NBER WORKING PAPER SERIES
ON
HISTORICAL FACTORS IN LONG RUN GROWTH

THE POOR AND THE DEAD: SOCIOECONOMIC STATUS AND MORTALITY IN THE U.S., 1850-1860

Joseph P. Ferrie

Historical Paper 135

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2001

The author thanks participants at Northwestern University's Economic History Workshop and Institute for Policy Research Faculty Seminar, the 2001 American Economic Association Meetings, and the NBER Conference on Health and Labor Force Participation Over the Life Cycle, particularly the discussants Richard Steckel and Irwin Rosenberg, and Deirdre McCloskey for extremely helpful comments. The views expressed herein are those of the author and not necessarily those of the National Bureau of Economic Research.

The Poor and the Dead: Socioeconomic Status and Mortality in the U.S., 1850-1860
Joseph P. Ferrie
NBER Historical Paper No. 135
July 2001
JEL No. N3, I1, J1
Development of the American Economy

ABSTRACT

Despite the significant research on aggregate trends in mortality and physical stature in the middle of the nineteenth century, little evidence on the individual-level characteristics associated with premature mortality has been presented. This essay describes a new project that links individuals from the mortality schedules to the population schedules of the 1850 and 1860 federal population censuses. This makes it possible to assess the link between individual and household characteristics and the probability of dying. The results reveal a strong and negative relationship between household wealth and mortality in 1850 and 1860 and a somewhat weaker negative relationship between occupational status and mortality in 1850. The findings suggest that even when the U.S. population was largely rural and agricultural, changes in the distribution of income and wealth would have had a large impact on mortality rates and life expectancies. Urbanization merely exacerbated already existing disparities in mortality by socioeconomic status.

Joseph P. Ferrie
Department of Economics
and Institute for Policy Research
Northwestern University
Evanston, IL 60208-2600
and NBER
ferrie@northwestern.edu

"*The Poor and the Dead: Socioeconomic Status and Mortality in the U.S., 1850-1860*"

Joseph P. Ferrie
Department of Economics
and Institute for Policy Research
Northwestern University
and
NBER

July 13, 2001

Introduction

Economic growth in the nineteenth century U.S. was not without its costs. The deleterious impact of economic growth on nineteenth century U.S. living standards can be seen in what has come to be known as "the antebellum paradox": though economic growth was strong and per capita incomes were rising before the Civil War, two indicators of living standards—average height

The author thanks participants at Northwestern University's Economic History Workshop and Institute for Policy Research Faculty Seminar, the 2001 American Economic Association Meetings, and the NBER Conference on Health and Labor Force Participation Over the Life Cycle, particularly the discussants Richard Steckel and Irwin Rosenberg, and Deirdre McCloskey for extremely helpful comments.

and life expectation at age 10 (e_{10})—were declining (Fogel, 1986). Several studies have examined the factors that influenced physical stature at the individual level in cross section, in order to understand how changes over time in urbanization, industrialization, and nutrition could explain the change in stature. But falling life expectation has received considerably less attention at the micro level.

This study introduces new evidence on the correlates of individual-level mortality, particularly socioeconomic status measured by occupation and household wealth, created by linking the mortality and population schedules of the 1850 and 1860 federal population censuses. The linkage makes it possible to calculate mortality rates by occupation, by household wealth, and by cause of death, and thereby uncover in two cross sections some of the factors underlying the changes in life expectation observed over the decades before the U.S. Civil War.

Research on the link between status and mortality in the late twentieth century U.S. has uncovered a substantial gap between high and low status individuals, a gap that is apparent in other health outcomes as well (Williams, 1998; Lantz *et al.*, 1998). Though a great deal of attention has now been devoted to explaining why those lower in status have worse outcomes, and the possibility that there is some reverse causation involved (with poor health leading to low status), such investigations lack an historical perspective (Smith, 1999). For example, though wide disparities in mortality rates by status were described in the early 1970s (Kitagawa and Hauser, 1973), we simply do not know whether the disparities observed over the last quarter century are large or small by historical standards. It is entirely possible that 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. This casts in an altogether different light the disparities observed today. Perhaps they are *not* merely the continuation of poor outcomes for poor people that generations have failed to erase.

I. What We Know About 19th Century Socioeconomic Status & Mortality

There is a consensus 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. Only three published studies and one on-going research project have attempted to assess the link between status and mortality for the second half of the nineteenth century.

The first of these studies estimated crude death rates in 1865 Providence, Rhode Island, with a comparison of the rates for taxpayers and non-taxpayers (Chapin, 1924). The annual crude death rate for taxpayers was 11 per thousand, while the corresponding rate for non-taxpayers was 25 per thousand. Though 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, though 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 household head (Preston and Haines, 1991, pp. 156-56). Though 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.

Though 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 household where their mother was enumerated had died or simply moved out. This study is also somewhat limited in the components of socioeconomic status that it can examine: though the household head's occupation was recorded, there was no information collected in the 1900 census on the value of the household's wealth. 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 study is that, like the 1865 Providence, Rhode Island study, it provides information at only one date (1900). Though 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 than 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 U.S. 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 the Preston and Haines study, the study 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 Early Indicators Project currently underway at the University of Chicago's Center for Population Economics, under the direction of Robert Fogel, is using information from Civil War military pension records to assess the link between socioeconomic status (among other factors) and later disability and premature mortality. Though 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. Further, it is limited to the northern population. The present study complements this work: though the data on causes of death is less precise, it covers the populations and regions missed in the Civil War 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, 2000, p. 8). Though 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.

II. The Data

This project links decedents from the mortality schedules of the 1850 and 1860 federal population censuses to the population schedules of those censuses. 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. Though 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, 1994), the data have never been examined at the individual level. Several difficulties have prevented their full exploitation.[1]

[1] These difficulties are summarized in Condran and Crimmins (1979).

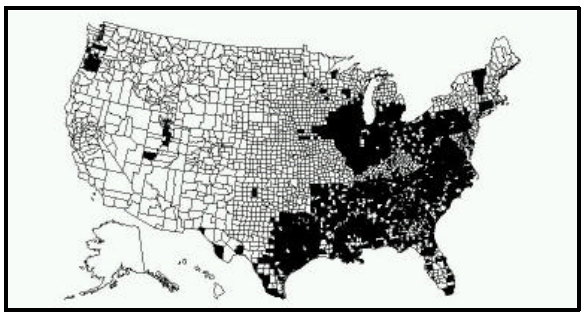

Figure 1: Counties for which 1850 Mortality Schedules Have Been Transcribed.

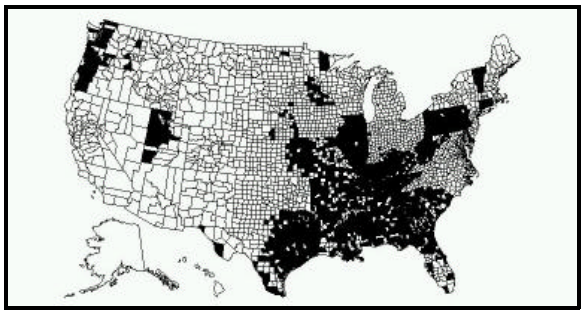

Figure 2: Counties for which 1860 Mortality Schedules Have Been Transcribed.

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 some states have not survived, some have not been microfilmed, and none had been available in machine-readable form until recently. Since the late 1980s, entries for just over 400,000 decedents from the 1850 through 1880 mortality schedules have been transcribed and computerized. These computerized transcriptions of the mortality schedules contain all of the information as it appeared on the forms submitted to the census office by census marshals.

Figures 1 and 2 show the counties for which the mortality schedules have been transcribed. For 1850 and 1860, the mortality schedules cover a significant fraction of the total U.S. Table 1 shows several records from the 1850 mortality schedules from Washington, D.C. to illustrate the range of information available from this source. There are four significant 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 under-reported, and that overall mortality is probably 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 under-enumeration 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"), but can be remedied to some extent by grouping diseases into several broad

Name	Age	Occupation	Sex	Month of Death	Cause of Death	Birthplace
Altamus, Thomas	1	None	M	Feb	whooping cough	DC
Anderson, Lewis	50	Laborer	M	Jan	smallpox	DC
Augustus, Milly	27	Slave	F	Mar	scarlet fever	MD
Duvall, Thomas	31	Coachman	M	Feb	pneumonia	MD

Table **1**: Sample Records from Mortality Schedules (1850) for Washington, D.C.

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 problems only if under-reporting or mis-reporting varies by status differently in the mortality and population schedules.

Though 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. After decedents from the 1860 mortality schedules were matched to their surviving households in the 1860 population schedules, the distribution of the months in which deaths occurred was calculated for high (total wealth>0) and low (total wealth=0) status households. Figure 3 shows that the distributions are similar except at ten and eleven months prior to the census[2]. Though the

[2] The month of death field in the computerized mortality schedules was only two characters in width, so it was not possible to distinguish deaths in June from those in July, or to distinguish deaths in March from those in May.

overall distributions are statistically indistinguishable, there may be some under-reporting of deaths by low status households, so if the results indicate a negative relationship between wealth and

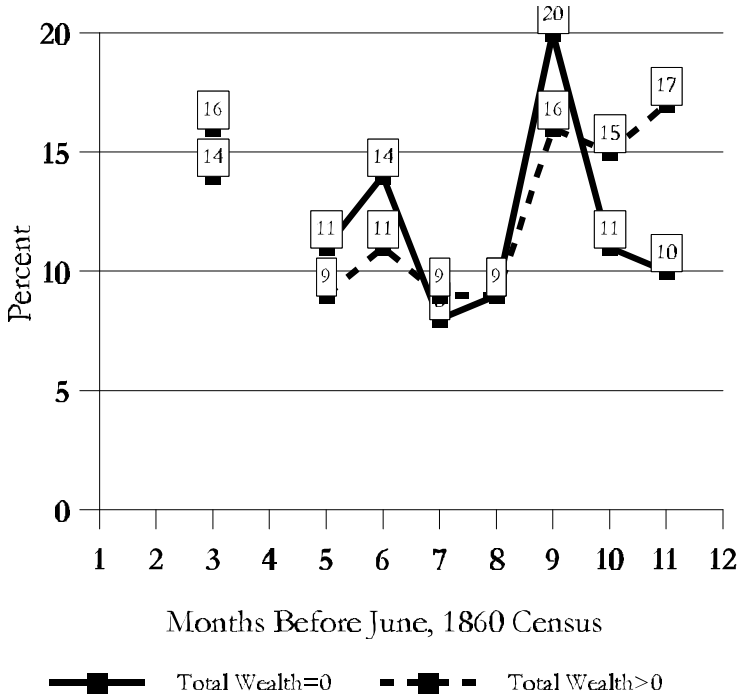

Figure 3. Distribution of Months of Death, 1860. For "Total Wealth=0," N=118; for "Total Wealth>0," N=392. The χ^2 statistic for the homogeneity of the two distributions is 5.5004 (p=0.5991).

mortality, the estimated magnitude may be biased downward from its true value. To the extent that there is a bias, then, it is against finding a negative relationship between status and mortality.

The advantages of using individual observations from the mortality schedules more than outweigh the shortcomings. For example, when combined with the information on household 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 covered 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 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 shows the information relating to status than can be obtained from the 1850-60 population schedules. Each piece of information is reported for each surviving member of the household.

1850	1860
occupation	occupation
real estate wealth	real estate wealth
	personal wealth
literacy	literacy
school attendance	school attendance
pauper	pauper
criminal	criminal
disabled	disabled

Table **2**: Variables in Population Schedules Related to SES.

The linkage process was guided by two considerations:

1. The census marshals collected the mortality information from households in the same order as that in which they collected the population information; comparing the mortality records sorted by order of visitation with similarly sorted population records makes it possible to link even individuals with the most common names.

2. In order to calculate mortality rates or perform multivariate analysis on the correlates of mortality, it is necessary to have either a reliable sample or a complete enumeration of the entire population of a county to which the mortality data can be linked.

The first consideration meant that the linkage of individual decedents to their surviving households could be done easily only where a source allowed the computerized mortality records to be sorted by order of visitation. Published transcriptions of the mortality schedules available for several states were used for this purpose and allowed decedents in these states to be linked back to their actual households. The second consideration further limited the analysis to those counties for which the 1850 or 1860 population schedules have been entirely transcribed. For 1850, it was possible to link decedents to surviving households for five counties in Illinois (Morgan, Jackson, Union, Saline, and Washington), one in Indiana (Howard), one in Alabama (Shelby), and two in Texas (Shelby and Ellis). For 1860, it was possible for three counties in Illinois (Perry, Shelby, and Vermilion) and two in Alabama (St. Clair and Tuscaloosa). The locations of these counties are shown in Figures 4 through 7.

For 1850, the presence of a large number of counties for which transcriptions of the census exist but for which it was not possible to sort the households on the mortality schedules in the order of their visitation suggested an additional form of linkage, in addition to linking individuals back to their surviving households. In this additional analysis, individual male decedents over age 15 whose occupations were reported in the mortality schedules were compared with individual survivors over age 15 whose occupations were reported in the population schedules from the same counties. Fifty counties were selected for which the 1850 population schedules have been entirely transcribed and for which decedents were included in the computerized mortality database. These counties are shown in Figures 8 through 18.

For the sample of decedents linked back to their households in 1850 and 1860, the linked data contains all the information for each decedent in the mortality schedules if the household experienced a death in the year preceding the census, as well as the following information for all

surviving members of their household and all members of households that did not experience a death (in addition to the status-related information shown in Table 2): age, sex, race, birthplace, household wealth, and occupation of the household head. The cause of death is also included for decedents.

For 1860, the overall linkage rate was 85.4%, but a year of age reduced the linkage rate by 0.17 percentage points (p=0.032), males were 4.3 percentage points more likely to be matched (p=0.096), the linkage rate in the three Illinois counties was 7.2 percentage points below that in the two Alabama counties (p=0.009), and those born outside their state of residence at death were 5.9 percentage points more likely to be linked (p=0.118). The linked sample contains 582 decedents and 59,809 survivors in 1850 and 758 decedents and 66,093 survivors in 1860.

For the 1850 analysis of working males, the linkage produced a sample of 810 adult male decedents with reported occupations and 92,755 adult male survivors with reported occupations in the 50 county area. The sample contains each individual's occupation, age, and birthplace, and the reported cause of death for decedents.

In both 1850 and 1860, the counties for which the analysis can be performed were determined largely by where genealogists had transcribed mortality and population schedules. These counties are uniformly rural. In 1850, only 4 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.[3]

III. Analysis of Socioeconomic Status and Mortality in 1850 & 1860

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 second is that in 1850 and 1860, individuals in households with greater wealth had lower mortality rates than individuals in households with less wealth. The exact mechanism through which these relationships operate will not be tested, but it seems reasonable to imagine that higher status individuals and households 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.

Though it will be necessary to control for a variety of individual, household, and community characteristics in addition to occupation and wealth, as well as to consider different causes of death separately, it will be useful to calculate some simple mortality rates without these controls as a first step. Table 3 presents *t*-tests on differences in mortality rates by the individual's occupation in 1850. Occupations are grouped into 4 broad categories: white collar (professional, managerial, clerical and sales, and government), craft, farmer, and laborer (including operatives and unskilled

[3] The 1860 mortality and population schedules for Albany, New York have been linked by David Davenport, and the author has linked most of the 1860 mortality and population schedules for Chicago, but both places presently lack a comparison population of survivors.

workers). Farmers and white collar workers had higher incomes than craft workers, who in turn had higher incomes than laborers. No attempt is made here to calculate separate mortality rates by cause of death.

The results provide only weak support for the hypothesis that occupational status (as a proxy for income) is a significant determinant of mortality rates: white collar workers did indeed have lower mortality rates than craft workers, but craft workers had substantially higher mortality rates than common laborers. There is little evidence that farmers had lower mortality rates than any of the other occupational groups. White collar workers had the lowest rates, followed by laborers, farmers, and craft workers.

There are several other likely influences on mortality rates that can be accounted for in order to isolate the role of status. The most important of these is obviously age. It is also possible to identify individuals born outside the state in which they resided at the time of the census. Such

	Mortality Rate (per thousand)	Difference (1 vs. 2)	*t*-statistic	probability	obs.
1850					
1. Farmer	8.6				
2. White Collar	7.1	1.5	1.25	0.211	63,037
1. Farmer	8.6				
2. Laborer	8.1	0.5	0.64	0.522	74,959
1. White Collar	7.1				
2. Laborer	8.1	-1.0	-0.73	0.464	22,090
1. Farmer	8.6				
2. Craft	10.1	-1.5	-1.59	0.110	71,475
1. White Collar	7.1				
2. Craft	10.1	-3.0	-1.89	0.059	18,606
1. Craft	10.1				
2. Laborer	8.1	2.0	1.79	0.074	30,528

Table **3**. Comparisons of Mortality Rates by Occupation, 1850.

people may have lower mortality if the process of in-migration selects for the most physically fit, 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 measures of both population density and the availability of transportation. The multivariate analysis includes both of these county level variables as well as regional dummies for the west (Indiana, Michigan, Illinois, Iowa, and Texas) and south (North Carolina, Virginia, Alabama, and Kentucky).

	Dependent Variable:		
Variable	Death From Any Cause	Death From Consumption	Death From Cholera
Age 25-34	0.0002	0.0001	0.0005**
Age 35-44	0.0021**	-0.0001	0.0004
Age 45-54	0.0064***	0.0002	0.0005
Age 55+	0.0178***	0.0016**	0.0010*
White Collar	-0.0016	0.0002	-0.0001
Farmer	-0.0018**	-0.0005**	-0.0001
Laborer	-0.0009	-0.0004**	0.0001
Born In State	0.0015**	0.0002	0.0001
South	-0.0039***	-0.0011***	-0.0006***
West	-0.0012*	-0.0006***	0.0001
Density	-0.0001***	-0.0000	-0.0001**
Transportation Access	0.0001	0.0002	-0.0002
Predicted Probability	0.0079	0.0009	0.0005
Pseudo R^2	0.0180	0.0282	0.0513
Observations	93,565	92,864	92,830

Note: The figures shown are partial derivatives. The sample consists of males age 15 to 65 who reported occupations in either the mortality or population schedules. For "Consumption" and "Cholera," only deaths from these causes and survivors are included; other deaths are excluded. Omitted categories for the categorical variables are "Age 15-24," "Craftsman," "Born Outside State of Residence," "North," "East," 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. Probability that the true partial derivative is zero: *** < 0.001 ** < 0.05 * < 0.10

Table **4**. Logistic Regressions on Mortality, 1850.

The first column of Table 4 presents a logistic regression with death (from any cause) as the dependent variable. As expected, there is a clear age pattern: the risk of death increases slightly

from the 15-24 category (the excluded group) to the 25-34 group, then increases by a greater amount in the 45-54 group, then nearly doubles in the age 55 and over group. Death rates were also lower in the south and the west. The occupation categories now reveal a pattern somewhat more consistent with a positive influence of status on mortality: the largest coefficients are for farmers and white collar workers (though only the former is statistically significant), and craft workers and common laborers are similar in their mortality. The density measure (population per square mile in the county) suggests that more densely populated places actually had lower mortality rates. Though this seems contrary to an impression of nineteenth century urban places as sites of prevalent disease, overcrowding, poor sanitation, and high death rates, keep in mind that the counties included in the analysis are all predominantly rural, while the small number of urban places within them are quite modest in size. The coefficient on density suggests, then, that within the range of densities found in these rural counties, mortality was lower in counties with more people per square mile (and thus more people living in small towns) than in places more like the frontier with very low densities and few amenities.

It is straightforward to identify two specific causes of death in the mortality schedule: consumption was the most frequently reported cause of death throughout the nineteenth century, while cholera was epidemic throughout the U.S. in 1849 (the second half of which falls within the twelve month window of the question asked by census marshals as they canvassed households). The second and third columns of Table 4 examine deaths separately from these two causes. For consumption, there are clear differences by occupation, though not the differences we would anticipate if occupation captured only the influence of income and its corresponding access to superior food, shelter, and sanitation. Rather, both farmers and laborers enjoyed lower consumption death rates than white collar and craft workers. This may reflect the importance of

the workplace (generally outdoors for farmers and common laborers, most of whom in these counties would have been farm laborers) and fewer workplace opportunities to come into direct contact with other people for these occupations rather than the role of the home environment in the disease's transmission. The age pattern of mortality was similar in magnitude to that for all causes of death, though statistically significant for consumption only in the oldest age cohort, perhaps as a consequence of the greater susceptibility of those already debilitated in some way. Lower consumption death rates prevailed in the south and the west.

The regression for deaths from cholera reveals no statistically significant relationship to occupation, though the coefficients suggest an ordering of death rates from laborers (highest) to

	Mortality Rate (per thousand)	Difference (1 vs. 2)	*t*-statistic	probability	obs.
1850					
1. Real Wealth > 0	12.3				
2. Real Wealth = 0	8.1	4.2	5.13	0.001	60,391
1860					
1. Total Wealth > 0	9.7				
2. Total Wealth = 0	27.5	-17.8	-12.64	0.001	66,851
1. Real Wealth > 0	10.1				
2. Real Wealth = 0	13.8	-3.8	-4.36	0.001	66,851
1. Personal Wealth > 0	9.5				
2. Personal Wealth = 0	26.8	-17.3	-12.97	0.001	66,851

Table **5**. Comparisons of Mortality Rates by Wealth, 1850 and 1860.

craft workers to farmers to white collar workers (lowest). That this disease is only weakly (if at all) related to occupation as a proxy for material living conditions is not surprising. Cholera is spread by contaminated water supplies (and also the consumption of contaminated shellfish), so it is quite

	Dependent Variable:	
Variable	Death From Any Cause	Death From Consumption
Age	-0.0014***	0.0000
$Age^2 \times 10^{-2}$	0.0043***	0.0000
$Age^3 \times 10^{-4}$	-0.0033***	0.0000
Male	0.0016***	0.0000
Born In State	-0.0045***	-0.0005**
Morgan County, IL	-0.0069***	
Jackson County, IL	-0.0012	0.0001
Union County, IL	-0.0027**	-0.0000
Saline County, IL	-0.0039***	-0.0002
Washington County, IL	-0.0009	-0.0000
Howard County, IN	-0.0018	0.0003
Shelby County, AL	-0.0040***	-0.0000
Shelby County, TX	-0.0011	-0.0000
Log(Real Wealth+$1.00)	-0.0004***	-0.0001*
Predicted Probability	0.0069	0.0003
Pseudo R^2	0.0574	0.0982
Observations	60,391	44,048

Note: 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 linked to households in the population schedules. For "Consumption," only deaths from this cause and survivors are included; other deaths are excluded. Wealth is measured at the household level. Omitted categories for the categorical variables are "Female," "Born Outside State of Residence," and "Ellis County, TX." Probability that the true partial derivative is zero: *** < 0.001 ** < 0.05 * < 0.10

Table **6**. Logistic Regressions on Mortality, 1850.

possible for an otherwise well-fed, well-housed, well-clothed individual to contract the disease.

Results from the 1850 and 1860 samples of individual decedents linked back to their surviving households are shown in Tables 5 thru 8. Simple comparisons of those in households with and without wealth (Table 5) reveal substantial differences in mortality by status. In both 1850 and 1860, individuals in households with no real estate wealth had mortality rates more than a third higher than individuals in households with any real estate wealth. In 1860, when it is possible to assess separately the impact of real and personal wealth, the impact of personal wealth is even

greater: mortality rates for individuals in households without personal wealth were 2.8 times higher than for individuals in households with any personal wealth.

In the multivariate regressions (Tables 6 and 7), for all causes of death in both years, the impact of wealth on mortality is large and statistically significant. The impact of real estate wealth is consistent in magnitude between 1850 and 1860, as can be seen by comparing column 1of Table 6 and column 2 of Table 7. In both 1850 and 1860, the impact of wealth on mortality can be seen for specific causes of death, even after controlling for an individual's personal and community characteristics: wealth (real estate wealth in 1850, total wealth in 1860) reduces mortality in the last columns of both Table 6 and Table 7.

The finding that personal wealth has a much more substantial impact on mortality than real estate wealth is borne out in the third column of Table 7: the impact of personal wealth on

	Dependent Variable:			
Variable	Death From Any Cause			Death From Consumption
Age	-0.0016***	-0.0016***	-0.0016***	0.0000***
$Age^2 \times 10^{-2}$	0.0047***	0.0048***	0.0046***	-0.0001**
$Age^3 \times 10^{-4}$	-0.0032***	-0.0033***	-0.0032***	0.0001
Male	0.0007	0.0007	0.0007	-0.0001
Born In State	0.0044***	0.0044***	0.0043***	0.0002**
Perry County, IL	0.0016	0.0015	0.0016	-0.0000
Shelby County, IL	0.0010	0.0001	0.0010	0.0004
Vermilion County, IL	0.0028***	0.0021**	0.0026***	0.0005*
St. Clair County, AL	0.0030***	0.0022*	0.0034***	0.0002
Log(Total Wealth + $1.00)	-0.0009***			-0.0001*
Log(Real Wealth + $1.00)		-0.0003***		
Log(Pers. Wealth + $1.00)			-0.0011***	
Predicted Probability	0.0074	0.0078	0.0074	0.0004
Pseudo R^2	0.0793	0.0700	0.0805	0.0738
Observations	66,851	66,851	66,851	66,150

Note: 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 linked to households in the population schedules. Total Wealth is measured at the household level. Omitted categories for the categorical variables are "Female," "Born Outside State of Residence," and "Tuscaloosa County, AL." Probability that the true partial derivative is zero: *** < 0.001 ** < 0.05 * < 0.10

Table **7**: Logistic Regressions on Mortality, 1860.

Dependent Variable: Death From Any Cause						
Variable	Infants	Age 1-4	Age 5-19	Age 20-44	Age 45-64	Age 65+
Age		-0.0112***	-0.0000	0.0002***	0.0009***	0.0023***
Male	0.0117	0.0057***	-0.0000	-0.0023**	0.0044	0.0002
Born In State	0.0200	0.0026	0.0024***	0.0043***	0.0136	0.0473
Perry County, IL	0.0320	0.0139**	-0.0010	-0.0014	0.0002	
Shelby County, IL	-0.0065	0.0054	0.0013	0.0009	-0.0047	-0.0007
Vermilion County, IL	0.0355**	0.0104**	0.0007	0.0007	0.0020	-0.0107
St. Clair County, AL	0.0128	0.0077	0.0010	0.0016	0.0061	0.0308*
Log(Total Wealth + $1.00)	-0.0063***	-0.0021***	-0.0005***	-0.0008***	-0.0005	-0.0026*
Predicted Probability	0.0522	0.1654	0.0035	0.0064	0.0131	0.0331
Pseudo R^2	0.0279	0.0751	0.0267	0.0232	0.0272	0.0729
Observations	2,531	9,290	26,194	22,258	5,500	952

Note: 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 linked to households in the population schedules. Total Wealth is measured at the household level. Omitted categories for the categorical variables are "Female," "Born Outside State of Residence," and "Tuscaloosa County, AL." Probability that the true partial derivative is zero: *** < 0.001 ** < 0.05 * < 0.10

Table **8**: Logistic Regressions on Mortality, 1860.

mortality is nearly four times greater than the impact of real estate wealth on mortality. This difference 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 coefficient on total wealth in the 1860 regression for infants (Table 8, first column), for example, implies that a change in income from the first quartile of the 1860 household distribution of wealth ($250) to the third quartile ($2900) reduces infant mortality by 15 per thousand, a drop of nearly a third at the mean predicted probability. Though these figures are well below previous estimates of infant mortality in the mid-nineteenth century U.S., reflecting the under enumeration of infant deaths in the mortality schedules, the difference between the rates for low and high wealth households is striking. As age increases, the effect of total wealth on mortality falls initially (from -0.0063 for infants to -0.0021 for age 1 to 4, to -0.0005 for age 5 to 19), then rises (to -0.0008 for age 20-44). The greatest impact, then, is for infants, which is consistent with a substantial role for material resources earliest in life.

Conclusions and Extensions

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-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.

Nonetheless, the results for both 1850 and 1860 support the view that 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. The results for wealth are much stronger than those for occupation as a proxy for status and are quite large in magnitude,

with a change in wealth from the first to the third quartile of the household wealth distribution leading to a dramatic one third drop in infant mortality. 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. Though these disparities may have widened with this movement, the mortality experience and its relationship to socioeconomic status after 1860 were different only in degree and not in kind from that seen here. Even on farms and in small towns, the more affluent experienced longer lives than their poorer neighbors.

References

Chapin, C.V. "Deaths Among Taxpayers and Non-Taxpayers, Income Tax, Providence, 1865." *America Journal of Public Health* 14 (1924): 647.

Condran, Gretchen A., and Crimmins, Eileen. "A Description and Evaluation of Mortality Data in the Federal Census: 1850-1900," *Historical Methods* 12 (1979): 1.

Craig, Lee, Palmquist, Raymond, and Weiss, Thomas. "Transportation Improvements and Land Values in the Antebellum United States: A Hedonic Approach." *Journal of Real Estate Finance and Economics* 16 (March 1998): 173-189.

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

Haines, Michael. "Estimated Life Tables for the United States, 1850-1900." NBER Working Paper H0059 (September 1994).

Haines, Michael, Craig, Lee, and Weiss, Thomas. "Development, Health, Nutrition, and Mortality: The Case of the 'Antebellum Puzzle' in the United States." NBER Working Paper H0130 (October 2000).

Kitagawa, Evelyn M., and Hauser, Philip M. *Differential Mortality in the United States: A Study in Socioeconomic Epidemiology* (Cambridge: Harvard University Press, 1973).

Lantz, P.M., House, J.S., Lepowski, J.M., Williams, D.R., Mero, R.P., and Chen, J. "Socio-economic Factors, Health Behaviors, and Mortality; Results from a Nationally Representative

Prospective Study of U.S. Adults." *Journal of the American Medical Association* 279 (June 3, 1998): 1703-1708.

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

Smith, James P. "Healthy Bodies and Thick Wallets: The Dual Relation Between Health and Economic Status." *Journal of Economic Perspectives* 13 (Spring 1999): 145-166.

Steckel, Richard H. "The Health and Mortality of Women and Children, 1850-1860." *The Journal of Economic History* 48 (June 1988): 333.

Williams, Redford B. "Lower Socioeconomic Status and Increased Mortality: Early Childhood Roots and the Potential for Successful Interventions." *The Journal of the American Medical Association* 279 (June 3, 1998): 1745.

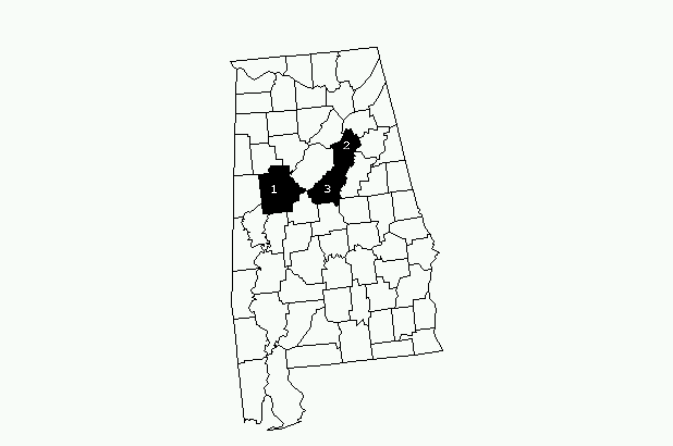

Figure 4. Alabama counties in 1850 and 1860 analysis by wealth: (1) Tuscaloosa, (2) St. Clair, (3) Shelby.

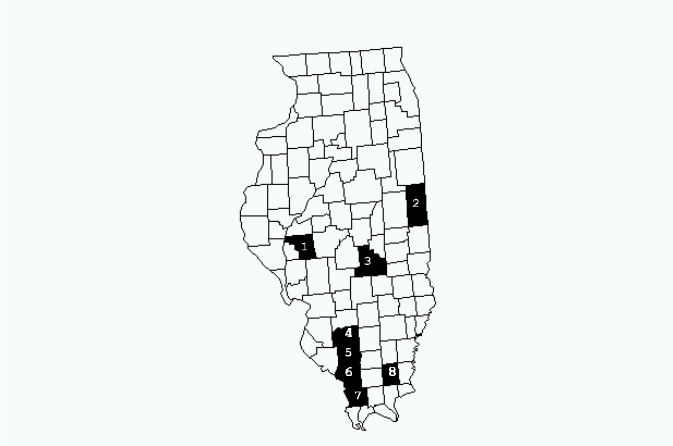

Figure 5. Illinois counties in 1850 and 1860 analysis by wealth: (1) Morgan, (2) Vermilion, (3) Shelby, (4) Washington, (5) Perry, (6) Jackson, (7) Union, (8) Saline.

Figure 6. Indiana counties in 1850 and 1860 analysis by wealth: (1) Howard.

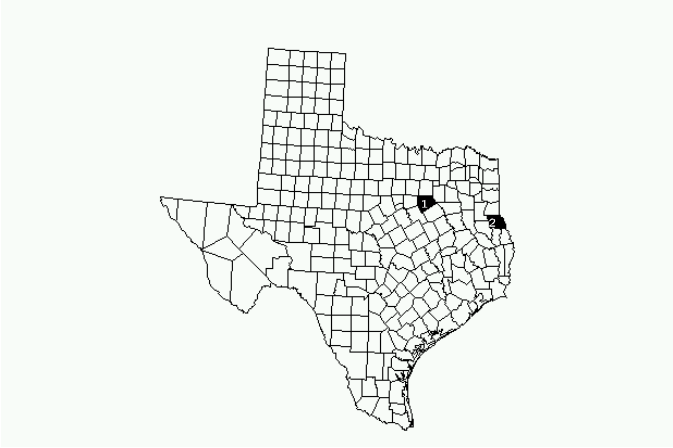

Figure 7. Texas counties in 1850 and 1860 analysis by wealth: (1) Ellis, (2) Shelby.

Figure 8. Alabama counties in 1850 analysis by occupation: (1) Madison, (2) Jackson, (3) Blount, (4) Jefferson, (5) Shelby, (6) Marengo, (7) Lowndes, (8) Wilcox, (9) Monroe, (10) Conecuh, (11) Washington, (12) Baldwin, (13) Henry.

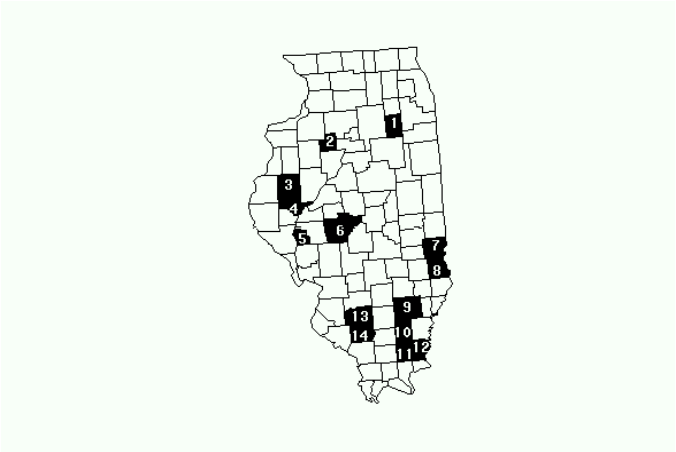

Figure 9. Illinois counties in 1850 analysis by occupation: (1) Grundy, (2) Stark, (3) McDonough, (4) Schuyler, (5) Scott, (6) Sangamon, (7) Clark, (8) Crawford, (9) Wayne, (10) Hamilton, (11) Saline, (12) Gallatin, (13) Washington, (14) Perry.

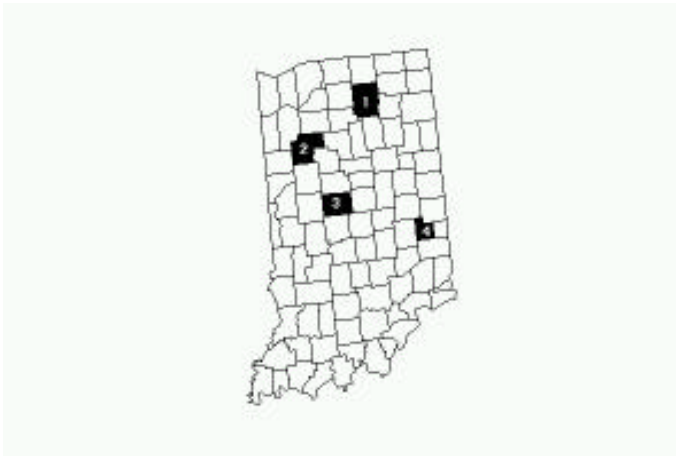

Figure 10. Indiana counties in 1850 analysis by occupation: (1) Kosciusko, (2) White, (3) Boone, (4) Fayette.

Figure 11. Iowa counties in 1850 analysis by occupation: (1) Cedar, (2) Appanoose.

Figure 12. Kentucky counties in 1850 analysis by occupation: (1) Spencer, (2) Simpson.

Figure 13. Michigan counties in 1850 analysis by occupation: (1) Lapeer, (2) Ionia.

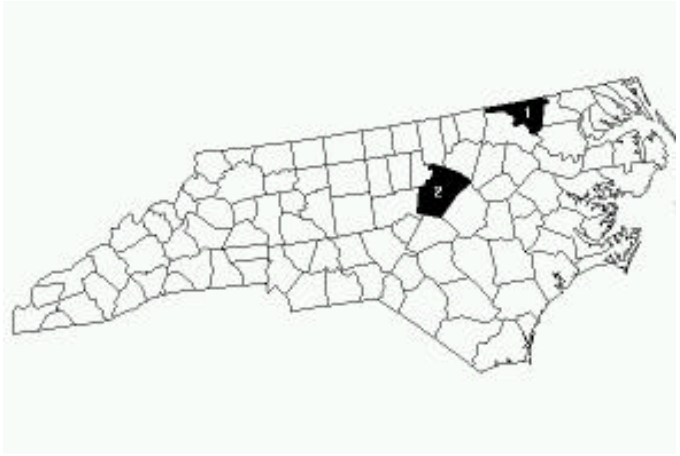

Figure 14. North Carolina counties in 1850 analysis by occupation: (1) Northampton, (2) Wake.

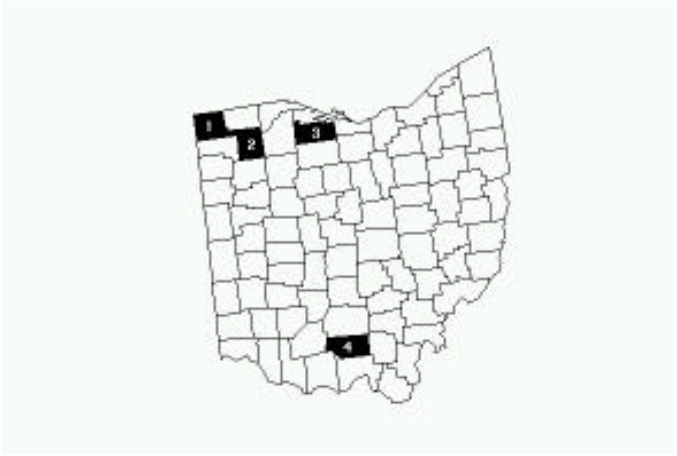

Figure 15. Ohio counties in 1850 analysis by occupation: (1) Williams, (2) Henry, (3) Sandusky, (4) Pike.

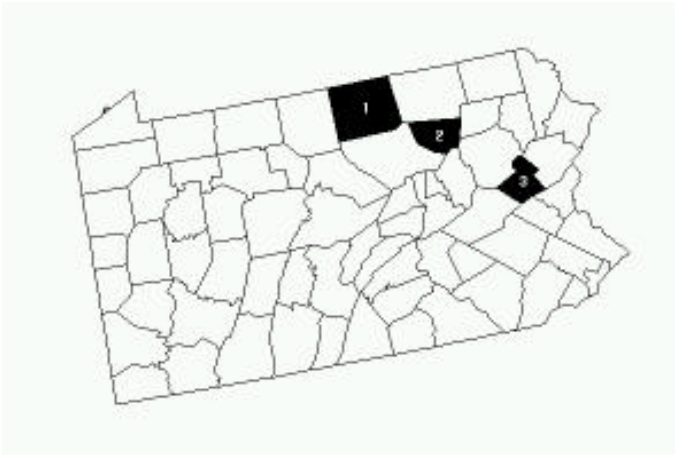

Figure 16. Pennsylvania counties in 1850 analysis by occupation: (1) Tioga, (2) Sullivan, (3) Carbon.

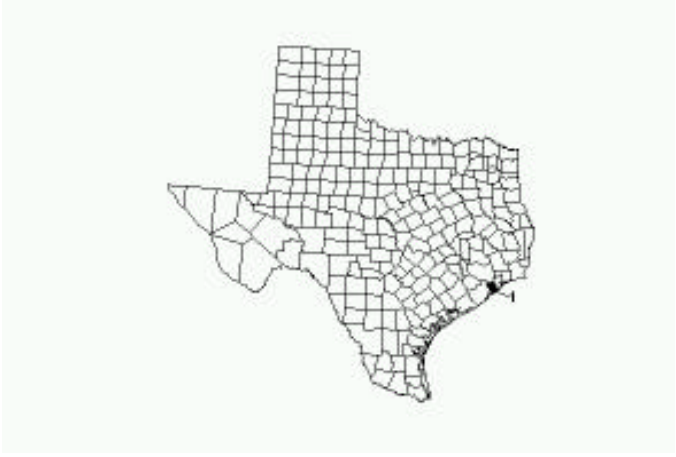

Figure 17. Texas counties in 1850 analysis by occupation: (1) Galveston.

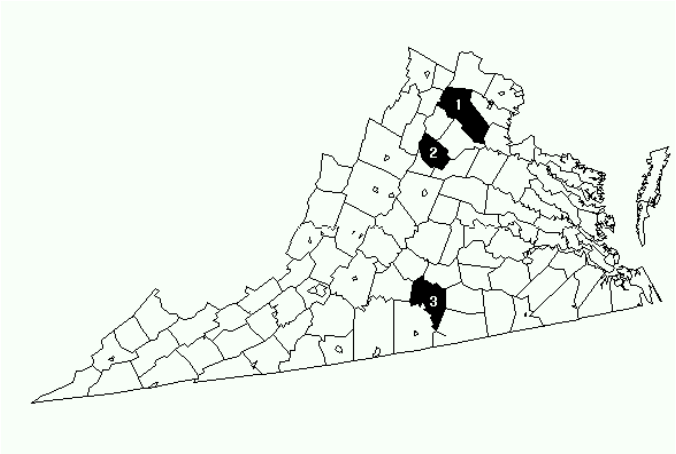

Figure 18. Virginia counties in 1850 analysis by occupation: (1) Fauquier, (2) Madison, (3) Charlotte.