age		–0.0101	0.0000	0.0000	0.0006
		(3.93)***	(0.02)	(0.25)	(2.88)***
(Illinois) • age		–0.0006	–0.0001	0.0001	–0.0002
		(0.19)	(0.48)	(0.81)	(0.48)
Illinois	–0.0049	0.0080	–0.0007	0.0001	0.0174
	(0.22)	(0.95)	(0.23)	(0.02)	(0.87)
Predicted probability	0.0481	0.0145	0.0033	0.0054	0.0136
Pseudo-R^2	0.0115	0.0558	0.0064	0.0126	0.0401
Observations	2,154	7,730	19,480	18,279	5,136

Note: Absolute value of z-statistics in parentheses.
***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

References

Case, Anne, Darren Lubotsky, and Christina Paxson. 2001. Economic status and health in childhood: The origins of the gradient. NBER Working Paper no. 8344. Cambridge, Mass.: National Bureau of Economic Research, June.

Chapin, Charles V. 1924. Deaths among taxpayers and non-taxpayers, income tax, Providence, 1865. *America Journal of Public Health* 14:647–51.

Condran, Gretchen A., and Eileen Crimmins. 1979. A description and evaluation of mortality data in the Federal Census: 1850–1900. *Historical Methods* 12:1–23.

Craig, Lee, Raymond Palmquist, and Thomas Weiss. 1998. Transportation improvements and land values in the antebellum United States: A hedonic approach. *Journal of Real Estate Finance and Economics* 16 (March): 173–89.

Currie, Janet, and Mark Stabile. 2002. Socioeconomic status and health: Why is the relationship stronger for older children? NBER Working Paper no. 9098. Cambridge, Mass.: National Bureau of Economic Research, August.

Fogel, Robert W., and Stanley L. Engerman. 1974. *Time on the cross: Evidence and methods.* Boston: Little, Brown.

Hahn, Marilyn Davis. 1983. *Alabama mortality schedule, 1850.* Easley, S.C.: Southern Historical Press.

———. 1987. *Alabama mortality schedule, 1860.* Easley, S.C.: Southern Historical Press.

Haines, Michael. 1998. Estimated life tables for the United States, 1850–1910. *Historical Methods* 31 (Fall): 149–69.

Haines, Michael, Lee Craig, and Thomas Weiss. 2000. Development, health, nutrition, and mortality: The case of the "antebellum puzzle" in the United States. NBER Working Paper no. H0130. Cambridge, Mass.: National Bureau of Economic Research, October.

Jackson, Ronald V. 1999. Accelerated indexing systems, compiler. AIS mortality schedules index [database on-line]. Provo, Utah: Ancestry.com. Available at http://ancestry.com/search/rectype/inddbs/3530a.htm.

Jacobson, Paul N. 1957. An estimate of the expectation of life in the United States in 1850. *Milbank Memorial Fund Quarterly* 35 (April): 197–201.

Kitagawa, Evelyn M., and Philip M. Hauser. 1973. *Differential mortality in the United States: A study in socioeconomic epidemiology.* Cambridge: Harvard University Press.

Lantz, Paula M., James S. House, James M. Lepowski, David R. Williams, Richard P. Mero, and Jieming Chen. 1998. Socioeconomic factors, health behaviors, and mortality: Results from a nationally representative prospective study of U.S. adults. *Journal of the American Medical Association* 279 (3 June): 1703–08.

Margo, Robert A. 2000. *Wages and labor markets in the United States, 1820–1860.* Chicago: University of Chicago Press.

Preston, Samuel H., and Michael R. Haines. 1991. *The fatal years: Child mortality in late nineteenth-century America.* Princeton, N.J.: Princeton University Press.

Smith, James P. 1999. Healthy bodies and thick wallets: The dual relation between health and economic status. *Journal of Economic Perspectives* 13 (Spring): 145–66.

Steckel, Richard H. 1988. The health and mortality of women and children, 1850–1860. *Journal of Economic History* 48 (June): 333–45.

Volkel, Lowell M. 1972. *Illinois mortality schedule, 1850.* N.p.

———. 1979. *Illinois mortality schedule, 1860.* Indianapolis, Ind.: Heritage House.

Williams, Redford B. 1998. Lower socioeconomic status and increased mortality: Early childhood roots and the potential for successful interventions. *Journal of the American Medical Association* 279 (3 June): 1745–46.