1-1997

Childhood Conditions that Predict Survival to Advanced Ages Among African Americans

Samuel H. Preston
University of Pennsylvania, spreston@sas.upenn.edu

Mark E. Hill
University of Pennsylvania

Greg Lee Drevenstedt
University of Pennsylvania

Recommended citation:
Preston, Samuel H., Mark E. Hill, and Greg L. Drevenstedt. 1997. "Childhood Conditions that Predict Survival to Advanced Ages Among African Americans." PARC Working Paper Series, WPS 97-03.
This working paper was published in a journal:
Preston, Samuel H., Mark E. Hill, and Greg L. Drevenstedt. 1998. "Childhood Conditions That Predict Survival to Advanced Ages Among African-Americans." *Social Science & Medicine* 47(9):1231-1246. https://doi.org/10.1016/S0277-9536(98)00180-4.

Childhood Conditions that Predict Survival to Advanced Ages Among African Americans

Abstract

This paper investigates the social and economic circumstances of childhood that predict the probability of survival to age 85. It uses a unique study design in which survivors are linked to their records in U.S. Censuses of 1900 and 1910. A control group of age and race-matched children is drawn from Public Use Samples for these censuses. It concludes that the factors most predictive of survival are farm background, having literate parents, and living in a two-parent household. Results support the interpretation that death risks are positively correlated over the life cycle.

Keywords

cohort mortality, longevity, African Americans, socioeconomic factors, geographic factors, oldest old

Disciplines

Demography, Population, and Ecology | Family, Life Course, and Society | Social and Behavioral Sciences | Sociology

Comments

Recommended citation:

Preston, Samuel H., Mark E. Hill, and Greg L. Drevenstedt. 1997. "Childhood Conditions that Predict Survival to Advanced Ages Among African Americans." PARC Working Paper Series, WPS 97-03.

This working paper was published in a journal:

Preston, Samuel H., Mark E. Hill, and Greg L. Drevenstedt. 1998. "Childhood Conditions That Predict Survival to Advanced Ages Among African-Americans." *Social Science & Medicine* 47(9):1231-1246. https://doi.org/10.1016/S0277-9536(98)00180-4.

Population Aging Research Center, University of Pennsylvania
Working Paper Series No. 97-03

Childhood Conditions that Predict Survival to Advanced Ages among African Americans*

Samuel H. Preston
Mark E. Hill
Greg L. Drevenstedt

Population Studies Center
University of Pennsylvania
3718 Locust Walk
Philadelphia, PA 19104-6298

**This research was supported by a grant from the National Institute of Aging, AG 10168. We are grateful to Irma Elo, Ira Rosenwaike, Andrew Foster, and Paul Rosenbaum for comments and suggestions. We also thank Averil Clarke and Benjamin Preston for research assistance.*
Direct correspondence to Samuel Preston, Population Studies Center, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA 19104-6298. E-mail: spreston@pop.upenn.edu.

ABSTRACT

This paper investigates the social and economic circumstances of childhood that predict the probability of survival to age 85. It uses a unique study design in which survivors are linked to their records in U.S. Censuses of 1900 and 1910. A control group of age and race-matched children is drawn from Public Use Samples for these censuses. It concludes that the factors most predictive of survival are farm background, having literate parents, and living in a two-parent household. Results support the interpretation that death risks are positively correlated over the life cycle.

KEYWORDS:

COHORT MORTALITY; LONGEVITY; AFRICAN AMERICANS; SOCIOECONOMIC FACTORS; GEOGRAPHIC FACTORS; OLDEST OLD

Studies of social and economic differentials in mortality typically relate circumstances at one moment in time to contemporary mortality risks. Literally hundreds of studies that date back more than a century show that, with rare exception, socially and economically disadvantaged groups suffer elevated risks of death (Williams, 1990; Feinstein, 1993). Such results are hardly surprising. Healthiness and longevity are nearly universal goals, and groups with more economic and social resources are better equipped to achieve these goals.

Recently, studies have begun to investigate the relationship between social and economic features of childhood and adult health and mortality. Individuals and cohorts exposed to disadvantaged circumstances in childhood are typically found to experience increased levels of morbidity, disability, and mortality when they are older adults (see Elo and Preston, 1992 and Mosely and Gray, 1993 for reviews).

African Americans are sometimes said to represent an exception to the prevailing positive correlation among death risks across different stages of life. Recorded death rates among African Americans have "crossed over" those of white Americans throughout the twentieth century. Despite much higher mortality at younger ages, African Americans have had lower recorded death rates than whites beginning at some age between 70 and 85 (Elo and Preston, 1994). A common explanation of this crossover is that only the hardiest blacks have survived to older ages; the weeding out of more vulnerable members of a cohort has resulted in an unusually healthy group of older blacks whose robustness is manifest in unusually low death rates. An alternative explanation is that data on older blacks are flawed by age misreporting and that correction of these inaccuracies would eliminate the crossover (Preston et al., 1996).

This paper investigates the association between social and economic conditions in childhood and the probability of surviving to age 85 among African Americans. It uses a unique case-control approach in which blacks who survived to age 85+ in 1985 are traced to their records from the censuses of 1900 or 1910, when they were children. They are then matched to a set of black children enumerated at the same age and census in order to identify childhood characteristics predictive of survival to age 85. Special attention is paid to whether factors associated with higher levels of child mortality are positively or negatively associated with survival to age 85.

Relations Among Death Probabilities across the Life Cycle

Will children who have been exposed to harsher health environments in childhood be more or less likely to survive from childhood to advanced ages? There are at least four mechanisms linking childhood conditions with adult mortality that would suggest an answer to this question. They fall conveniently into the typology shown in Table 1. Two mechanisms would suggest that harsher health conditions in childhood would be associated with higher adult mortality and two with lower. Within each direction of influence, one mechanism is direct, representing a physiological influence of childhood health environment on adult mortality, and one is indirect and non-physiological.

A direct relationship that would produce a positive link between childhood death probabilities and adult death probabilities can be termed “scarring.” Certain conditions and diseases acquired in childhood may, in a sense, permanently impair the survivors and leave an imprint on death rates at all subsequent ages. For example, tuberculosis, hepatitis B, and rheumatic heart disease are diseases that are often acquired in childhood but that manifest themselves in elevated death rates throughout life (Elo and Preston, 1992). Low birth weight or growth retardation in childhood has also been hypothesized to affect death rates from chronic diseases, especially cardiovascular diseases and diabetes, in adulthood (Barker, 1992).

On the other hand, a direct mechanism that would produce an inverse association between childhood mortality risks and those of adulthood is acquired immunity. Individuals who are more frequently exposed to diseases to which immunity can be acquired, such as influenza, would be expected to have lower death rates from these diseases at older ages. That this possibility is more than academic in the context under consideration is indicated by analysis of Union Army troops in the Civil War. During their time in service, recruits from Ohio who were drawn from healthier areas were much more likely to succumb to diseases for which immunity can be acquired than were other recruits (Lee, 1996).

According to two other hypothesized mechanisms, observed relations between death risks across the life cycle would be indirect, attributable to their joint association with other variables. An indirect mechanism that would produce positive associations can be termed "correlated environments." Those who are born into advantaged socioeconomic circumstances are likely to retain some of those advantages throughout life (e.g., Featherman and Hauser, 1978; Mare, 1990). Better incomes, diets, health habits, and access to health care would be expected to reduce mortality at all stages of life. Attributions of positive correlations across the life cycle to scarring mechanisms are often flawed by failure to account for the possibility that the environments of children and adults are correlated (Elford, Whincup and Shaper, 1991).

Finally, an indirect negative association between mortality risks in childhood and mortality risks at older ages would result from selection, the survival-of-the-fittest phenomenon noted above. An individual who survived unusually poor health conditions in childhood might be expected to be unusually well endowed with some set of (usually unobserved) genetic or congenital traits that enhance survival across the life cycle. Such a result would require that vulnerabilities to the diseases of childhood be positively correlated across individuals with vulnerabilities to the predominantly chronic diseases of adulthood. Thus, the two “indirect” mechanisms reflect the idea that death risks are affected both by environmental and individual-level characteristics, each of which is at least moderately persistent across the life cycle. One emphasizes environmental factors and the other, individual factors.

Health Conditions among Children at the Turn of the Century

In order to ascertain whether variables associated with child mortality operate in the same direction as those associated with postchildhood mortality, it is first necessary to know how these variables affected the risk of death among children. Child mortality was high at the turn of the century: the probability of dying before age five is estimated to have been .161 for whites and .255 for blacks (Preston and Haines, 1991: 86). Analysis of questions asked of women about the number of children they had borne and the number surviving in the census of 1900 showed that being black, having an illiterate mother, having an illiterate father, and living in a medium-sized or large city significantly elevated the risk of child death in a multivariate analysis (Ibid., chapter 4). The occupational categories of husbands that had the lowest child mortality were farmers and salesmen. A multivariate replication of this analysis using a larger data set drawn from the 1910 census of population showed that these variables were also significantly associated with child mortality in 1910 (Preston, Ewbank, and Hereward, 1994: Table 3B.1). In addition, the 1910 analysis investigated child mortality among women who had no husband present in the household; such women had significantly elevated risks of child death. Families living in a dwelling that they owned rather than rented had significantly lower child mortality in both censuses, although the ownership difference was significant in 1900 only for those living on farms.

Perhaps the only variable whose effect is surprising in the present context is city living, which is no longer associated with excess mortality. But at the turn of the century, it was, after race itself, the single most powerful variable affecting child mortality, whether measured by variance explained or by explained variance forfeited when the variable is excluded from a multivariate model (Preston and Haines, 1991: 170-76). The urban penalty is confirmed by death registration data. Condran and Crimmins (1980) estimate for nine death registration states in 1900 that life expectancy at birth was 44.6 years in urban areas and 54.1 in rural. Indirect confirmation of the rural/urban health difference is provided by data on height; in the late nineteenth century, farmers and men from rural areas were significantly taller than urban men (Costa and Steckel, 1995). Nevertheless, residents of urban areas were probably better fed than rural residents at the turn of the century by virtue of better transportation networks and better methods of food preservation (Preston and Haines, 1991: 44-47). Statistics on the nutritional status of drafted men during World War I indicate that rural draftees were 44% more likely to suffer from malnutrition than their urban counterparts and 6.5 times more likely to suffer from pellagra, a disease caused by niacin deficiency (Love and Davenport, 1920: 352-53). Therefore, urban residence becomes an especially interesting indicator of disease environment and not simply of general socioeconomic circumstances. The most plausible interpretation of the urban mortality disadvantage is that, before the deployment of effective public health measures, the greater density of habitation and sharing of common resources such as water supplies facilitated the spread of communicable diseases.

Research Design

We developed a cross-sectional case-control design to investigate the influence of childhood characteristics on survival to very advanced age, which we defined as age 85. African Americans surviving to ages 85+ represent an unusually successful group in terms of longevity. Among African Americans born in 1899-1903 who survived to age 8, only about 12.3% survived to age 85.[1]

We have drawn a sample of African Americans who survived to age 85 on January 1, 1985. In particular, we acquired the death certificates of all 1038 native-born African Americans who died at ages 85+ from 1 January to 14 January, 1985. The sample was drawn with the assistance of the National Center for Health Statistics, which identified the death certificates of all individuals who met our specifications. We then acquired copies of all death certificates from the individual states. A sample of death certificates has one overriding advantage for our purposes: it contains information on place of birth and father's and mother's name, essential information for purposes of linkage with early records. That decedents at ages 85+ form a representative sample of all persons aged 85+ is indicated by the absence of social and economic differentials in mortality at ages 85+. Preston and Taubman (1994: 286-87) report for both blacks and whites that "educational differences in mortality are virtually absent among the population aged 85 and older." The Appendix demonstrates that this conclusion applies to other social and economic characteristics as well.

An effort was made to link each of these 1038 individuals to their records in the U.S. Censuses of 1900 or 1910, when they were children. Record linkage relied heavily on microfilm copies of Soundex index cards, which catalog individuals enumerated in the census using a phonetic coding system.[2] Items on the death certificate that were critical to successful record linkage were the individual's name, father's name, mother's name, and state of birth. When available, other useful linkage information included city and county of birth, month of birth, and state of residence at the time of filing for a social security number, indicated by the first three digits of the number. When important death certificate information was missing or incomplete, a staff member of the Social Security Administration (SSA) attempted to secure it from SSA files. This procedure substantially reduced the number of subjects lacking critical linkage variables. Age and year of birth were not used as linkage variables because the data were originally collected to study the quality of age reporting on death certificates. For further details regarding record linkage procedures see Preston et al. (1996).

Soundex records from the census of 1900 or 1910 were located for 622 of the 1038 persons (59.9%) reported to be aged 85+ according to the death certificate. Of these linked subjects, 435 were linked to the census of 1900 and 187 were linked to the census of 1910. We excluded from the data set 12 of the linked subjects who were found to be younger than age 85 in 1985 according to their age at the early census. We also excluded 19 persons who were reported to be older than 19 at the census in order to focus the study on childhood conditions. Thus, 591 subjects passed these restrictions. However, the original census records (rather than the Soundex records) could not be found for 9 of these individuals. Therefore, our sample of longevous cases consists of 582 subjects, or 56.1% of the original group of decedents.

For each case identified in a childhood census, five African-American control subjects of the same reported age were drawn from the public-use sample of census manuscripts (PUMS file) for the same census year. Control subjects were drawn from the 1900 or 1910 PUMS files using random numbers applied to individuals stratified by age and census. We considered matching on sex as well but chose not to in order to investigate whether the research design produced reliable estimates of sex differentials in mortality. Sex is the only variable for which external information is available on survival, in the form of national life tables drawn from vital statistics data.

The average age of cases and controls at childhood enumeration is 8. Among members of the nonwhite cohort born in 1899-1903 who survived to age 8 but died before age 85, approximately 15% of deaths occurred between ages 8 and 25, 21% between 25 and 45, 29% between 45 and 65, and 35% between 65 and 85 (see note 1). Thus, the survival experience under investigation primarily reflects factors that manifest themselves above age 45, where about 64% of deaths occurred.

Our design has several advantages compared with that of previous research, in which adults surviving to advanced age are asked about their childhood conditions and certain health outcomes are tabulated according to self-reported early life conditions (Mare, 1990; Kaplan and Salonen, 1990; Lundberg, 1993; Peck, 1994; Lynch et al., 1994):

1) We do not rely upon retrospective reporting of early life conditions, which is subject to numerous errors and biases. Instead, we use information actually recorded in childhood.

2) We do not need to be concerned with issues of selectivity in reporting, the fact that those who are providing information in a health survey are not likely to be representative of the cohort from which they derived because some members of the cohort have already died. Instead,

we track the survival experience of a cohort all the way from a childhood enumeration to age 85.

3) We have verified the ages of those reported to have survived to a very old age and eliminated those whose age was overstated.

Statistical Procedures

Our goal is to estimate univariate and multivariate (partial) relative survival probabilities measuring the association between various childhood factors and survival to advanced age. When certain conditions are met, this goal can be achieved using cross-sectional samples of a birth cohort at two points in time, t_1 (when cohort members are children) and t_2 (when cohort members are of advanced age). In a cohort closed to all forms of migration, the maximum likelihood estimate of the probability of survival from t_1 to t_2, P, is estimated by the ratio

$$P = N_2 / N_1 \qquad (1)$$

where N_t is the number of cohort members alive at time t (DeGroot 1975: 199-200, 284-85). However, when the cohort is identified in two cross-sectional samples, we could use equation (1) to estimate the probability of survival only when the same sampling fraction is employed at both t_1 and t_2 (assuming that migration is trivial). When the sampling fractions differ, P is estimated by multiplying N_2 by an adjustment factor k, where $k > 0$, that is reflective of the different sampling fractions:

$$P = k{*}N_2 / N_1 \qquad (2)$$

As in conventional case-control studies, we are interested in assessing *relative* risks. Therefore, knowledge of k is unnecessary because it cancels out in the ratio of two probabilities:

$$\begin{aligned} RSP &= P\{x{=}1\} \,/\, P\{x{=}0\} \\ &= (k^*N_2\{x{=}1\} \,/\, N_1\{x{=}1\}) \,/\, (k^*N_2\{x{=}0\} / N_1\{x{=}0\}) \\ &= N_2\{x{=}1\}^* \, N_1\{x{=}0\} \,/\, N_1\{x{=}1\}^* \, N_2\{x{=}0\} \qquad (3) \end{aligned}$$

where *RSP* is the relative survival probability and x is a dichotomous explanatory variable.

Equation (3) is the cross-product ratio, the basic measure of association in case-control studies. In our statistical analysis, we use this measure to estimate relative survival probabilities for our explanatory variables. Because the cross-product ratio is directly linked to parameters in logit regression (Agresti, 1990), we use logit regression to estimate partial relative survival ratios in multivariate models that control for potentially confounding factors. Therefore, our analytical approach employs the conventional statistical methods of case-control studies (Breslow & Day, 1980). However, in this case the cross-product ratio estimates the relative probability of survival rather than the more conventional relative odds ratio. Put differently, we are predicting whether a particular observation is a case (a survivor) or a control (a randomly-chosen age-matched child in the same early census) according to the child's characteristics. Using logistic regression to make this prediction provides the (log of the) relative survival probabilities. As appropriate when any characteristics of the cases are explicitly matched to those of controls, we use conditional logistic regression analysis to estimate relative survival probabilities (Breslow and Day, 1980).

Sample Selection Bias

A potential source of bias in estimating relative survival probabilities as in equation (3) is that we were only able to link 59.9% of the 1985 death certificate sample to records in an early census. If those not linked have characteristics that differ systematically from those who were linked, bias would be introduced.

The Appendix investigates the extent to which characteristics available on the death certificate predicted successful linkage. A logistic regression equation predicting the likelihood of linkage was not significant at a 5% level. Nevertheless, two sets of variables achieved significant coefficients: state of birth and marital status. It is likely that state of birth differences reflect the relative completeness of the censuses in the various states. Because state of birth is associated with linkage success, we controlled for residence state in the early census to avoid any biases that may be produced by correlations between this variable and variables of interest in the multivariate analyses. However, we do not attempt to interpret state of residence coefficients because any real effects would be confounded with differentials in linkage rates. The significant marital status contrast between widowed (the largest category) and "not reported" is likely to reflect incompleteness of information on the death certificate in the latter case, a factor that would impede linkage success. It is not clear whether this variable is in any way associated with early childhood conditions. Reassuringly, the main death certificate variable indicative of socioeconomic standing, occupation, was not associated with significant differentials in linkage success.

The Appendix demonstrates that weighting our cases (and their corresponding controls) by the inverse of their probability of linkage, estimated from the logistic regression equation in Appendix Table 2, has a very small effect on estimated relative survival probabilities for the variables investigated below. In other words, it does not appear that differentials in linkage success, to the extent that we can observe them from information on the death certificate, play a major role in the relative risk estimates shown below.

Data and Variables

We constructed a data file that contains individual- and household-level information from census manuscripts for cases and controls, endeavoring to code information for cases in the same way that information was coded in the PUMS for controls. The public use samples were drawn from the Integrated Public Use Microdata System (IPUMS), Version 1.0 (Ruggles and Sobek, 1995). Since we were dealing with two censuses, it was necessary to employ comparable coding systems for both census years. Fortunately, the schedules used in the censuses of 1900 and 1910 were virtually identical except for the occupational fields. The engineers of the IPUMS project had already established comparable coding systems for many variables in 1900 and 1910, and we constructed comparable codes for the remainder.

Children were linked to mothers, fathers, and household heads primarily through relation-to-head codes. We used the same protocol used to establish a linkage in the IPUMS, even though the death certificate sometimes supplied information that would have supported an alternative link. Names of incorporated places were converted into population size categories by reference to published data (U.S. Bureau of the Census, 1910). Unincorporated places were put together with the smallest incorporated places, those smaller than 1000, in a category labeled "rural". We also added a variable indicating county-level population density in 1900 derived from a special data file of historical census statistics (ICPSR 0003). Three density categories are employed: low, medium, and high density. Low density counties are those below the median density for all counties in the sample ($\leq$42.2 persons/sq. mi); medium density counties were between the 50th and 90th percentiles (>42.2 and $\leq$132.6 persons/sq. mi); and high density counties were the top 10 percentile (>132.6 persons/sq. mi).

Coding of household and individual-level variables was for the most part straightforward. For purposes of this analysis, an individual was considered literate if he or she could both read and write. Occupational codes used in the IPUMS were applied to cases and controls and grouped into four major occupational categories: farmers (most of whom were sharecroppers or cash tenants); farm laborers; skilled blue-collar and white-collar occupations; and semi- and unskilled occupations, which includes unclassifiable and unspecified occupations.[3]

Univariate Results

Tables 2 and 3 present results from univariate conditional logit regressions that predict the probability of survival from childhood to age 85. They also show the percentage distributions of cases and controls among the categories of each variable.

Residential Characteristics. The relations between residential characteristics and the probability of surviving to age 85 shown in Table 2 are consistent with those for childhood mortality. Childhood residence on farms, in smaller places, and in less dense counties is associated with greater postchildhood survival chances, although the association with size of place is not significant. Children who did not live on farms were only 68% as likely to survive to age 85 as were farm children, a highly significant differential. Whether the farm was owned or rented was not predictive of subsequent mortality, but ownership of the dwelling was a significant factor for non-farm dwellings. The highest mortality category was children who lived in a non-farm rental unit. This was also true of children who died before 1900 (Preston and Haines, 1991: 144).

Having both parents present in the household increased a child's probability of survival to age 85. Children who lived in single-parent households had survival probabilities only 66.3% as high, and children who did not live with either parent only 48.6% as high, as children living in households with both parents present. Separate tabulations for mothers and fathers show that the presence of each was powerfully and significantly associated with subsequent survival.

The only two previous studies that have investigated this factor also found significant effects of family disruption during childhood on subsequent health. Using retrospective reports in Sweden, Lundberg (1993) found that living with only one parent and reporting that there had been "dissension" in the family were both significantly associated with self-reported poor health in adulthood. The associations with adult mortality during a 3 1/2-year follow-up period were in the same direction for both variables but were not significant. Schwartz et al. (1995) investigated adult mortality among the Terman sample of gifted white children in California born around 1910. Of the many childhood social and psychological factors investigated in the study, the largest mortality differential was associated with parental divorce. Children of divorced parents had a subsequent mortality hazard rate that was 30-40% higher than that of other children.

Clearly, our results provide additional support for this link at a different end of the social scale. However, a cautionary note is in order. It is possible that the estimated advantage of living with both parents is overestimated in our results because our chances of linking survivors to an early census were greater if both parents were present in the household. The analysis in the Appendix bears on this issue. Appendix Table 2 shows that death certificates with missing or incomplete information on mother's and father's names, conditions that are presumably more likely when parents were missing from the household, were in fact less likely to be linked to an early census. However, neither variable is a significant predictor of linkage success. Furthermore, Appendix Table 3 shows that the coefficients representing the effect on mortality of presence of mother or father in the household were scarcely altered when cases and controls were weighted inversely by the probability of linkage success. This probability reflects the availability of parental names as well as other characteristics on the death certificate. Thus, it appears that any bias produced by differentials in linkage success is small.

Characteristics of the household head. Table 3 includes univariate results for the household head's, mother's, father's, and subject's characteristics. If a husband and a wife were both present in the household, the convention of census enumeration at the turn of the century was to list the husband as head of household. Among the cases, 81.3% lived in households headed by their fathers compared with 73.9% of the controls. Consistent with results just reported, children growing up in such households had probabilities of survival that were significantly greater than those of other groups. So were those of children in households headed by a married person.

More than half of the cases (54.5%), but less than half of controls (46.9%), lived in households with a literate head. Living with an illiterate head lowered a child's probability of survival by 27%. The investigation of the Terman sample found no relationship between adult survival and mother's or father's education, but the variance of education in this sample of gifted children was small (Schwartz, et al., 1995). The only other study to investigate the influence of parental education on adult health, a study of the 1946 British birth cohort, found a significant negative relationship between maternal and paternal education and the probability of being in very poor health at age 36 (Kuh and Wadsworth, 1993).

About half of the children in the censuses of 1900 and 1910 lived in a household headed by a farmer. These children had a higher chance of reaching age 85 than those in any other occupational group. The few offspring of skilled workers were not significantly disadvantaged, but children of unskilled workers and farm laborers were; their survival chances were only 69% and 60% of those of farmers' children, respectively.

The differential between offspring of farmers and farm laborers may be economic in origin. A Finnish study has shown that offspring of small farmers and the landless in East Finland (but not West Finland) had significantly higher adult death rates from coronary heart disease than offspring from larger farms (Notkola et al., 1985). However, economic histories of the black population do not suggest that farmers earned higher incomes than farm laborers at the turn of the century. For the most part, black farmers did not own the land they farmed but rented it or shared its products with the owner. Black farm laborers, on the other hand, earned wages within 10% of those of their white counterparts. Their mobility allowed them to seek out the best-paying opportunities. Most farm laborers were unmarried and highly mobile; marriage and parenthood would typically force them into a sharecropping arrangement, which often involved a reduction in income (Wright, 1986: 94-98).

One economic advantage of sharecropping is that risks of bad weather and poor crops would be shared with the farm owner. Sharecropping may have also permitted the building up of creditworthiness and capital. Therefore, it is possible that the disadvantage faced by children of farm laborers was connected more with the rootlessness and riskiness of the occupation than with any persistent shortage of resources.

Mother's characteristics. Consistent with results shown above, survival probabilities were much lower for children whose mothers were divorced/widowed or never married than for mothers who were married. We have grouped the widowed with the divorced because of evidence that widowhood was seriously overreported for black women in the census of 1910, especially when a child but no husband was present in the household (Preston, Lim, and Morgan, 1992). Table 3 also shows that children of literate mothers had a 36% (1/.735 - 1) greater chance of surviving to advanced ages than children of illiterate mothers.

The censuses of 1900 and 1910 asked ever-married women how many children they had borne and how many of them had survived to the time of the census. From this information, we were able to construct an index of mortality among the subject's siblings. We subtracted one from the reported number of children ever borne (B) and the number surviving (S) to account for the survival of the subject to the time of the census. We then used information on the age of the mother, an indicator of the duration of children's exposure to the risk of death, to construct an expected proportion dead among her children (p) (see Preston and Haines, 1991, for details on the construction). The ratio of the reported number of deaths (B-S) to the expected number ((N-1)p) provided an index of mortality among the subject's siblings. Because the distribution of the variable is sharply discontinuous, we have dichotomized it into none or fewer than expected (<1.0) and more than expected (>1.0) categories. Results indicate that when mortality among siblings is greater than expected the subject had a significantly lower chance of surviving to advanced ages by some 20%. Because the childhood mortality experience of siblings reflects to some degree the same household-level mortality hazard, this finding is an indirect indication that mortality risks are positively correlated over the life course.

Father's characteristics. Father's literacy appears to play an even greater role in longevity than mother's literacy, with children of illiterate fathers having a relative survival probability of 0.664 versus 0.735 for children of illiterate mothers. The results for the father's occupation are similar to those for heads (since most heads were fathers), again indicating an advantage for those children with farmer fathers and a disadvantage for offspring of farm laborers and unskilled workers. These results clearly underscore the importance of the father's socioeconomic status for survival to extreme old age.

It should be noted that we also examined the age of the father and of the mother at the subject's birth, a variable constructed by subtracting the subject's age from that of his or her parent. Neither variable showed any relationship to the subject's longevity. A recent article in <u>Science</u> (1997) reported unpublished results of Gavrilov and Gavrilov indicating that females with very old fathers (aged 50+) had lower life expectancies. We find no such relationship; offspring of fathers aged 50+ had a higher chance of reaching age 85 than offspring of fathers below age 20 or aged 40-49, and were within 1% of the survival chances of fathers in the modal category 20-29 (results not shown).

Subject's characteristics. In the research design section, we mentioned that cases and controls were matched on age, race, and census at childhood enumeration. They were not matched on sex so that we could perform a validity test of our results. The only characteristic on which we have survival information for the pertinent cohorts of African Americans is sex. The National Center for Health Statistics (Moriyama and Gustavus, 1972) produced a cohort life table for nonwhites born in 1899-1903. This life table survived the cohort to age 68 and has been extended to age 85 using period life tables (see footnote 1). The probability of surviving from age 8 (the mean age of our subjects at census enumeration) to age 85 was .15163 for females and .08783 for males, giving a male/female survival ratio of .579. With females as the reference group, we find that the relative survival chance for males in our case-control study is .594 (Table 3). Thus, our sex differential in survival is quite close to that in the cohort life table produced from vital statistics.

The 1900 and 1910 censuses collected data on reading and writing ability for children aged 10 and older. Though nearly 60% of cases and controls were below age 10, investigating the association between the subject's own literacy and his or her survival chances seemed worthwhile. Table 3 shows that cases were more likely than controls to be able to read. The survival probability of children unable to read was only 68.8% as high as that of those who could.

Although living on a farm appears to be an important factor for survival, we hypothesized that the farm advantage would be greater for males than for females. Such a result has earlier been shown in Sweden for persons aged 16-74 in 1980: adult male offspring of farmers had a 21% lower death rate than male offspring of non-farmers, whereas farm background had a slightly adverse (although statistically insignificant) effect for females (Peck, 1994: Table 1). To examine this interactive relationship, we combined sex and farm background in a joint variable. The farm/nonfarm and male/female dichotomies divide the population almost ideally into quartiles; between 23% and 27% of the controls appear in each of the four cells, providing a robust basis for estimation. With females who lived on farms as the reference group, females who did not live on farms were 16.2% less likely to survive to age 85, a disadvantage that is not statistically significant. Males who lived on farms were 26.9% less likely to have been in the longevous group compared with females who lived on farms, whereas males who did not live on farms were 66.4% less likely to survive. When males off farms are made the reference group (not shown), those growing up on the farm were 2.17 times more likely to survive to age 85 than males growing up off the farm, a differential that is significant at $p < .001$.

This survival differential between males who grew up on and off the farm is extraordinary. What might account for it? The result is unlikely to reflect important biases in linkage probabilities, since any farm/non-farm linkage differential for males also ought to apply to females. It also appears unlikely that disinvestment in the education of farm girls relative to farm boys is responsible since, among subjects aged 10+, girls had higher literacy rates than boys in both farm and non-farm settings (results not shown).

One reason for this result may be that males who grew up on farms were more likely to become farmers themselves. Male death rates for farmers have been below the average for all occupational groups in all investigations of which we are aware, although the advantage appears smaller for blacks than for whites (e.g., Moriyama and Guralnick, 1956: Tables 2 and 3; Kitagawa and Hauser, 1973: Table 3.2). Some indication of intergenerational occupational retention can be gleaned from our sample of death certificates. Of the male decedents whose usual occupation was listed as "farmer" on the death certificate, 84.4% lived on a farm in childhood, compared with 70.9% of all male decedents. Whether farm work is protective because of persistent demands for physical exertion, less damaging life styles, selection ("healthy worker") effects, or some other factor is unknown. The Swedish study had data on various intervening factors that might explain the survival advantage for males with a farm background, including exercise levels, smoking, living alone, and adult occupation. Introducing all these variables into a regression equation, however, left the male farm background advantage unchanged at 21% (Peck, 1993: Table 4).

It is possible that males raised on farms were less likely to live in cities during adulthood and were therefore spared the markedly higher mortality of urban black men. The adult mortality penalty from urban living for cohorts born around the turn of the century is shown in Table 4, which presents selected age-specific death rates by urban and rural residence for nonwhites in 1940. During this era, the death rates in urban areas were consistently higher than those in rural areas for both black males and females. However, the urban/rural differentials for males substantially exceeded those for females. It is useful to note that a 40% difference in age-specific death rates, about the level of urban/rural differences for males shown in Table 4, produces a survival differential from age 8 to age 85 of about .38 for urban residents relative to rural residents.[4] This relative survival ratio is even lower than that which we have uncovered for males. Thus, the evidence on differentials in adult death rates near the middle of the century is roughly consistent with the very large discrepancy in survival rates that we have uncovered.

Another way to interpret the sex/farm results in Table 3 is that, for persons raised on farms, males were 73% as likely to survive to age 85 as females; for persons reared off the farm, males were only 40% as likely to survive as females. Thus, our results may bear on the very large contemporary sex differential in black mortality, one of the largest sex differentials in any population. In 1990, life expectancy at birth for black males fell short of that of black females by 9.1 years, compared with 6.7 years for whites in that year and 2.5 years for blacks in 1900-02 (National Center for Health Statistics, 1994: Table 6.4). The widening of the sex differential for blacks in the course of the century is clearly consistent with the reduced fraction of the population having farm origins, although many other factors are doubtless involved in this change.

Multivariate Results

We used conditional logit regression to estimate partial relative survival probabilities adjusted for the confounding effects of other variables. In conducting the multivariate analysis, we faced one problem of multicollinearity. Because a large fraction of farmers lived on farms and a large fraction of farm laborers did not, estimated coefficients became very unstable when both farm residence and the farmer/farm laborer occupational classifications were simultaneously introduced. Our solution was to combine the occupational categories of farmer and farm laborer while retaining the farm/non-farm distinction. It should be recognized that results pertaining to the farm/non-farm variable may reflect occupational distinctions as much as residential ones. Urban status was not significant when population density was included and has been dropped from the models.

Table 5 presents results from the multivariate analyses.[5] Three models were estimated. Model I was estimated for the full sample and includes residential/household and head of household characteristics; Model II includes only those subjects whose mother was present in the household and estimates the effect of household and mother's characteristics; and, Model III includes only those subjects who lived with both parents and examines both mother's and father's characteristics. All three models include subject's sex and reading ability and control for state of residence at the early census (state of residence results not shown).

Model I shows that the variables that were significant in univariate analysis tend to operate in the same direction and to retain their significance in multivariate analysis. Those living in counties with the highest density have a significantly lower probability of surviving (RSP=.630). Children living on farms had a 46% higher probability of surviving to age 85 than those off the farm. A sex/farm residence interaction term is highly significant. The estimate implies that males who lived on a farm were 2.48 times more likely to survive to age 85 (1.697*1.460) than males who lived off a farm, once other variables are controlled. This compares with an advantage of 2.17 times in the univariate analysis. Thus, the relative survival advantage that males gain from a farm background is markedly increased when confounding factors are controlled.

Children living with neither parent have a significantly lower probability of survival than those living in households headed by fathers. In contrast to univariate results, however, children living in households headed by their mothers actually have slightly higher probability of survival than those with father heads. If the mother is unmarried, however, the advantage disappears (RSP=.904=1.189*.791). If attention is confined to households with a mother present, as in Model II, survival chances are significantly lower when the mother is unmarried (RSP=.601).

Once the categories of farmer and farm laborer are combined and other variables are added to the model, there are no significant survival differences associated with head's occupation. However, the literacy of the head continues to influence a child's survival chances. Children from households with illiterate heads are only 73% as likely to reach age 85 as when the head is literate. A wealth effect is also apparent in the significant coefficient on homeownership, which is shown (through the interactive variable combining ownership and farm residence) to be limited to persons living off the farm.

The basic pattern of results is retained when analysis is restricted to children living with their mother (Model II) or with both parents (Model III). Growing up on a farm continues to be an advantage, especially for males, and owning a home retains its advantage if the home is not a farm. Both mother's and father's literacy contribute significantly to a child's survival chances;

when both parents are present, the literacy of the father is slightly more important than that of the mother. The index of mortality among siblings, included in Models II and III, has an RSP nearly identical to that of the univariate result, suggesting that children with an excessive number of sibling deaths have an 18-20% lower survival probability. However, this variable is not significant in the multivariate analysis.

Discussion

A principal question addressed in the paper is whether and how early life factors that are predictive of mortality in childhood influence postchildhood mortality. The answer is that, in general, the socioeconomic and residential factors examined here influence mortality in the same direction in both childhood and adulthood. Illiteracy of father and mother, dense living conditions, and absence of the father from the household raised child mortality at the turn of the century and reduced the chances of surviving from childhood to age 85. The consistency of results is made more secure by the fact that the childhood analysis and the adult analysis apply to roughly the same cohort, persons born between 1885 and 1900.

In order to cast additional light on the relationship between childhood and adult mortality, we have created a new variable for each child in the study who was living with both parents. This variable is the *predicted* relative level of child mortality in a family with the set of characteristics recorded in the census. It is constructed by applying to these characteristics a micro-level multivariate regression equation predicting relative child mortality levels (actual vs. expected child deaths) in 1900 (Preston and Haines, 1991: Table 4.4).[6] The predicted level of child mortality functions as a summary measure of household conditions as they bear on the risk of child death.

In particular, it is a weighted average of those conditions, with higher weights (i.e., larger coefficients) attached to those circumstances that are most important in predicting child mortality.

The predicted level of child mortality for the household is a continuous variable with a mean of 1.00 for households in the 1900 census. When the binary response variable indicating case or control status is regressed against this variable using conditional logit regression, its coefficient is -.901 (significant at $p < .001$). This large negative coefficient implies that when the predicted level of child mortality varies from, say, 0.5 to 1.5 of its average value, the probability of surviving to age 85 declines by 59.4% (100*[exp(-.901) - 1]). Thus, children who were exposed to the most unhealthy childhood environments were far less likely to reach age 85 than those living in more favorable environments.

In terms of the schema of Table 1, our results suggest that the correlation between mortality risks in childhood and those in adulthood is positive rather than negative. We cannot determine whether the positive relations result from some form of scarring or from correlated environments, the two mechanisms identified on the table, because we lack data that would help us identify the nature of the linkages. The fact that such a large sex difference was uncovered in the impact of farm background provides weak evidence against the importance of scarring, since it seems unlikely that early environments could have had radically different physiological effects on boys and girls.

For one purpose, the source of the positive association may be immaterial. Whichever mechanism dominates, the level of child mortality would have predictive value for the level of adult mortality in the same cohort. There is no reason to expect, as the survival of the fittest hypothesis would predict, that the decline in child mortality will dampen the decline in adult mortality because more and more impaired persons are surviving. Quite the contrary: the steady advance of child survival during the twentieth century augurs well for adult survival in the twenty-first.

NOTES

[1] This estimate is based upon a cohort life table for nonwhites born during the period 1899-1903 (Moriyama and Gustavus, 1972), which tracked survival to age 68. The survival history of the cohort to age 85 is then completed by using official period life tables for nonwhites in 1973 (ages 68-75) and 1978 (ages 75-85) (National Center for Health Statistics, 1975, 1980). Because the survival rates for nonwhites include the experience of Asian and Pacific Islanders, a group with exceptionally low mortality, the percentage of African Americans surviving to age 85 is almost certainly lower than the above estimate.

[2] These records are organized by state of residence and then alphabetically by the surname and first name of the household head. Soundex index records list the household's geographic location and the full name, relationship to the household head, state of birth, and age of each household member. Index records for the 1900 census also include month and year of birth. Persons with surnames different from the household head are listed on separate “individual” cards.

[3] We gratefully acknowledge the consultation of Ann Miller, who assisted in the design of the occupational classification scheme used in this study.

[4] As noted earlier, the male survival probability from age 8 to age 85 in the relevant cohort life tables was .08783. A 40% increase in death rates at all ages from 8 to 85 would produce a survival probability of $.08783^{1.4} = .03320$. Thus, cumulated over a lifetime, the urban penalty

would produce a survival ratio in urban relative to rural areas of .378 (=.03320/.08783). A similar application of the approximate female urban excess of 25% in Table 4 to the female survival probability of .15163 produces a much higher relative survival probability of .624, which is also below the female value that we have estimated. These hypothetical calculations are designed to show that the large differentials in survival to age 85 shown in Table 3 are not discrepant with the more moderate urban/rural adult mortality differentials shown in Table 4.

[5] In order not to lose observations in which information about certain characteristics was missing, we have created missing data categories for each variable but do not present results for these categories.

[6] The composite child mortality variable was constructed for each record (case or control) in which both parents were present in the household. Preston and Haines (1991; Table 4.4) estimated a weighted least squares regression model that predicted the ratio of actual to expected child deaths for native-born women in the 1900 census. The coefficients from their model for all native-born women in the 1900 PUMS were combined with individual-level data on mother's, father's, household, and ecological characteristics for individuals in our sample. Certain characteristics were omitted, such as ancestry, ability to speak English, and mother's employment status, because they were unavailable or not relevant to our sample. These omissions affect only the constant term in the predicted child mortality index and as a result do not affect the coefficient or interpretation of the composite child mortality variable.

Table 1. Typology of Relations between Mortality Risks in Childhood and Mortality Risks in Adulthood

Direction of Relation	**Direct, Physiological**	**Indirect, Associational**
Positive	Scarring	Correlated Environments
Negative	Acquired Immunity	Selection

TABLE 2. Relative Probabilities of Surviving from Childhood to Age 85 According to Characteristics in U.S. Censuses of 1900 and 1910: County, Household, and Household Head Characteristics

		% Distribution of Sample		Maximum Likelihood Estimate of Relative Survival Probability	
Characteristic		**Cases (N=582)**	**Controls (N=2,910)**	**Relative Survival Probability**	**95% Confidence Interval**
Residential characteristics					
Farm status	Non-farm	37.29	46.49	---	---
	Farm	62.03	52.41	1.478**	1.229-1.777
Urban status:	Rural	81.79	79.59	---	---
	1,000-24,999	11.00	11.62	0.921	0.691-1.226
	25,000+	7.22	8.80	0.799	0.568-1.123
Population density:	Low density	52.58	49.07	---	---
	Medium density	40.21	39.55	0.943	0.782-1.137
	High density	6.70	10.62	0.583**	0.407-0.834
Homeownership:	Rent	66.67	70.14	---	---
	Own	28.52	24.36	1.229*	1.007-1.502
Farm/homeownership:	Own farm	17.53	14.98	---	---
	Rent farm	43.99	37.22	1.023	0.792-1.321
	Own home (non-farm)	10.65	9.21	1.008	0.705-1.441
	Rent home (non-farm)	22.68	32.16	0.605**	0.457-0.801
Household characteristics					
Parental Co-residency:	Both present	81.27	72.54	---	---
	One parent present	13.06	17.39	0.663**	0.511-0.861
	Neither present	5.67	10.07	0.486**	0.333-0.711
Mother in household:	Present	91.41	86.80	---	---
	Absent	8.59	13.20	0.610**	0.446-0.834
Father in household:	Present	84.19	75.67	---	
	Absent	15.81	24.33	0.574**	0.451-0.731

*p<0.05, **p<0.01; --- reference category.

Note: Percentages may not add to 100.00 due to missing data or rounding.

TABLE 3. Relative Probabilities of Surviving from Childhood to Age 85 According to Characteristics in U.S. Censuses of 1900 and 1910: Mother, Father, and Subject Characteristics

Characteristic		% Distribution of Sample		Maximum Likelihood Estimate of Relative Survival Probability	
		Cases	Controls	Relative Survival Probability	95% Confidence Interval
Household head's characteristics		**(N=582)**	**(N=2,910)**		
Relationship to subject:	Father	83.33	73.85	---	---
	Mother	7.73	9.90	0.681*	0.490-0.946
	Other relative	6.53	12.30	0.466**	0.329-0.662
	Subject (self)/non-relative	2.41	3.95	0.524*	0.294-0.933
Marital status:	Married	88.14	82.34	---	---
	Divorced/widowed	9.79	14.71	0.622**	0.464-0.833
	Never married	2.06	2.96	0.641	0.347-1.184
Literacy:	Literate	54.47	46.94	---	---
	Illiterate	44.50	52.03	0.735**	0.614-0.881
Occupational category:	Farmer	60.82	50.93	---	---
	Farm laborer	10.31	14.19	0.604**	0.449-0.813
	Skilled	5.33	4.60	0.975	0.647-1.471
	Unskilled	21.48	26.05	0.688**	0.550-0.859
Mother's characteristics		**(N=532[a])**	**(N=2,526[a])**		
Marital status:	Married	90.60	83.57	---	---
	Divorced/widowed	8.46	13.78	0.560*	0.402-0.780
	Never married	0.94	2.65	0.354*	0.142-0.885
Child deaths:	None/fewer than expected	63.91	60.21	---	---
	More than expected	27.63	31.39	0.803*	0.645-0.998
Literacy:	Literate	48.87	41.57	---	---
	Illiterate	48.50	55.54	0.735**	0.606-0.891
Father's characteristics		**(N=490[a])**	**(N=2,202[a])**		
Literacy:	Literate	58.37	49.73	---	---
	Illiterate	40.61	49.23	0.664**	0.541-0.815
Occupational category:	Farmer	63.67	55.77	---	---
	Farm laborer	10.41	14.53	0.577**	0.415-0.802
	Skilled	5.51	5.13	0.905	0.576-1.420
	Unskilled	19.39	22.98	0.730*	0.564-0.946
Subject's characteristics		**(N=582)**	**(N=2,910)**		
Sex:	Female	62.37	49.86	---	---
	Male	37.63	50.14	0.594**	0.493-0.714
Sex/living on farm:	Female, living on farm	35.57	26.15	---	---
	Female, not living on farm	26.46	23.13	0.838	0.661-1.062
	Male, living on farm	26.46	26.25	0.731**	0.578-0.924
	Male, not living on farm	10.82	23.37	0.336**	0.248-0.455
Reading ability (age 10+):	Can read	73.08	67.01	---	
	Cannot read	22.22	28.97	0.688*	0.488-0.968

[a] Sample sizes reflect the number of households where mothers or fathers are present.

*p<0.05, **p<0.01; --- reference category. Note: Percentages may not add to 100.00 due to missing data or rounding.

Table 4. Death Rates in 1940 for Nonwhite Males and Females by Urban/Rural Residence: Selected Age Groups[a]

Area of Residence	**Females**	**Males**	**Female/Male Ratio**
	Ages 35-44 (born 1895-05)		
Urban	1254.6	1478.0	0.849
Rural	1069.4	1088.5	0.982
Urban/Rural Ratio	1.163	1.358	
	Ages 45-54 (born 1885-1895)		
Urban	2342.7	2830.0	0.828
Rural	1810.3	1996.4	0.907
Urban/Rural Ratio	1.294	1.418	
	Ages 55-64 (born 1875-1885)		
Urban	4014.9	4748.7	0.845
Rural	3090.4	3212.0	0.962
Urban/Rural Ratio	1.300	1.478	

[a]Urban areas defined as cities, towns, and other incorporated areas containing 2,500 or more residents; rural defined as all other areas.
Source: Linder & Grove (1943: Tables 24 and III).

TABLE 5. Adjusted Relative Probabilities of Survival to Age 85 for Selected Childhood Characteristics

Characteristic		Model I Full Sample (N=3,492)	Model II w/ Mothers (N=2,851[a])	Model III w/ Mothers & Fathers (N=2,223[a])
Residential/Household characteristics				
Population density:	Low density	---	---	---
	Medium density	0.986	1.056	1.041
	High density	0.630*	0.735	0.738
Farm status:	Non-farm	---	---	---
	Farm	1.460*	1.442*	1.489
Homeownership:	Rent	---	---	---
	Own	1.546*	1.403	1.346
Farm*Ownership[b]		0.578*	0.620*	0.584*
Household head's characteristics				
Relationship to subject:	Father	---		
	Mother	1.189		
	Other relative	0.540**		
	Subject (self)/non-relative	0.566		
Martial status:	Married	---		
	Not married	0.791		
Literacy:	Literate	---		
	Illiterate	0.728**		
Occupational category:	Farmer/Farm laborer	---		
	Skilled	1.244		
	Unskilled	1.004		
Mother characteristics				
Marital status:	Married		---	NI
	Unmarried		0.601**	NI
Child deaths:	None/fewer than expected		---	---
	More than expected		0.802	0.815
Literacy:	Literate		---	---
	Illiterate		0.733**	0.769*
Father characteristics				
Literacy:	Literate			---
	Illiterate			0.687**
Occupational category:	Farmer/Farm laborer			---
	Skilled			1.121
	Unskilled			0.938
Subject's characteristics				
Sex:	Female	---	---	---
	Male	0.407**	0.417**	0.398**
Sex*Farm status[c]		1.697**	1.594*	1.619*
Log-likelihood		-973.14	-828.40	-667.52
Chi-square (df)		139.34 (33)	116.56 (29)	104.47 (33)

[a] Reflects effective sample size; some matched cases-control groups dropped from analyses due to no in-group variance.
[b] Interaction term has a value of 1 for persons who lived on farms that were owned, and 0 otherwise.
[c] Interaction term has a value of 1 for males who lived on farms, and 0 otherwise.
*p<0.05, **p<0.01; --- reference category; NI - not included because all mothers were married when the father was present.
Note: Estimates not shown for categories with missing values on a characteristic and for state of residence controls.

Appendix: Assessing Bias in the Sample of Linked Cases

We make the assumption that the survivors to age 85 in this analysis (the cases) are an unbiased sample of all blacks born before 1900 who survived to age 85 in 1985. There are two main circumstances that could have produced a biased sample: we have sampled from persons dying above age 85 in 1985, who may not be representative of the living population; and we have excluded from our study sample those whom we were unable to link to an early census record. This Appendix addresses both sources of bias.

1) The Implications of Drawing a Sample of Decedents

The existence of social and economic differentials in mortality would cause decedents to be a biased sample of the population from which they were drawn. We use data from the National Longitudinal Mortality Study Public Use File Release 2 (NLMS) to determine whether social and economic differentials are present among the oldest African Americans during the period 1979-90. The NLMS is a nationally-representative sample of the non-institutionalized population. (For a description of the NLMS sample see Rogot et al., 1992.) To test for important socioeconomic differentials, we estimated logit regression models predicting the log odds of dying during a 9-year follow-up period for persons aged 85 to 97 at the time of the initial survey. About 64% of NLMS subjects aged 85 to 97 died during the follow-up period.

Appendix Table 1 presents logit regression results for models predicting the log odds of death during the nine-year period for black Americans aged 85 to 97 (column 1) and for whites and black Americans aged 85 to 97 (column 2). Results for age and sex show that females enjoyed lower odds of dying than did males; further, the odds of death increased with age (although the large age coefficient was not statistically significant in the African American model, presumably because of the small sample size). Educational attainment, family income, and household size were not significantly associated with mortality in either model. Of the socioeconomic indicators included in these analyses, only Standard Metropolitan Statistical Area residence (SMSA) was associated with the risk of death in the African American model: African Americans who lived in central cities within SMSA's had slightly lower odds of dying than their counterparts who did not live in SMSA's (odds ratio of 0.71). Those living in non-central city areas of SMSA's also had lower odds of death than the non-SMSA reference group, but the coefficient was not statistically significant. Residence was not significantly associated with mortality in the full sample. These results indicate that the socioeconomic profile of extremely aged African American decedents differs little from that of the underlying population of survivors at ages 85+.

2) The Implications of Non-Linkage

A second concern about the representativeness of our cases stems from the non-linkage of decedents to early census records. Of the 1,038 African Americans aged 85+ in our death certificate sample, 582 (56.1%) were located in a 1900 or 1910 census schedule and also found to meet our set of census age criteria (i.e., age in the early census was less than 20 and census-implied age at death was greater than or equal to 85). Because these 582 individuals may differ in important ways from the population of interest, questions regarding non-linkage bias arise. To explore the depth of this potential problem, we used logit regression to estimate the predicted probability that a member of our original sample of 1038 decedents became a "case" in this study. Predictor variables were constructed from information provided on death certificates. Appendix Table 2 presents the estimated coefficients of the logit model.

Results presented in Appendix Table 2 indicate that this model has only weak predictive value, suggesting that non-linkage did not bias the sample of cases used in this paper in ways that are associated with variables that we can observe. Remarkably, the likelihood ratio χ^2 for the overall model is barely significant at even the $p < 0.10$ level: χ^2 (df 30) = 40.27; p = 0.0997. Of the variables included in the model, only state of birth had a non-trivial effect on the likelihood of inclusion in the study sample. The state-of-birth effect may reflect differences in the quality of census enumeration by state or migration patterns that affected linkage success. In the multivariate analyses presented in the body of the paper, we include controls for state of residence at the time of the childhood census to adjust for any bias that may have been produced by these geographic differentials.

The estimated coefficients for sex, age, marital status, occupational status, and two dummy variables indicating missing parental names (included because parental names aid census linkage) were relatively small and statistically insignificant. The one exception was "marital status not reported," which had a strong negative effect on the odds of linkage success (odds ratio of 0.15 compared with widowed persons, the most prevalent group). Unreported marital status appears to be a marker of poor quality death certificate information (e.g., erroneous reporting state of birth, parental names and/or other linkage information), reporting that would have reduced the chance of linkage to an early census.

We used the logit regression equation presented in Appendix Table 2 to estimate probabilities of linkage for each of the 582 cases used in our study. The inverse of the probability was used as a sampling weight in order to estimate weighted relative survival probabilities that adjust for selection bias. The weight estimated for a given case was also assigned to his or her five matched controls. Appendix Table 3 presents unweighted and weighted risk ratios estimated using logit regression. The relative survival probabilities presented here are univariate cross-product ratios estimated from conventional contingency tables (Feinberg, 1980). The similarity of the weighted and unweighted estimates provide no evidence of substantive bias resulting from non-linkage.

Appendix Table 1. Coefficients for Logistic Regression Models Predicting Death among Americans aged 85-97: 1979-1988

Characteristic	**Blacks Only (N = 821)**	**Whites & Blacks (N = 11,369)**
Age	0.6780	0.7726**
Age^2	-0.0036	-0.0038**
Female	-0.3992**	-0.6350**
African American		-0.2173**
Education		
0-8 years	0.0398	-0.0464
9-11 years	-0.1427	-0.0796
12 years	-	-
13-15 years	0.0543	-0.1084
16+ years	0.0245	-0.0781
Economic status		
Log of family income	0.1181	0.0081
Household size	-0.0508	0.0150
Residence		
Not in SMSA	-	-
SMSA, central city	-0.3441*	-0.0634
SMSA, not in central city	-0.2732	0.0112
Intercept	-31.5763	-36.7567
Log Likelihood	-540.55	-7118.85
Likelihood Ratio Chi^2 / df	26.073 / 11	636.0 / 13

*p<.05, **p<.01 (two-tailed tests). -- indicates reference category.
Source: National Longitudinal Mortality Study, Public Use File (Release 2).

Appendix Table 2. Coefficients for Logit Regression Model Predicting Census Linkage (N=1,038)

Characteristic	**Coefficient**
Sex	
Female	--
Male	0.2049
Death certificate age	-0.0089
Marital status	
Divorced	0.0314
Married	0.0125
Never-married	0.2276
Widowed	--
Not reported	-1.8886*
Occupational Status[a]	
White collar	0.0711
Skilled crafts	0.0103
Operative	-0.0581
Nonfarm laborer	--
Farmer	0.1908
Farm laborer	-0.9702
Service/domestic	0.1733
Homemaker	0.1657
Unclassifiable	0.4412
Not reported	-0.0390
State of birth	
Alabama	0.3530
Arkansas	0.4801
Florida	-0.4697
Georgia	0.3396
Louisiana	0.0997
Mississippi	0.2445
North Carolina	0.7097*
South Carolina	0.4906
Tennessee	0.9638*
Texas	0.5971*
Virginia	0.8410*
Non-Southern states	0.7553*
Other Southern states	--
Not reported	-0.6398
Father's name missing/incomplete	
No	--
Yes[b]	-0.2448
Mother's first name missing	
No	--
Yes	-0.2750
Constant	0.4684
Log Likelihood	-691.68
Likelihood Ratio Chi^2 *(df = 30)*	40.27

* p<.05 (two-tailed tests). -- indicates reference category.

[a]Based on the scheme used for the Census of 1970 (U.S. Bureau of the Census, 1971). White collar refers to professional, managerial, sales, and clerical workers.

[b]Indicates that the father's first and/or last name was missing on the death certificate.

Appendix Table 3. Unweighted and Weighted Relative Survival Probabilities for Selected Characteristics

Characteristic		Unweighted Relative Survival Probability	Weighted Relative Survival Probability
Household characteristics			
Urban status:	Rural	--	--
	Urban	0.8684	0.8535
Home ownership:	Rent	--	--
	Own	1.2316*	1.2430*
Farm status:	Farm	--	--
	Non-farm	0.6775**	0.6784**
Head of household Characteristics			
Sex:	Male	--	--
	Female	0.6474**	0.6563**
Marital status:	Married	--	--
	Divorced or widowed	0.6220**	0.6174**
	Never married	0.6517	0.6861
Literacy:	Literate	--	--
	Illiterate	0.7372**	0.7449**
Occupational status:	Farmer	--	--
	Farm laborer	0.6082**	0.6283**
	Skilled	0.9685	0.9850
	Unskilled	0.6904**	0.6734**
Mother characteristics			
Mother present:	Yes	--	--
	No	0.6182**	0.6058**
Marital status:	Married	--	--
	Divorced or widowed	0.5663**	0.5738**
	Never married	0.3268*	0.3018*
Literacy:	Literate	--	--
	Illiterate	0.7426**	0.7678**
Child mortality ratio (continuous scale)		0.8907*	0.8859*
Father characteristics			
Father present:	Yes	--	--
	No	0.5839**	0.5874**
Literacy:	Literate	--	--
	Illiterate	0.7029**	0.7013**
Occupational status:	Farmer	--	--
	Farm laborer	0.6273**	0.6448**
	Skilled	0.9404	0.9770
	Unskilled	0.7390*	0.7156*
Subject characteristics			
Sex:	Male	--	--
	Female	0.6000**	0.5678**
Reading ability:	Can read	--	--
	Cannot read	0.7033*	0.7457

* p<.05, ** p<.01 (two-tailed tests). -- indicates reference category.

REFERENCES

Agresti, Alan. 1990. Categorical Data Analysis. New York: John Wiley & Sons.

Barker, D.J.P., editor. 1992. Fetal and Infant Origins of Adult Disease. British Medical Journal, London.

Breslow, Norman E. and N. E. Day. 1980. Statistical Methods in Cancer Research, Vol. I: The Analysis of Case-Control Studies. IARC Scientific Publications, Lyon.

Condran, Gretchen A. and Eileen Crimmins. 1980. "Mortality Differentials between Rural and Urban Areas of States in the Northeastern United States 1890-1900" Journal of Historical Geography Vol. 6(2): 179-202.

Costa, Dora L. and Richard H. Steckel. 1995. "Long-Term Trends in Health, Welfare, and Economic Growth in the United States" National Bureau of Economic Research. Historical Paper 76, Cambridge, MA.

DeGroot, Morris H. 1975. Probability and Statistics. Reading, MA: Addison-Wesley.

Elford, Jonathon, Peter Whincup and A.G. Shaper. 1991. "Early Life Experience and Adult Cardiovascular Disease: Longitudinal and Case-Control Studies" International Journal of Epidemiology Vol. 20 (4): 833-844.

Elo, Irma T. and Samuel H. Preston. 1992. "Effects of Early-Life Conditions on Adult Mortality: A Review" Population Index Vol. 58(2):186-212.

Elo, Irma T. and Samuel H. Preston. 1994. "Estimating African-American Mortality from Inaccurate Data" Demography Vol. 31 (3): 427-458.

Featherman, David L. and Robert M. Hauser. 1978. Opportunity and Change. Academic Press, New York.

Feinberg, Stephen E. 1980. The Analysis of Cross-Classified Categorical Data, 2nd ed. MIT Press, Cambridge MA.

Feinstein, Jonathon S. 1993. "The Relationship between Socioeconomic Status and Health: A Review of the Literature" The Milbank Quarterly Vol. 71(2):279-322.

Haines, Michael R. and Samuel H. Preston. 1997. "The Use of the Census to Estimate Childhood Mortality." Historical Methods Vol. 30 (2): 77-96.

Kaplan, George A. and Jukka T. Salonen. 1990. "Socioeconomic Conditions in Childhood and Ischemic Heart Disease during Middle Age" British Medical Journal Vol. 301: 1121-1123.

Kuh, D.J.L. and M.E.J. Wadsworth. 1993. "Physical Health Status at 36 Years in a British National Birth Cohort" Social Science and Medicine Vol. 37. No.7: 905-916.

Kitagawa, Evelyn M. And Philip M. Hauser. 1973. Differential Mortality in the United States: A Study in Socioeconomic Epidemiology. Harvard University Press, Cambridge, MA.

Lee, Chulhee. 1996. "Socioeconomic Background, Disease, and Mortality among Union Army Recruits: Implications for Economic and Demographic History" Manuscript, Department of Economics, University of Chicago.

Linder, Forrest E. and Robert D. Grove. 1943. Vital Statistics Rates in the United States: 1900-1940. Government Printing Office, Washington, DC.

Love, Albert G. and Charles B. Davenport. 1920. Defects Found in Drafted Men. Government Printing Office, Washington, DC.

Lundberg, Olle. 1993. "The Impact of Childhood Living Conditions on Illness and Mortality in Adulthood" Social

Science and Medicine Vol. 36(8): 1047-1052.

Lynch, John W., George A. Kaplan, Richard D. Cohen, Jussi Kauhanen, Thomas W. Wilson, Nicholas L. Smith, and Jukka T. Salonen. 1994. "Childhood and Adult Socioeconomic Status as Predictors of Mortality in Finland" The Lancet Vol. 343: 524-527.

Mare, Robert D. 1990. "Socio-economic Careers and Differential Mortality among Older Men in the United States" pp. 362-387 in Jacques Vallin, Stan D'Souza, and Alberto Palloni, editors. Measurement and Analysis of Mortality: New Approaches. Clarendon Press, Oxford.

Moriyama, Iwao M. and L. Guralnick. 1956. "Occupational and Social Class Differences in Mortality" pp. 60-73 in Milbank Memorial Fund, Trends and Differentials in Mortality. Milbank Memorial Fund, New York.

Moriyama, Iwao M. and Susan O. Gustavus. 1972. Cohort Mortality and Survivorship: United States Death-Registration States, 1900-68. Analytic Studies Series 3 No. 16. National Center for Health Statistics, Hyattsville, MD.

Mosley, W. Henry and Ronald Gray. 1993. "Childhood Precursors of Adult Morbidity and Mortality in Developing Countries: Implications for Health Programs" pp. 69-100 in James N. Gribble and Samuel H. Preston, Editors. The Epidemiological Transition: Policy and Planning Implications for Developing Countries. National Academy Press, Washington, DC.

National Center for Health Statistics. 1975. Vital Statistics of the United States, 1973. Volume II. National Center for Health Statistics, Hyattsville, MD.

National Center for Health Statistics. 1980. Vital Statistics of the United States, 1978. Volume II. National Center for Health Statistics, Hyattsville, MD.

National Center for Health Statistics. 1994. Vital Statistics of the United States, 1990. Volume II. National Center for Health Statistics, Hyattsville, MD

Notkola, V., S. Punsar, M.J. Karvonen, and J. Haapakoski. 1985. "Socioeconomic Conditions in Childhood and Mortality and Morbidity Caused by Coronary Heart Disease in Adulthood in Rural Finland" Social Science and Medicine Vol. 21 No.5: 517-23.

Peck, Maria Nystrom. 1994. "The Importance of Childhood Socio-economic Group for Adult Health" Social Science and Medicine Vol. 39(4): 553-562.

Preston, Samuel H., Suet Lim and S. Philip Morgan. 1992. "African-American Marriage in 1910: Beneath the Surface of Census Data" Demography Vol. 29(1): 1-15.

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

Preston, Samuel H., Irma T. Elo, Ira Rosenwaike, and Mark Hill. 1996. "African American Mortality at Older Ages: Results of a Matching Study," Demography Vol. 33 (2): 193-209.

Preston, Samuel H., Douglas Ewbank, and Mark Hereward. 1994. "Child Mortality Differences by Ethnicity and Race in the United States: 1900-1910" pp. 35-82 in Susan Cotts Watkins, editor. After Ellis Island: Newcomers and Natives in the 1910 Census. Russell Sage Foundation, New York.

Preston, Samuel H. and Paul Taubman. 1994. "Socioeconomic Differences in Adult Mortality and Health Status" pp. 279-318 in Linda Martin and Samuel H. Preston, editors. Demography of Aging. National Research Council, Washington, DC.

Rogot, E, P.D. Sorlie, N.J. Johnson, and C. Schmitt. 1992. A Mortality Study of 1.3 Million Persons by Demographic, Social, and Economic Factors: 1979-1985 Follow-up. National Institutes of Health, Washington, DC.

Ruggles, Steven and Matthew Sobek. 1995. Integrated Public Use Microdata Series (IPUMS).Version 1.0. Social History Research Laboratory, University of Minnesota, Minneapolis.

Schwartz, Joseph E., Howard S. Friedman, Joan S. Tucker, Carol Tomlinson-Keasey, Deborah Wingard, and Michael H. Criqui. 1995. “Sociodemographic and Psychosocial Factors in Childhood as Predictors of Adult Mortality” American Journal of Public Health Vol. 85. No. 9: 1237-45.

Science. 1997. “The Perils of Late-Age Procreation” Science Vol. 276. June 6, 1997: 1503.

U.S. Bureau of the Census. 1910. Report of the Thirteenth Census. 1910, Volumes I & II. Government Printing Office, Washington, DC.

Williams, David R. 1990. “Socioeconomic Differentials in Health: A Review and Redirection” Social Psychology Quarterly Vol. 53(2): 81-99.

Wright, Gavin. 1986. Old South, New South: Revolutions in the Southern Economy since the Civil War. Basic Books, New York.