NBER WORKING PAPER SERIES
ON
HISTORICAL FACTORS IN LONG RUN GROWTH

EARLY INDICATORS OF LATER WORK LEVELS, DISEASE, AND DEATH

Robert W. Fogel

Larry T. Wimmer

Historical Paper No. 38

National Bureau of Economic Research
1050 Massachusetts Avenue
Cambridge, MA 02138
June 1992

NBER Historical Paper #38
June 1992

EARLY INDICATORS OF LATER WORK LEVELS, DISEASE, AND DEATH

ABSTRACT

This paper summarizes a collaborative project designed to create a public-use tape suitable for a prospective study of aging among a random sample of 39,616 men mustered into 331 companies of the Union Army. The aim of the project is to measure the effect of socioeconomics and biomedical factors during childhood and early adulthood on the development of specific chronic disease at middle and late ages, on labor force participation at these later ages, and on elapsed time to death. This paper surveys the nature of and quality of the data and data sources to be included in the study, discusses the characteristics of a subsample of recruits from 20 companies recently recruited, looks at questions of representativeness of Union Army recruits to the Northern white male population, and finally examines several issues involving questions of possible selection bias due to linkage failure.

Robert W. Fogel
Center for Population Economics
The University of Chicago
1101 East 58th Street
Chicago, IL 60637
and NBER

Larry T. Wimmer
Department of Economics
149 Faculty Office Building
Brigham Young University
Provo, UT 84602
and NBER

A. Introduction

From its inception in 1920 the National Bureau of Economic Research (NBER) has been a major sponsor of economic research involving questions of long-term growth and policy. Between 1930 and the late 1960s such work was conducted mainly at the macro level and based on the Bureau's pioneering work in the development of national income accounts and related measures of macroeconomic behavior. In 1978, a new NBER program called the Development of the American Economy (DAE) was launched which is concerned mainly with the long-term changes that have occurred at the microeconomic level. To understand the sources of the long-term decline in savings, investment, and male labor-force participation; the factors influencing the rate of technological change; the long-term shifts in the demographic structure of the population; the interaction between socioeconomic behavior and biomedical conditions; and the more recent increase in married, female labor-force participation, we need to know much more about long-term patterns of microeconomic behavior than we have known in the past. Research at the microeconomic level, however, has been inhibited by the absence of suitable data. Attention has turned therefore to the problem of discovering new data sources and determining the feasibility and cost effectiveness of constructing new data sets capable of clarifying the relationships between the current and past behavior of families and firms.

NBER initiated the preparatory work for the project described in this paper in 1978 and since 1981 has worked on it in conjunction with the Center for Population Economics of the University of Chicago. During the subsequent 12 years of research, NBER has brought together a team of 32 investigators (7 of whom are involved in this project) from 22 institutions in the United States and Great Britain. These investigators are drawn from medicine, epidemiology,

biology, demography, statistics, economics, and history. This capability to attract social and natural scientists from various disciplines and universities, and to bring their diverse skills to bear on a given problem is one of the principal resources of NBER. The pilot studies have been aimed at determining whether the creation of the projected data sets is economically feasible and whether it is likely that such data sets will yield the desired information. The results to date have been very encouraging on both counts.

B. The Search for New Data Sources

The traditional source of information regarding the household structure, occupation, residence, and demographic composition of families has been the federal census manuscripts. While we have used these records extensively ourselves, and continue to do so, they possess several shortcomings. Names are subject to frequent spelling errors by the enumerators who wrote down the names, birthplace often contradicts prior information, and age shows heaping around five and ten-year intervals. Finally, regardless of the quality of the record, many questions cannot be answered by examination of a single source at a point in time. Questions involving time, the life cycle, and socioeconomic mobility, require linking over several census years. These linkages are costly and prone to error due to the frequency of changes in residence, the different spelling of names, or the problem of multiple individuals with the same name. The alternative to linking is to use the limiting assumption of cross-sectional studies that in general a random sample of 50 year old heads of households in 1910 is equivalent to a similar sample of 30 year olds in 1890-- that the principle difference is simply the 20 year time lapse. But this assumption disguises many of the most important issues which are precisely the result of decisions within an individual family linked over decades and even generations.

One of the most promising data sources for the study of the family, and particularly of the process of aging, is the military and pension records of the Union Army found in the National Archives of the United States, Washington D. C. The regimental records (RR) filed by regiment, by company, contain information on each recruit's pre-induction residence, age, height, occupation, ethnicity, and other personal identifying characteristics. A history of each recruit's war-time experience is established by combining two sources: the military service records (MSR) and the carded medical records (CMR). These two include a history of his military ranks, companies of service, battles, wounds and injuries; his illnesses and hospitalizations; and his desertions, leaves, capture, and death or discharge with dates and location of each. The most unique of all the military records are the Civil War pension records (PR) which cover almost the entire adult life of most Union Army recruits who survived until 1890 and were not deserters. While these unique records vary in completeness, for the typical pensioner there are over 100 pages of separate documents, testimonials, and affidavits covering their medical, occupational, residential, and demographic history from enlistment until death. The largest pension we have found contains over 800 pages of material, the smallest around 20 pages. Each pension record is designed to verify the reliability of the data whether submitted by the applicant, or by family members, friends, military commanders, family physicians, employers, or attorneys. In addition, each pension contains periodic medical examinations of the claimant conducted by a team of three physicians, each of whom was certified by the Pension Bureau. Unlike the census manuscripts, information such as name, age, residence, height, and occupation is repeated numerous times in many different documents.

These extraordinary records form the basis of a project designed to create a public-use tape suitable for a prospective study of aging among a random sample of 39,616 men mustered into 331 companies of the Union Army. To these military records, economic and demographic information on the households in which these Union Army recruits were raised is being added, including parental ethnicity and wealth from the manuscript schedules of the 1850 and 1860 federal censuses. Indexes of the prevalence of particular diseases in the counties in which recruits were raised during their developmental ages are being constructed from surveys conducted by the AMA and other public health organizations of the time, from disease specific annual morbidity rates at widely scattered military installations, and from disease-specific death rates by counties from the records of the federal and state mortality censuses between 1850 and 1870. Information on the households or institutions in which these veterans lived at late ages will be obtained from the schedules of the 1900 and 1910 federal censuses as well as from the pension records. The cost of linkage over census years will be reduced and the task of verification improved by the availability of multiple addresses and other supporting information from the military service and pension records.

A sample of 5,000 rejectees from the Union Army is being analyzed to measure the biases arising from the army's screening process. Samples from the manuscript schedules of the 1890 and 1910 censuses are also being used to measure biases created by both the recruitment and pension screenings. The tape containing all this information, when completed, will be available for public use and deposited with the ICPSR in Ann Arbor, Michigan.

The information in the tape, covering virtually the entire life cycle of the recruits, will permit the study of the following issues:

1. The effect of nutritional status, socioeconomic factors, and exposure to disease during developmental and middle ages on the morbidity and mortality rates of white males at middle and late ages, using height at age of recruitment as an index of nutritional status during developmental ages, and the weight-for-height of the same individuals later in life as the index for later ages.

2. The effect of exposure to warfare during late adolescence and early adulthood on employment, morbidity, and mortality rates among white males who survived to middle and late ages, with controls for injuries, illnesses, and exposure to stress during service.

3. The effect of host and environmental factors on the probability that recruits would contract specific diseases and on the probability of dying from these diseases before being mustered out.

4. The effect of youthful exposure to virulent environments on the likelihood of developing chronic diseases and on the capacity to work during midadulthood and late ages, with controls for environmental conditions, for age at the time of exposure, and for the duration of the exposure.

5. The effect of different diseases and other disabilities on labor force participation rates and relative earnings at late ages during a period of time when the culture encouraged individuals to be self-supporting to the fullest extent possible.

6. The nature, cost, and effectiveness of arrangements for care of the aged, by the nature of their disabilities and by the income, occupations, family circumstances, and residences of those with whom they were lodged.

7. The effect of pensions on age-specific employment rates under a law that did not penalize regular earnings (i.e. without a substitution effect), and with controls for the status of health, among white males aged 50 to 75.

8. The effect of changes in the views of physicians, public officials, and others (including the self-perceptions of the aged) on the definition of disabilities, on the capacity of individuals to function independently and to work, and on the shift in the care of the disabled from families to institutions.

C. Progress Report

The research effort since early 1987 has focused on providing the evidence needed to demonstrate the usefulness of these data sources at the National Archives and the feasibility of creating a life-cycle sample by linking information from the 10 principal data sets. To resolve those issues, a judgment subsample of 20 companies was drawn from the 331 companies that comprise the total random sample of Union Army recruits. The subsample was chosen in such a way as to guarantee that most of the problems that beset linking would be confronted. Fifteen of these companies came from Ohio and New York, which provided many of the most active units in the war. Within those states, companies were drawn from both rural and urban areas. Two companies were drawn from Michigan, a state that kept poor records on enlistments. One company was drawn from each of three other states: Illinois, Pennsylvania, and New Jersey. Collectively, these states accounted for about 60 percent of Union Army enlistments (Gould 1869, pp. 24-25).

Substantial progress has been made in analyzing the characteristics of this life-cycle subsample. Table 1 indicates the 10 principal data sources from which the linked sample was constructed. It also summarizes the main variables or categories of variables that are included in this life-cycle sample and indicates the main sources for particular categories. The category entitled "conditions diagnosed at each examination" represents numerous diseases which if listed would have taken several pages.

Table 1
The Sources and Principle Variables (or
Categories of Variables) of the Life-Cycle Subsample

Part A
The Principal Data Sources from Which
the Variables are Obtained

Main Sources

C50 = manuscript schedules of U.S. census of 1850
C60 = manuscript schedules of U.S. census of 1860
C00 = manuscript schedules of U.S. census of 1900
C10 = manuscript schedules of U.S. census of 1910
CMR = carded medical service records
MSR = military service records
PR = pension records
PHR = public health records
RH = regimental histories
RR = descriptive books of Union Army regiments

Supplementary Sources

MO = muster-out rolls
PA = pension payout cards
RE = rejection rolls
S55 = state censuses for 1855
S65 = state censuses for 1865
S75 = state censuses for 1875

Part B
Principle Variables (or Categories of Variables)
in the Linked Life-Cycle Data Set

Preservice Data

A) for recruit and his family (C50), C60, RR, PR)
- year of birth
- age at enlistment
- height at enlistment
- geographic origin (urban/rural, county/state/region)
- place of birth & of enlistment
- occupation in 1850, 1860, and at enlistment
- date of enlistment (early/middle/late in war)
- household wealth in 1850, 1860
- ethnicity of parents
- family size in 1850 and 1860
- birth order among surviving siblings
- migration history of parental family
- literacy of individual and family member

B) ecological variables (C50, C60, S55, S65, PHR)
- causes of mortality and morbidity in nation as whole
- causes of mortality and morbidity in locality from early childhood to recruitment
- nature of water supply and sewage system
- soil characteristics (including prevalence of marshes and swamps

C) rejection data (RE)
- causes of rejection
- relation to above

Table 1 (continued)

Service Record

A) morbidity and mortality
(RR, CSR, MSR, PR)
illnesses and hospitalizations
(cause, treatment duration, outcome)
battle injuries
other accidents and trauma

B) potential stress
(RR, MSR, CMR, MO, PR, RH)
rank at enlistment
combat experience
killed
wounded (no. of battles)
fired on (no. of battles)
in zone (no. of battles)
severity of each battle
movement of company between battles
total losses in company & in regiment
cumulative score

POW
when
where
duration
conditions
Tour of Duty
duration
stage of war
transfers
Record
desertion or AWOL
citations for bravery
reprimands or punishments
type discharge
promotions/demotions

C) status (see above)
(RR, MSR, MO, PR)
rank over time
duties over time

Post Service Data

A) Health of veteran (PR)
testimonial history of
health before pension
application
complaints of veteran at
each examination by
Bureau surgeons
height, weight, pulse,
respiration, urinalysis
at each exam by surgeons
conditions diagnosed at
each examination
date of becoming bedridden
dates of admission to
veterans hospitals or homes
& the diagnoses on entry
date & cause of death

B) Occupations and work
(PR, C00, C10)
occupations at each exam
and in 1900 & 1910
surgeons' estimates at
each exam of degree of
impairment for manual
labor (from 1/32 to 32/32)
Pension Bureau's estimate
of overall impairment for
manual labor (after each
application)
months worked in 1900 & 1910

C) Personal & family (PR, PA,
C00, C10)
marriages, divorces, deaths of
spouses
births & deaths of children
residences at examinations &
when receiving payments from
Pension Bureau
household structure in
1900 & 1910

Several random samples have been drawn from the 1910 census rolls to probe the issue of linkage failures and other sources of sample selection bias. For example, persons who identified themselves as Union Army veterans in 1910 have been divided into those who had pensions and those who did not, and differences in the characteristics of these groups are being investigated. Samples have also been drawn of white males of the same birth cohorts who were not Union Army veterans and of blacks who were veterans. The last sample was drawn to investigate the feasibility of a corresponding study of blacks from the pension and Union Army records.

Software has been developed to guide the data collectors and to assure the quality of the data retrieval. There are software packages for each of the principal data sources. The main achievement in this area is the program for the management of the diagnostic information in the surgeons' certificates found in each recruit's pension application. These programs have been designed with the aid of the physicians in the project and are geared to the International Classification of Diseases (9th rev.). The package includes a dictionary of medical terms in use during 1840-1940 and a spelling checker which provides the most likely meanings of words only partially legible and of abbreviations. Similar dictionaries have been developed for place names. Another feature of the software is a program for checking the consistency of information that is retrieved (e.g. persons cannot die before they are born).

D. Evaluation of the Military Records for Socioeconomic Studies

The representativeness of the military records for studies of socio-economic and bio-medical behavior is discussed in four general parts: (1) some characteristics of Union Army recruits as a sample of the northern white male population during 1861-1865; (2) the cluster sample of 331 companies (with 39,616

individuals) and its representativeness with respect to the white population of the Union Army; (3) the quality of the information in the military and pension records; and (4) sample selection biases due to linkage failure and other causes.

1. Some Characteristics of Union Army Recruits as a Sample of the Northern White Male Population

The aggregate figure for recruits into the Union Army is about 2,659,000, of whom 2,480,000 were white. However, these figures are not of individuals but of enlistments. Some individuals re-enlisted one or more times. It is estimated that there were about 370,000 white re-enlistments, so that the total number of white individuals who served in the Union Army is about 2,110,000 (Gould 1869, p. 25). Recruiting was organized at the state level, and the state was broken into recruiting districts that corresponded to Congressional districts. Quotas for recruits were established for each Congressional district based on estimates of available males of military age (18-45). During the first two years of the war, recruiting was conducted on a voluntary basis, with large bounties (often in excess of per capita income) offered to those who signed up. In March, 1863 Congress enacted a law making all men aged 20-45, except in certain exempt categories, subject to a draft. The principal grounds for exemption were previous service, poor health, payment of a commutation fee, and presentation of a substitute (U.S. Provost Marshall General 1866).

Of approximately 5,080,000 white males of military age during 1861-1865 who were at risk to serve in the Union Army, about 1,650,000 were exempt for reasons of health. About 87,000 were exempt by payment of the commutation fee and 144,000 provided substitutes. Additional numbers were exempt because of family dependents, religious scruples, and alien status. Faced with a choice of being drafted without a bounty or volunteering and receiving a bounty, volunteers

remained the principal form of recruiting from March 1863 through the end of the war. All told, less than 65,000 men were drafted into service (Gould 1869, pp. 2-4; Baxter 1875, II, Table 16; U.S. Provost Marshal General 1866, pp. 39-46; Murdock 1971).

Aside from health, the principal basis for the selection of recruits was age. Table 2 shows the 9 cohorts most at risk to be examined by surgeons for service were born between 1837 and 1845. The proportions of these age cohorts that were examined ranged from 61 percent (for those born in 1837) to about 98 percent (for those born in 1843). Among those in the 9 prime birth cohorts, an average of 75 percent was examined and 4 out of 5 of these men were mustered into service. Men in these age cohorts accounted for about 55 percent of the soldiers of the Union Army. Most of the remainder were born between 1822 and 1832. Of these older age cohorts, about 50 percent were examined by surgeons, but more than 1 out of 2 of these older examinees were rejected for reasons of health.

Men who paid commutation fees or provided substitutes generally were not examined, but they presumably were in good health since they could have requested an examination prior to being called for service if they expected to be exempt by reason of health. All men offered as substitutes were identified as such; so the variable "substitute" can enter into the analysis of rejects and of recruits.

The main findings regarding how the recruits compared with the total northern white male population of military are that the foreign-born served in approximately the same proportions are natives; the recruits came from households with the same average wealth as the northern male population as a whole; and they reflected the geographic distribution of the northern population.

Table 2

The Estimated Share of the White Male Population of the North During 1861-1865 that was Examined by Surgeons and Mustered into the Union Army, by Birth Cohort

1 Year of Birth	2 Age in 1861	3 Number at Risk to be Recruited	4 Percentage Examined	5 Percentage Accepted
1845	16	229,700	63.8	55.4
1844	17	227,090	78.0	66.4
1843	18	224,670	97.7	81.4
1842	19	221,250	84.6	68.9
1841	20	217,300	80.2	63.2
1840	21	213,310	75.1	57.8
1839	22	209,290	70.0	52.6
1838	23	205,270	65.2	47.8
1837	24	201,090	60.8	43.5
1832-1836	25-29	941,410	52.3	33.6
1827-1831	30-34	826,630	41.0	23.4
1822-1826	35-39	701,070	41.6	19.6

Sources: Columns 3 and 5 computed from Gould 1869, pp. 80-81; column 4 computed from column 5, using the average rejection rate for disease in Baxter 1875 (II, Table 16) and an age adjustment schedule from a least-square regression. About 10 percent of white northerners who enlisted were born after 1845 or before 1822.

2. The Cluster Sample of 331 Companies and Its Representativeness With Respect to the White Population of the Union Army

The construction of the life-cycle sample began with the drawing of a sample of white recruits who were mustered into the Union Army. During 1981 such a sample was randomly drawn from the surviving regimental records of the Union Army at the National Archives. The technique employed was a one-stage cluster sampling procedure. A cluster sampling procedure does not bias the estimates of the parameters of the population being sampled, but it makes the sample variance larger than it would be in a sample based on the individual recruits (Cochran 1953). However, a sample based on companies has three advantages over one based on individual recruits. First, since the principal objective of this project is not point or interval estimates of means or comparable descriptive statistics, but of multivariate analysis of the relationship between factors inducing early age stress and variables reflecting middle and late age health and behavior, moderately increased variance in the sample is an asset rather than a liability. Second, sampling by companies rather than individuals greatly reduces the cost of linking individuals to other military records and to the pension records. Third, a sample of companies makes it possible to separate company effects of exposure to military stress from individual effects.

The sampling frame was the complete set of companies in the list of all white regiments and other independent organizations presented in Dyer (1908). A number was assigned to each of more than 20,000 companies and these numbers were arranged in the order in which they were drawn from a random number generator. The descriptive books of the regiments containing the designated companies were requested from the National Archives in the order that they were drawn. If a particular book had not survived, the book corresponding to the next random

number was called. Once a book was obtained, all of the information on all the recruits in the designated company was typed into a portable terminal with storage capacity for a day's work (about 400 observations). At the end of the day all of the information in the terminal was transmitted to the computer at Chicago where it was cleaned, coded, and organized into working files. This process was continued until our sample of 39,616 recruits was obtained.

The result of this work yielded 331 companies in 284 regiments, so that about 11 percent of the regiments and other independent organizations, covering all of the states except Rhode Island from which the Union Army recruited white troops, are represented. The 39,616 individuals are a 1.6 percent random sample of all whites mustered into the Union Army (Dyer 1908). This sample of 331 companies with 39,616 recruits will hereafter be referred to as the "recruits" sample.

Table 3 presents a number of statistics that can be used to assess how representative the recruits sample is of the Union Army. Lines 1-6 compare estimates of some key behavioral characteristics. In each of these comparisons the difference between the sample estimate and the figure obtained from the aggregate source is less than one percent (varying between 1 and 9 per thousand). Lines 7-10, which compare the geographic distribution in the recruits sample and in the aggregate source, show that the North Central region is somewhat over represented and New England is somewhat under represented. This was due to the differences in the proportion of the regiments in the two regions whose descriptive books were deposited in the National Archives. The issue could be addressed either by postweighting or by adding additional New England companies (chosen by a random procedure) to the recruit sample. However, for the multivariate procedures currently contemplated, the size of the New England subsample is adequate. Various experiments with postweighting produced results

that were virtually the same as the internal weights, a finding anticipated by the closeness of the statistics computed from the recruit sample to those in the aggregate sources reported in lines 1-6 of Table 3.

3. The Quality of the Information in the Military and Pension Records

Several issues bearing on the quality of the information in military and pension records have been raised. The first relates to the range and relative reliability (compared with the federal censuses) of the information in the military and pension records. The results of our investigation of this question follows.

On a number of variables (such as wealth of parental households in 1860, order of birth among surviving siblings in 1860, number of months worked in 1899-1900 and 1909-1910, and persons living in veteran's household in 1900 and 1910) there is no counterpart in the military and pension records. It is to get this information that we are linking the military and pension records to the census schedules.

On the other hand, for most variables in the military and pension records, there are no counterparts in the federal censuses. As mentioned in the introduction, there is, for example, no information in the federal censuses on military service except for the question of whether individuals had served in the Union or Confederate armies, which appears in the censuses of 1890, 1900, and 1910. By contrast the military and pension records provide detailed information on military service history, on medical history during and after service, on the date and cause of death of veterans, and on marriages and divorces (or spouses' death), and on birth and deaths of the veterans' children.

Table 3

The Representativeness of the Recruits Sample With Respect to White Population of the Union Army

	1 Estimated from recruits sample	2 Estimated from aggregate source
Test statistic sources		
1. Proportion of recruits who are native-born (percent)	74.9	75.5
2. Desertion rate (percent)	9.08	9.16
3. Mean height of recruits aged 25-29 (inches)	68.04	68.01
4. Mean height of recruits aged 30-34 (inches)	68.08	68.05
5. Mean height of recruits aged 35 or over (inches)	67.89	67.96
6. Proportion of recruits who died during war (percent)	11.96	12.05
7. Proportion of recruits from New England (percent)	6.3	12.8
8. Proportion of recruits from Middle Atlantic (percent)	29.2	31.7
9. Proportion of recruits from North Central (percent)	54.2	45.7
10. Proportion of recruits from all other states and territories (percent)	10.1	9.9

Note: The first 3 regions are defined as follows: New England, ME, NH, VT, MA, RI, CT; Middle Atlantic, NY, NJ, PA; North Central, OH, IN, IL, MI, WI, MN, IA, MO, KA.

Sources: All of the entries in column 2, except for lines 2 and 6, were computed from Gould 1869, p. 27. The numerator of entry in column 2, line 2 is from U.S. Provost Marshal General 1866, p. 89, the denominator is from Gould 1869, pp. 25-26 and consists of the sum of the different white soldiers plus black soldiers from states subject to the draft since the aggregate source did not report desertions by race. Line 6 of column 2 is computed from Dyer 1908, p. 11.

The best variables for the comparisons of relative quality are name, place of birth, and age. The findings with respect to the reliability of the evidence in the census and in pension and military records are as follows:

	Census records	Pension records
Name	Subject to frequent spelling errors by enumerator who wrote the name down or to variants of the spelling of the name.	Numerous documents are included giving the name, including many written in the hand of the individual. Discrepancies in spelling investigated by Pension Bureau in the course of determining eligibility of the individual. Documentary support includes marriage certificates, enlistment and discharge papers, and affidavits by neighbors, officers, and others who had personal knowledge of the individual's identity.
Place of birth	State or country of birth only; provided by person answering questions for all residents of household. Errors because respondent did not know place of birth of all residents in households, or because town was sometimes substituted for	Place of birth given repeatedly in numerous documents for country, state, county, and town and frequently for such smaller geographic units as townships. Support by affidavits by persons who knew the individuals and other documents such as marriage certificates, death certificates,

	country. Recruits who immigrated after the 1860 census will not be covered.	because of spelling errors for small European principalities, family Bibles, etc. Discrepancies investigated by Pension Bureau. All immigrants covered regardless of date of arrival.
Age	Errors due to lack of knowledge of respondent of all residents in household as well as to recording errors, age heaping and poor memory.	Reported in numerous documents such as surgeons' certificates, death certificates, and family Bibles and substantiated by various affidavits. Discrepancies investigated by Pension Bureau.

About 80 percent of all of the information in the pension records pertain to veterans between the ages of 25 and 65, so there is no void of information between these ages in the pension file. For veterans who lived to 1910, there is an average of 6.2 residences reported in the pensions. The average interval between these reports is 4.0 years. For medical examinations by surgeons there is an average of 4.5 per survivor to 1910 and the average interval between examinations is 5.5 years. Reporting of occupations is somewhat more sporadic, but one-third did not change occupations between enlistment and 1910. Of the two-thirds who did, 50 percent retired before June 1, 1910. Among veterans surviving to 1910, there is an average of 2.6 reports on occupation and the average interval between reports is 9.6 years.

Another question concerns the completeness of the death dates in the pension records. In the 20-company subsample described in Table 1, 95 percent of the

pensioners have death dates. With the aid of the information in the payout cards we believe the figure can be made even higher.

A second question relates to the quality of the cause-of-death information in the civil death certificates. This issue is of less importance in this study than in epidemiological studies that depend primarily on death certificate information. The key variables in our analysis of waiting time to death will not be the listed causes of death, but the chronic diseases and other stresses that the individual experienced prior to death, e.g. the extent to which childhood malnutrition and extreme military stress at early adulthood combined with given chronic conditions before given ages (50, 60) affect the probability of dying during the next 10 years--or the waiting time to death. In some of these analyses information on cause of death may be used as an additional variable, with the aim of assessing its usefulness. But it will not play a primary role in the analyses.

Finally, there are issues involving the quality of the medical histories in the pension records. The medical histories fall into two categories: (1) testimonial evidence by the veteran, supported by affidavits from private physicians, on his health prior to examination by a surgeon of the Pension Bureau; and (2) diagnostic information in successive examinations reported in the certificates of Bureau surgeons.

Testimonial evidence has been examined in two ways: its consistency with the diagnoses of the Bureau surgeons; and the capacity of early-age variables to predict the chronic diseases reported in testimonials. These preliminary analyses suggest that the testimonial evidence is useful. We have not yet been able to determine whether variables obtained from the testimonies, when added to

waiting-time regressions based on the diagnoses of Bureau surgeons, reduces the unexplained variance.

4. Sample Selection Biases Due to Linkage Failure

Not all of the individuals in the recruits sample can be linked across the 10 data sets which together comprise the life-cycle sample. The investigators in this project are concerned with possible biases that might arise from linkage failure. To investigate these issues and to determine linkage rates, the judgement subsample of twenty companies was chosen from the 331 companies in the recruits sample in a way that would reflect the full range of linkage problems. Table 4 indicates the percentage of the individuals at risk to be linked to each of the other nine data sets that comprise the life-cycle sample.

The first three lines indicate that linking rates from the recruits to the other three military data sets are very high, varying between 86 and 100 percent. Line 4 shows that the linkage rate to the pension sample is also high, with 85 percent of the eligible recruits having been found (the ineligible groups were mainly deserters and persons who died during service without eligible dependents). Undoubtedly a portion, of the 15 percent deemed eligible-to-be-linked but not found were in fact ineligible, falling into such categories as "died without a war-related disability eligible for a pension." Part of the linkage failure is due to the disappearance of some pension records from the Archives and other records stored at the Veterans Administration have not been made available to date.

Lines 5 to 8 give the linkage rates to the manuscript schedules of the four federal censuses. Contrary to some expectation the linkage rates are higher in the 1900 and 1910 censuses (73 and 65 percent respectively) than in the 1860 and

Table 4

Linkage Rates for the Nine Data Sets Which Together With the Recruits Sample Form the Life-Cycle Sample

	Data Set	1 Number at risk to be linked	2 Number linked so far	3 Percentage linked (Col. 3 ÷ Col.2) x 100
1.	Military service record (MSR)	2,357	2,311	98
2.	Carded medical record (CMR)	2,240	1,933	86
3.	Regimental histories (RH)	2,357	2,357	100
4.	Pension record (PR)	1,957	1,672	85
5.	U.S. Census of 1910 (C10)	593	385	65
6.	U.S. Census of 1900 (C00)	1,036	753	73
7.	U.S. Census of 1860 (C60)	2,217	914	47
8.	U.S. Census of 1850 (C50)	1,855	823	42

1850 censuses (47 and 42 percent respectively). We are currently making use of the pension pay-out records, which were only recently located. We expect that with these records, which indicate the addresses to which the monthly pension checks were sent, we will be able to raise the linkage rates for the 1900 and 1910 censuses to the 75 to 85 percent range.

It should be noted that some of the eligible individuals linked to the 1900 census were not linked to the 1910 census and vice versa. The proportion of eligible individuals linked to at least one of these two censuses is 82 percent. Similarly, the proportion of individuals who were linked to at least one of the two pre-war censuses is 60 percent.

The relatively low linkage rates for the 1860 and 1850 censuses are due to the absence of adequate soundexes for these censuses, compounded by the fact that indexes for this period cover only head of households at a time when our recruits were young, most often living within their parents households. A commercial genealogist has recently constructed soundexes for these years which we hope to be available soon. With these soundexes the linkage rates should rise to the 55 to 65 percent level for each census, and the proportion linked to the wealth and family information in at least one of the prewar censuses should exceed the 70 percent level. That level of linkage should be adequate for the analytical purposes of this project. If for some reason it is not, the level of linkage can be increased by making use of the Civil War enrollment lists at the National Archives, a special census of men between ages 20 and 46 conducted in 1863 as a basis for the draft. These lists give addresses that could aid in the location of individuals not found through the soundex.

Table 5 investigates both OLS and logit regressions aimed at identifying the factors that affect the odds of linking the individuals in the recruits sample

to the other data sets in the life-cycle sample. The eleven behavioral variables used as predictors are attributes obtained from the recruits sample. The main finding of these regressions is that being foreign born was the principal nonrandom factor accounting for the failure of linkage to the 1850 and 1860 censuses. In linking to the 1900 and 1910 censuses, having been foreign born is much less important in explaining linkage failure than in the prewar case. The discrepancy is due primarily to the fact that about two-thirds of the foreign-born recruits arrived in the U.S. after June 1, 1850 and about 7 percent arrived after June 1, 1860 and hence were not covered by the census. The behavioral factors do not explain much of the variation in the odds of linking in either the prewar or the postwar censuses. The chi-square and R-square values are especially low in the postwar census, with the behavioral factors accounting for less than 3 percent of the variation in the probability of making a link.

In the case of the pension records, "died during the war" and being a deserter are the principal reasons for the nonrandom linkage failure. Deserters were in most cases ineligible for a pension and many of those who died during the war had no dependent who was eligible. The foreign dummy is also significant in part because many of the foreigners who died during the early postwar years had no eligible dependents or were used behind the front and so were less likely to incur war-related disabilities. However, those who survived to be eligible under the pension law of 1890 were as likely to be linked to pension records as natives.

In the case of the military service records, two variables are statistically significant at the 5 percent level, but the magnitude of the coefficients is small and the chi-square and R-square values are very low, indicating the linkage failure was due almost exclusively to random factors, such as loss of records.

Table 5

LOGIT and OLS Regressions on the Probability of Linking Individuals in the Recruits Sample to the Other Data Sets in the Life-Cycle Sample

(for 20 companies)

Variable	CEN50 [M=57.0; N=1669]		CEN60, CEN00, CEN10 [M=82.3; N=859]		MSR [M=96.8; N=2317]		PEN [M=71.0; N=2317]		CMR [M=63.3; N=2317]	
	LOGIT	OLS	LOGIT	OLS	LOGIT	OLS	LOGIT	OLS	LOGIT	OLS
INTERCEPT	-4.502*	-0.271	-1.684	0.676~	2.386	0.934*	-5.153*	0.186	0.610	-0.071
	(1.602)	(0.315)	(2.756)	(0.367)	(3.228)	(0.097)	(1.469)	(0.235)	(1.424)	(0.253)
AGE1861	0.013~	0.003~	-0.023	-0.004	0.003	0.000	0.035*	0.006*	0.006	0.001
	(0.008)	(0.002)	(0.014)	(0.002)	(0.016)	(0.000)	(0.007)	(0.001)	(0.006)	(0.001)
HEIGHT	0.036~	0.008~	0.020	0.003	0.014	0.000	0.035~	0.006~	0.026	0.005
	(0.021)	(0.005)	(0.039)	(0.006)	(0.048)	(0.001)	(0.020)	(0.004)	(0.018)	(0.004)
FREQFNAM	0.020	0.004	0.002	0.000	-0.059^	-0.002^	-0.007	-0.001	-0.004	-0.001
	(0.013)	(0.003)	(0.022)	(0.003)	(0.026)	(0.001)	(0.012)	(0.002)	(0.011)	(0.002)
FREQLNAM	0.273^	0.059	0.174	0.025	2.302~	0.031	-0.213	-0.042	0.069	0.013
	(0.266)	(0.058)	(0.556)	(0.078)	(1.438)	(0.020)	(0.255)	(0.047)	(0.245)	(0.051)
DUMFOR	-1.431*	-0.325*	-0.244	-0.038	-0.595^	-0.022^	-0.795*	-0.154	-0.105	-0.023
	(0.157)	(0.033)	(0.265)	(0.041)	(0.282)	(0.010)	(0.127)	(0.024)	(0.124)	(0.026)
DUMFARM	0.503*	0.114*	0.317~	0.046*	0.244	0.009	0.481*	0.085*	0.478*	0.101*
	(0.110)	(0.025)	(0.192)	(0.028)	(0.259)	(0.008)	(0.106)	(0.019)	(0.097)	(0.020)
DUMT3YRS	0.955*	0.216*	-0.044	-0.006	0.124	0.007	0.079	0.015	1.236*	0.288*
	(9.218)	(0.048)	(0.258)	(0.038)	(0.327)	(0.011)	(0.146)	(0.026)	(0.134)	(0.028)
DUMSUB	-0.974	-0.182	-1.430	-0.304		0.048	0.047	0.014	1.384	0.267^
	(0.679)	(0.124)	(1.017)	(0.192)		(0.043)	(0.551)	(0.103)	(0.651)	(0.111)
DUMDIED	-0.026	-0.004			0.414	0.009	-1.268*	-0.253*	0.172	0.029
	(0.156)	(0.034)			(0.452)	(0.011)	(0.140)	(0.027)	(0.152)	(0.029)
DUMWOUND	0.139	0.030	0.274	0.035	0.172	0.005	0.380^	0.072^	1.365*	0.226*
	(0.165)	(0.036)	(0.332)	(0.044)	(0.421)	(0.012)	(0.163)	(0.029)	(0.196)	(0.031)
DUMDES	-0.135	-0.031	-1.095*	-0.207*	0.203	0.007	-2.030*	-0.440*	-1.217*	-0.285*
	(0.203)	(0.045)	(0.400)	(0.072)	(0.447)	(0.014)	(0.177)	(0.033)	(0.167)	(0.035)
CHISQ/RSQ	143.39	10.57	17.12	2.29	19.36	0.84	318.09	14.23	221.39	12.75

Note: 17 observations with DUMSUB=1 were dropped to obtain MSR LOGIT results. For MSR, DUMSUB=1 perfectly predicts a successful link.

Table 5 (concluded)

Notes: The variables are defined as follows:

AGE1861 = Age on January 1, 1861 (in years)

HEIGHT = Height at enlistment (in inches)

FREQFNAM = Commonness of first name, measured by the relative frequencies in RR (in percent)

FREQLNAM = Same as above for last names (in percent)

DUMBFOR = Dummy: 1 if foreign-born; 0 otherwise

DUMFARM = Dummy: 1 if a recruit's occupation prior to enlistment was reported to be a farmer; 0 otherwise

DUMT3YRS = Dummy: 1 if the term of enlistment was for 3 years or more; 0 otherwise

DUMSUM = Dummy: 1 if the recruit was a substitute; 0 otherwise

DUMDIED = Dummy: 1 if died during the military service; 0 otherwise

DUMWOUND = Dummy: 1 if wounded during the military service; 0 otherwise

DUMDES = Dummy: 1 if deserted during military service; 0 otherwise

Chi-square has 11 degrees of freedom in columns 1, 5, 7, and 9, and 10 degrees in column 3.

The RSQ is given in percent.

The symbols in column headings are:

M = mean of dependent variable (in percent)
N = number of observations
CEN50, CEN60 = the 1850 and 1860 censuses
CEN00, CEN10 = the 1900 and 1910 censuses
MSR = military service records
PEN = pension records
CMR = carded medical service record

Significance levels are coded as * = 1%; ^ = 5%; ~ = 10%

In the case of the medical service records, desertion is the principal, nonrandom factor accounting for linkage failure, due largely to the short period that deserters were at risk to require medical care in service.

The predictability of the factors which explain linkage failure suggests that any biases introduced by censoring can be corrected by reweighting subsamples having the relevant characteristics (this applies to subgroups overrepresented as well as those underrepresented). However, tests reveal that height has very little effect on estimates of key parameters. For example, the height at enlistment of those who were linked to the 1900 and 1910 samples differed from those who were not by just one-tenth of an inch and the standard deviations were quite similar. Since virtually the full range of variation remains, and censoring hadvery little effect on means (which reenforces the conclusion that linkage failures were due overwhelmingly to random factors), it is unlikely that the multivariate procedures that will be employed in the analyses of the data will be significantly affected by the linkage failures.

Several other tests of the representativeness of the linked sample were undertaken. One of these concerned the wealth distribution of all adult males (age 20 and over) in the households to which the recruits were linked in the 1860 census. The mean wealth was $2,187 (ln = 7.69) and the SD of the logs is 1.45. These figures are quite close to Soltow's (1975) random sample of northern males (age 20 and over) in 1860: mean = $2,040 (ln = 7.62) and SD of logs = 1.87. Although the variation is somewhat reduced in the recruit subsample, the mean is not significantly changed. Moreover, the range of wealth covered by the linked sample extends from zero to the top 4 hundredths of one percent of Soltow's distribution (i.e. less than the top 0.04 percent of the wealth distribution is missing).

Still another test of the linked subsample can be performed by comparing the causes of death among veterans linked to the 1910 census with the causes of death reported by the death registration states. To standardize for age, the comparison was limited to veterans aged 70-74. The result is shown in Table 6. The chi-square value for the differences between the two distributions (7.79 with 5 degrees of freedom) is not significant at the 0.10 level or higher.

Although linkage failure to the censuses appears to be largely random with respect to the variables considered here, we do not assume that this situation will pertain to other key variables, and will reweight should the evidence indicate that such a procedure is called for.

The most difficult problems of inference related to screening problems stem from the varying date of entry into the pension records. The governing principle in dealing with such data is that individuals are not at risk for most purposes until they applied for the pension. Life tables constructed on this principle for the period circa 1900 are similar to the mortality schedules constructed from the death registration data but somewhat lower, as is to be expected, since the areas covered by death registration in 1900 were still concentrated in the high-mortality states and cities (Preston, Keyfitz, and Schoen 1972; Preston and Haines 1991).

E. Conclusions

The most important aspect of this project is that it will permit, at a relatively low cost, the prospective study of a larger and more representative sample of white males, over virtually their entire life cycle, than it has hitherto been possible to study (Migdal, Ables, and Sherrod 1981; Dawber 1980; Shock 1984). Unlike many other prospective studies, the individuals in this sample are not limited to a single, relatively homogeneous community. The sample

Table 6

Distribution of Causes of Death of
White Males Ages 70-74

Cause of Death	1 Death registration area		2 Veteran's sample 1910	
	number	%	number	%
Infectious diseases	1,194	7.1	12	8.0
Neoplasms	1,836	10.9	10	6.7
Cardiovascular	10,044	59.4	85	56.7
Influenza, pneumonia, bronchitis	2,536	15.0	32	21.3
Diarrheal	278	2.6	2	1.3
Violence, including motor vehicles	1,008	6.0	9	6.0
Total classified (does not include other and unknown)	16,896	100.0	150	100.0

Note: Column 1. The source is Preston, Keyfitz, and Schoen 1972.

covers individuals who were born in 22 states and 30 foreign countries, and who were reared in households whose socioeconomic circumstances ranged from the deepest poverty to substantial wealth. Unlike the subjects of other prospective studies, many of those covered by the public-use tape were highly mobile geographically. They were also exposed in childhood and at young adult ages to a far wider range of disease and sources of stress than is usual for the individuals covered by other prospective studies.

A special feature of this sample, stemming primarily from the pension records, is the opportunity it provides for studying the synergism between socioeconomic and biomedical conditions. The information on the wealth, occupation, family size, birth order, ethnicity and other socioeconomic characteristics of the households in which the recruits were raised, make it possible to measure the impact of these factors (both separately and jointly with pre-enlistment exposure to disease), on the probability that they would contract various diseases while in service, the probability that they would die from these diseases, and for those who survived military service, the probability that they would develop particular chronic diseases by given ages.

Where individuals "prematurely" developed chronic conditions we can estimate the decrease in economic productivity at middle and late ages attributable to early-life conditions. By relating such late-age disabilities to early-life socioeconomic and biomedical conditions it is possible to measure the full life-cycle economic costs of these early-life events, including the human capital cost of premature mortality.

It will also be possible to investigate the effect of nutritional status during the developmental ages on morbidity and mortality at later ages. The public-use tape will contain two measures of nutritional status. One is height-

by-age at recruitment, for ages ranging from the mid-teens to the mid-forties. Stature at given ages during growing years and at maturity has been shown to be the single best measure of nutritional status during developmental ages, and final heights are an index of the cumulative nutritional experience during the developmental ages (Habicht, Yarborough, and Martorell 1979; Fogel 1991). The other measure is weight-for-height which will be available for veterans at numerous ages between the time they left the services and their deaths.

The particular period spanned by the individuals in this study adds still another dimension absent from most other prospective studies of aging. The recruits were born mainly during the decades 1820-1850 in which public health had declined to its lowest level since the beginning of the eighteenth century (Duffy 1953; Pope 1991; Smillie 1955; Fogel 1986, 1989 and 1991b). Since reforms were quite uneven, it is possible to study the effects of the reforms (or the lack thereof) on the health and productivity of the veterans. For example, the tape provides an opportunity to measure the effect of exposure to lead poisoning in New York and other cities which employed lead pipes to distribute fresh water to some of its districts (N.Y. Dept. Health 1872).

A detailed study of the recorded outbreaks of disease during the Civil War will provide a unique opportunity for testing various public health theories (Woodward 1863). In the case of measles, for example, one would expect a large percentage of the urban recruits to have had measles as children and hence to have been relatively immune to the disease. On the other hand, many of the men from sparsely settled rural areas may have escaped previous exposure to measles. Moreover, since native-born rural men were about two inches taller at maturity than their urban counterparts, they enjoyed a superior nutritional status during their developmental ages. These points plus the proposition that the course of

measles is influenced by nutrition suggests two hypotheses that can be tested against the Union Army medical records. Urban recruits should have had a lower incidence of measles than rural recruits, but a higher case-fatality rate.

The military data also provide an opportunity for studying the effects of wartime trauma, including stress, on both physiological conditions at later ages and, for a given physiological status, on the deviations from predicted capacity to work. Such information may shed light on both the health and economic consequences of stress that may be experienced by veterans of the Vietnam War at later ages. In this connection, it should be noted that the array of experiences within the Civil War sample is very wide. Many men in the sample were exposed to numerous battles, wounded in various degrees, often more than once, and some suffered as prisoners of war. On the other hand, about 15 percent of the recruits served for 100 days or less. Many in this latter category belonged to militia units formed exclusively from men of their locality and never saw battle. Many were recruited (and released) before the major battles or after Lee's surrender and thus escaped not only the stress of combat but even a substantial threat of exposure to combat (Hamersly 1888; Dyer 1908).

The public-use tape will also provide hitherto unavailable information needed to evaluate the potential for increasing labor force participation rates at ages 65 and over from their currently low levels of 18 percent to levels of 60 to 75 percent which prevailed before 1930 (Moen 1987a and 1987b; U.S. Bur. Lab. Stat. 1984; cf. Random and Sutch 1986 and 1989). The numerous proposals for moving in this direction as a solution to the long-term crisis in social security cannot be assessed adequately merely from cross-sectional analysis of recent labor force data. Analyses based on recent data cannot explain the factors that have made the income and price elasticities of the demand for leisure since 1960 so much

higher than they were during the preceding century (Moen 1987a). To evaluate the possibility of increasing LFPR by the magnitudes now being proposed, it is necessary to treat as variables characteristics often taken as given in recent econometric studies, such as changes in the organization of production and both societal and individual attitudes toward work at older ages.

An important byproduct of this study is the usefulness of the information it will contain for epidemiologists specializing in the diseases of the third world countries today. Both during their developmental ages and in the Union Army the principal diseases afflicting the recruits were those that currently predominate in the less developed nations. The careful clinical descriptions of the progress of these diseases in the medical histories of the recruits, the repeated follow-up examinations by surgeons of the Pension Bureau, and the large amount of information on the socioeconomic and biomedical conditions of the men both before service and afterwards, may provide useful hypotheses (and opportunities to test them) that are of relevance to the formulation of public health policies for less developed nations today.

Creation of this data set is made possible by the availability of the military and pension records at the National Archives, the generous and capable assistance of Archival personnel, and more recently the technological advances in mainframe and lap-top computers which make the retrieval, storage, and analysis of these data feasible and cost effective.

ACKNOWLEDGEMENTS

An earlier version of this paper was presented before the section on Government Statistics of the American Statistical Association, August 20, 1991. This paper draws heavily from work done on a collaborative project jointly sponsored by the National Bureau of Economic Research (NBER) and the Center for Population Economics of the University of Chicago (CPE). The preliminary research has been supported by grants from the Walgreen Foundation, the Exxon Foundation, the University of Chicago, and Brigham Young University. Recently funding for the project has been received through the National Institute on Aging. Major contributors to the project in addition to the above authors include Michael Haines (Colgate University, NBER), Clayne L. Pope (Brigham Young University, NBER), Irwin Rosenberg (Tufts Medical School), Nevin S. Scrimshaw (Harvard University), and James Trussell (Princeton University, NBER). The findings presented in this paper are tentative and subject to change. Not to be cited, quoted or reproduced without permission. Computer tapes of the new data compiled under this project will be made available for public use when they are completed by depositing them with the ICPSR in Ann Arbor, Michigan.

Literature Cited

Achenbaum, W.A., 1978. Old Age in the New Land: The American Experience Since 1790, Baltimore: Johns Hopkins University Press.

Baxter, J.H., 1875, Statistics, medical and anthropological, of the Provost-Marshal-General's bureau, derived from records of the examination for military service in the armies of the United States during the late war of the rebellion of over a million recruits of drafted men, substitutes, and enrolled men. Washington, D.C.

Cochran, W.G., 1953, Sampling Technique, Wiley, New York (2nd ed. 1959).

Dawber, T.R., 1980, The Framingham Study: The Epidemiology of Arteriosclerotic Disease, Harvard Univ. Press, Cambridge, MA.

Dyer, F.H., 1908, A Compendium of the War of the Rebellion, Dyer Pub. Co., Des Moines, 3 vols.

Fogel, R.W., and Engerman, S.L., 1974, Time on the Cross: The Economics of Slavery, 2 vols., Little, Brown, Boston.

Fogel, R.W., 1986a, Nutrition and the Decline in Mortality Since 1700: Some Additional Preliminary Findings, mimeo, NBER Working Paper No. 1802.

Fogel, R.W., 1989, Without Consent or Contract: The Rise and Fall of American Slavery, W.W. Norton, New York.

Fogel, Robert W., 1991, "Biomedical Approaches to the Estimation and Interpretation of Secular Trends in Labor Productivity, Equity, Morbidity, and Mortality in Western Europe and America, 1780-1980," typescript, Univ. of Chicago.

Gould, B.A., 1869, Investigations in the military and anthropological statistics of American soldiers, Cambridge, MA: Riverside Press.

Graebner, W., 1980, A History of Retirement: The Meaning and Function of an American Institution, 1885-1978, New Haven: Yale University Press.

Haber, C., 1983, Beyond Sixty-Five: The Dilemma of Old Age in America's Past, New York: Cambridge University Press.

Hamersly, T.H.S., 1888, Complete Army and Navy Register of the United States of America from 1776 to 1887, T.H.S. Hamersly, pub., New York.

Migdal, S., Ables, R.P., and Sherrod, L.R., 1981, An Inventory of Longitudinal Studies of Middle and Old Age, Soc. Sci. Res. Council, New York.

Moen, J.R., 1987, "The Labor of Older Men: A Comment," Journal of Economic History, Vol. 47, no. 3(Sept.), pp. 761-767.

Moen, J.R., 1987a, "Essays on the Labor Force and Labor Participation Rates in the United States From 1860 Through 1950, University of Chicago, Unpublished Ph.D. thesis.

Moen, J.R., 1988, "From Gainful Employment to Labor Force: Definitions and a New Estimate of Work Rates of American Males, 1860 to 1980.

Murdock, E.C., 1971, One Million Men: The Civil War Draft in the North, Heffernan Press, Worcester, MA.

New York Dept. of Health, 1872, Annual Report.

Pope, C.L., 1986, Native adult mortality in the U.S., 1770-1870, mimeo, Brigham Young University.

Preston, S.H., Keyfitz, N., and Schoen, 1972, Causes of Death: Life Tables for National Populations, New York.

Preston, S.H., and Haines, M.R., 1991, Fatal Years: American Child Mortality in the Late Nineteenth Century, Princeton Univ. Press, Princeton, NJ.

Ransom, R.L. and Sutch, R., 1986, "The Labor of Older Americans: Retirement of Men On and Off the Job, 1870-1937," Journal of Economic History, Vol. 46, no. 1(March), pp. 1-30.

Ransom, R.L. and Sutch, R., 1989, "The Trend in the Rate of Labor Force Participation of Older Men, 1870-1930: A Reply to Moen," Journal of Economic History, Vol. 49, no. 1(March), pp. 170-183.

Shock, N.W., et al., 1984, Normal Aging: The Baltimore Longitudinal Study of Aging, NIH Pub. No. 84-2450, G.P.O., Washington, D.C.

Smillie, W.G., 1955, Public Health: Its Promise for the Future, Macmillan, New York.

Soltow, Lee, 1975, Men and Wealth in the United States 1850-1870, Yale University Press, New Haven, CN.

Stone, D., 1984, The Disabled State, Philadelphia: Temple University Press.

U.S. Pension Bureau, 1873-1910, Annual Report of the Commissioner of Pensions, Washington, DC: G.P.O.

U.S. Provost-Marshal-General, 1866, Final Report, U.S. House of Representatives, Exec. Doc. No. 1, 39th Cong. 1st Sess., Ser. Nos. 1251, 1252.

Yasuba, Y., 1961, Birth Rates of the White Population in the United States, 1800-1860, Johns Hopkins Univ. Press, Baltimore.