# *"The Past as Prologue: The Effect of Early Life Circumstances at the Community and Household Levels on Mid-Life and Late-Life Outcomes"*

Joseph Ferrie, NORTHWESTERN UNIVERSITY AND NBER,
Karen Rolf, UNIVERSITY OF NEBRASKA,
and Werner Troesken, UNIVERSITY OF PITTSBURGH AND NBER

December 19, 2006

**Abstract**

The impact of circumstances early in life on later life outcomes in the U.S. can now be explored using data linking individuals from Social Security records back to the manuscript schedules of federal population censuses shortly after their birth. This allows us to assess the effect of individual, household, and community level influences on longevity. We find that these all had substantial affects on age at death, as did characteristics measured at enlistment into the U.S. Army in World War Two.

## Introduction

Health outcomes are among the most intractable forms of inequality in modern society. Though inequalities in income or education or access to housing can be addressed with various forms of redistribution in the very short run, health is the product of a lifetime of influences. Modern research emphasizes the extent to which events very early in life – even events *in utero* – can have a significant impact on health much later in life. We follow a large population of more than 40,000 individuals born between 1900 and 1930 from their appearance in the U.S. federal population censuses until their deaths and their appearance in the Social Security Death Index (SSDI) to assess the impact of early life-circumstances on longevity and cause-specific mortality, as well as how the impact of early life circumstances has evolved across these 30 years of birth cohorts.

Though other studies have examined early-life influences on later-life outcomes, until now none have done so with such detailed information on early-life circumstances at the individual, family, and neighborhood levels, or done so with a study that is prospective in design. Our work does not suffer from the individual's inexact recollection of their own early-life circumstances, nor does it suffer from the attrition bias that occurs in long-run longitudinal studies. The large number of birth cohorts we examine make it possible to assess how early-life influences changed from the beginning of the twentieth century until the early Great Depression. This allows us to more accurately project the longevity and late-life health of cohorts that are now entering their mid-seventies, and thereby evaluate the extent to which health inequality will remain a concern for policymakers over the coming decades.

The impact of circumstances early in life on health outcomes late in life has been a subject of increasing interest since the 1970s. Particularly since the work of Barker and a number of co-authors (see Doblhammer 2003 for a survey of the literature), attention has focused on the environment faced by individuals not just in the years immediately after birth but even *in utero* in shaping their health decades later. A shortcoming of much of the research in the area is the limited range of information available on the early life experiences of individuals whose later life health can be observed: no modern longitudinal datasets span more than 40 years. Information is often collected retrospectively, or small, opportunistic samples that may not be very representative are exploited.

We offer two improvements on previous work: (1) a large sample (2) that is created prospectively. The sample allows us to assess the impact on one particular late life health outcome, longevity, of circumstances both very early in life and at approximately age 25. Recent research by Clay and Troesken (2006) has established a relationship between the environment (crowding and impure water) and health fifteen years later at the city level for the early nineteenth century U.S. We will examine the same time period, as well as outcomes that occur as late as 2005, and focus on circumstances that can be observed not just at the community level but also at the household and individual level.

**The Data**

The dataset was created in two steps: (1) first, 137,224 males who were under age 5 in 1900, 1920, and 1930 and under age 2 in 1910 were drawn from the 1900-30 IPUMS files (Ruggles *et al.*, 2004) – large, nationally-representative samples extracted from the U.S. decennial population censuses; (2) then, these individuals were sought in the Social Security Death Index (which reports the name, date of birth, and date of death for individuals who were in the Social Security system between 1965 and 2005) and in the World War Two enlistment records of the U.S. Army.[1] A total of 18,195 males were linked from one of the pre-1940 censuses to the Social Security records; of these, 1,218 were also linked to the army records. The linkage rate of 13.3% from the IPUMS to the SSDI represents roughly a quarter of those who – based on survivorship by age of the white, native-born male U.S. population – were still alive in 1965 and therefore at risk to enter the Social Security Death Index.[2]

---

[1] Only males are included because, at present, the linkage process requires that the individual's surname did not change between their appearance in the pre-1940 census and their appearance in the Social Security records. Later work will overcome this shortcoming by exploiting information on the names of each decedent's parents in the larger set of Social Security records from which the Death Index is abstracted.

[2] The greatest source of linkage failure, apart from the absence of anyone at all with the correct name, birth month, and birth year in the SSDI, was the inability to differentiate among several individuals who were identical along these three dimensions. At a later stage in the project, additional information will be provided by the Social Security Administration (the full names of both parents, and the detailed place of birth) that will make it possible to resolve these ambiguities. The linked data now also contains an unknown number of "false positives" (individuals from the IPUMS matched to the wrong individual in the SSDI), but these, too, will be eliminated when the additional

The census data, together with information from the National Center for Health Statistics on cause-specific mortality by state and year, provide information on circumstances before and after birth; the enlistment data provide information on subsequent health (measured by Body Mass Index or BMI); finally, the Social Security records make it possible to calculate the age at death for members of the sample.

**Preliminary Analysis**

In Tables 1 through 4, we present the results of our preliminary analysis of the sample. The outcomes examined are age at death as a function of early life circumstances (conditional on having survived to enter the Social Security system and having died by November, 2005 when the Social Security file ends), BMI at enlistment in World War Two, and age at death as a function of both early life circumstances and BMI at enlistment.[3]

The simplest analyses of longevity, with only a small number of control variables, are shown in Table 1. Among individuals who survived to age 70 and died by age 95 (Column 1), longevity was increasing in year of birth, but was sharply lower among those born in the second quarter of the year and somewhat lower among those born in the third quarter. This is broadly consistent with work by Doblhammer (2003) who also found that births in the second quarter of the year resulted in shorted lives than births in the fourth quarter of the year, though we are able to control for a much larger set of early life circumstances. Those born in April, May, and June lived lives nearly 31 weeks shorter

---

data are added. The only substantial bias introduced by the inability to resolve ambiguous links is that the sample is biased toward those with more unusual combinations of name, birth year, and birth month.

[3] Tables 1 through 4 present Ordinary Least Squares (OLS) regressions. All of the substantive findings that follow were also generated when a hazard model was estimated instead.

than those in the excluded category (born October to December). We are investigating the source of this difference. Among individuals who died between ages 50 and 75 and were observed in the 1920 and 1930 censuses, the only significant impact on longevity comes from residence in the South: a reduction of nearly half a year compared to those in the Northeast.

The elimination of the quarter of birth effect in going from Column 1 to Column 2 could result from differences across cohorts in the impact of early life circumstances on mortality, or from differences in the range of age at death: if early circumstances have an impact only on causes of death for which individuals are at heightened risk only after having survived causes of death not so sensitive to early circumstances, the pattern we observed would be expected. Finally, Column 3 reveals the substantial penalty paid in longevity terms by those who were born during the influenza pandemic of 1918-19: an individual born in 1918 survived more than two years fewer than one born in 1915. This is consistent with the long-lingering effects of exposure to the pandemic described in Almond (2006).

Two controls for the mortality environment faced before and after birth are included in Table 2: the state-level mortality rate in the year of birth, and a measure of mortality at the household level. The latter was constructed using the questions for mothers on the number of children ever born and surviving in the 1900 and 1910 censuses. The mortality measure is 1 - (children surviving/children born), so a higher value indicates a more adverse mortality regime at the household level. This could indicate circumstances within the household (either genetic or environmental), or it could reflect the broader community-level health environment. When both the state-level and household-level measures are entered, the latter has a large, negative impact on longevity. The effect of quarter of birth, however, persists. In the post-1910 data, when household-level mortality is unavailable, there is no substantial impact from the state-level mortality measure.

The full set of available control variables is included in Table 3. There are three noteworthy findings: (1) even with a broader set of controls, the effect of quarter of birth remains; (2) the household-level mortality variable remains significant; and (3) the characteristics of the father and mother have little impact on longevity – the only exceptions are occupation of the household head in Columns 5 and 6 (white collar household heads were associated with longer and farmer household heads with shorter life spans after age 50 among those dying by age 85), and the almost universally negative effect of having a native-born father.

Table 4 provides an analysis of both the effect of early circumstances on BMI at the time of enlistment in the army (Columns 1 to 3) and the effect of BMI and early circumstances on longevity (Columns 4 to 6). The army enlistment data provide height and weight for individuals who entered the service between July, 1940 and March, 1943. BMI was lower outside the Northeast, higher in larger cities, and lower among those born during the influenza pandemic. Those born between July and September had higher BMI at enlistment than those born at other times of year. When additional controls are introduced in Column 3, the presence of a household head in a skilled occupation was associated with higher BMI while the sons of native-born fathers had lower BMI.

The effect of BMI on longevity is non-linear: for example, in Column 5, longevity achieves a maximum at a BMI of 24. Figure 1 plots this relationship. The other early life circumstances in Columns 4 through 6 have little impact on longevity, apart from the presence of an English-speaking father (associated with greater longevity) and the presence of a native-born father (associated with shorter longevity). As these regressions control for BMI at around age 25, which captures some of the effect of influences on health up to that age, these latter effects must be generated by early life influences that persist and have an impact over and above the impact that shapes health by young adulthood.

Finally, Table 5 returns to the issue of the season of birth. In Table 1, it was unclear whether the elimination of the season of birth effect in going from the 1900-10 sample to the 1920-30 sample was the result of (1) the elimination of the effect as changes occurred between these two periods, or (2) the fact that the two samples covered different age ranges (which would produce a difference in the season of birth effect if longevity is only effected by season of birth at older ages). In Table 5, the analysis is restricted to individuals linked from the 1920 IPUMS to the SSDI, with separate regressions for two age groups: those 50-69 at death and those 70-84 at death. Whether the local (state) mortality rate in the year of birth is included (Columns 1 and 2) or excluded (3 and 4), two striking differences by age at death emerge: the negative effect on longevity of birth during the influenza pandemic is present for those who died younger but not for those who died older, and the effect of season of birth is present for those who died older but not for those who dies younger (though it is now those born in the first quarter of the year who live the shortest lives among those who died between 70 and 84, as opposed to Table 1 where those born in the second and third quarters who were linked from 1900 and 1910 to the SSDI had the shortest lives conditional on surviving to age 70).[4] This suggests that a least some impact of season of birth remained for those born as late as 1915-19.

---

[4] Three additional effects appear when the data from the 1920 sample are disaggregated by age at death in Table 5: (1) mother's age has a positive but diminishing effect on longevity for those who died between ages 70 and 84 (longevity achieving a maximum at a mother's age of 31 years); (2) higher-order births are associated with shorted longevity among those who die younger (though only when controls for the local mortality rate at birth are included); and (3) those located in more crowded households (measured by the number of persons in the dwelling, as we do not know the floor area of each dwelling) at the time of the 1920 census lived 10 to 12 weeks fewer for each additional co-resident present if they survived to age 70 than otherwise identical individuals.

## Conclusions

For more than twenty years, evidence has mounted that the effects of circumstances very early in life can be observed in outcomes very late in life. The data we have begun to collect allow us to identify these effects in a large, nationally-representative population, and to control for a host of early life circumstances simultaneously. In doing so, we illustrate the extent to which public health measures and improvements in general living conditions can have pay-offs that may be fully realized only decades after the expenditures they entail. We also illustrate the importance of taking account of the full range of circumstances encountered across an individual's lifetime in determining how their current health status and the costs of addressing that status have been generated.

The strongest finding we offer is that season of birth is a strong predictor of longevity, though perhaps more so for those who survive to older ages (over 70) than for those who die younger (under 70), and BMI at age 25 is a strong predictor of longevity even after accounting for numerous early life circumstances. Influences at the household and community levels (mortality, crowding, region of residence) have an impact on longevity as well. With additional data in hand (both a larger linked population and detailed information on cause of death), we will be able to describe the mechanisms through which these effects operate. Finally, additional data on earnings, disability, and interactions with the Medicare system will be linked to the individuals whom we have followed from very early in life until death, making it possible to assess the effect of early circumstances on a broader range of later life outcomes.

**References**


Almond, Douglas (2006). "Is the 1918 Influenza Pandemic Over? Long-Term Effects of *In Utero* Influenza Exposure in the Post-1940 U.S. Population." *Journal of Political Economy* 114: 672-712.

Clay, Karen, and Werner Troesken (2006). "Deprivation and Disease in Early Twentieth Century America." NBER Working Paper No. 12111 (March).

Doblhammer, Gabriele (2003). "The Late Life Legacy of Very Early Life." Max Planck Institute for Demographic Research Working Paper WP2003-030 (September).

Ruggles, Steven, Matthew Sobek, Trent Alexander, Catherine A. Fitch, Ronald Goeken, Patricia Kelly Hall, Miriam King, and Chad Ronnander (2004). Integrated Public Use Microdata Series: Version 3.0 [Machine-readable database]. Minneapolis, MN: Minnesota Population Center [producer and distributor].

Table 1. OLS Regression on Age at Death: Basic Specification

| | (1) 1900-10, Age 70-95 | (2) 1920-30, Age 50-75 | (3) 1920, Age 50-85 |
|---|---|---|---|
| Midwest | -0.099 | 0.272 | 0.367 |
| | (0.49) | (1.09) | (0.57) |
| South | -0.357 | -0.453* | -0.580 |
| | (1.63) | (1.66) | (0.79) |
| West | -0.191 | 0.214 | 1.156 |
| | (0.54) | (0.58) | (1.14) |
| City Pop. 100,000+ | -0.314 | -0.310 | -0.656 |
| | (1.41) | (1.23) | (1.00) |
| City Pop. 10,000-99,999 | -0.168 | 0.283 | -0.001 |
| | (0.68) | (1.04) | (0.00) |
| Year of Birth | 0.076*** | -0.005 | |
| | (5.22) | (0.22) | |
| Born Jan-Mar | -0.245 | 0.188 | 0.670 |
| | (1.14) | (0.67) | (0.87) |
| Born Apr-Jun | -0.594*** | -0.012 | 0.778 |
| | (2.71) | (0.05) | (1.03) |
| Born Jul-Sep | -0.372* | -0.168 | 0.628 |
| | (1.74) | (0.61) | (0.87) |
| Family Size | 0.035 | -0.043 | -0.128 |
| | (0.94) | (0.86) | (0.93) |
| Born 1916 | | | -1.039 |
| | | | (0.99) |
| Born 1917 | | | -1.318 |
| | | | (1.33) |
| Born 1918 | | | -2.153** |
| | | | (2.28) |
| Born 1919 | | | -1.674* |
| | | | (1.87) |
| Constant | 80.486*** | 65.640*** | 73.242*** |
| | (242.44) | (96.94) | (57.30) |
| Observations | 7369 | 4738 | 1218 |
| Adjusted R-squared | 0.00 | 0.00 | -0.00 |

Absolute value of t statistics in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 2. OLS Regression on Age at Death: With Mortality Measures

| | (1) 1900-10, Age 70-95 | (2) 1900-10, Age 70-95 | (3) 1920-30, Age 50-75 | (4) 1920, Age 50-85 |
|---|---|---|---|---|
| Midwest | -0.090 | -0.933 | 0.153 | -0.681* |
| | (0.44) | (1.61) | (0.52) | (1.66) |
| South | -0.329 | -2.091* | -0.400 | -1.324*** |
| | (1.48) | (1.86) | (1.19) | (3.06) |
| West | -0.180 | -1.345 | 0.224 | 0.143 |
| | (0.50) | (1.50) | (0.53) | (0.26) |
| City Pop. 100,000+ | -0.313 | -0.519 | -0.285 | -0.081 |
| | (1.40) | (1.14) | (1.04) | (0.25) |
| City Pop. 10,000-99,999 | -0.154 | 0.537 | 0.230 | -0.087 |
| | (0.61) | (0.99) | (0.76) | (0.24) |
| Year of Birth | 0.076*** | 0.092 | -0.016 | |
| | (5.14) | (1.14) | (0.60) | |
| Born Jan-Mar | -0.258 | 0.984* | 0.355 | 0.218 |
| | (1.19) | (1.83) | (1.11) | (0.56) |
| Born Apr-Jun | -0.623*** | -1.540*** | 0.358 | 0.458 |
| | (2.81) | (2.72) | (1.14) | (1.21) |
| Born Jul-Sep | -0.398* | 0.213 | -0.184 | 0.195 |
| | (1.85) | (0.37) | (0.59) | (0.51) |
| Family Size | 0.036 | 0.003 | -0.079 | -0.122* |
| | (0.94) | (0.03) | (1.40) | (1.82) |
| Household Mortality Rate | -0.713 | -5.019*** | | |
| | (1.06) | (2.77) | | |
| State Mort Rate at Birth | | -0.160 | -0.041 | -0.143 |
| | | (1.20) | (0.83) | (1.51) |
| Born 1916 | | | | -0.102 |
| | | | | (0.22) |
| Born 1917 | | | | 0.110 |
| | | | | (0.24) |
| Born 1918 | | | | 0.123 |
| | | | | (0.20) |
| Born 1919 | | | | -0.414 |
| | | | | (1.00) |
| Constant | 80.546*** | 83.056*** | 66.589*** | 74.491*** |
| | (236.59) | (29.71) | (53.99) | (49.63) |
| Observations | 7209 | 1100 | 3661 | 4789 |
| Adjusted R-squared | 0.00 | 0.03 | 0.00 | 0.00 |

Absolute value of t statistics in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 3. OLS Regression on Age at Death: With Full Set of Controls

| | (1) 1900-10, Age 70-95 | (2) 1900-10, Age 70-95 | (3) 1920-30, Age 50-75 | (4) 1920-30, Age 50-75 | (5) 1920, Age 50-85 | (6) 1920, Age 50-85 |
|---|---|---|---|---|---|---|
| Midwest | -0.284 | -1.096* | 0.403 | 0.453* | 0.088 | -0.318 |
| | (1.35) | (1.77) | (1.32) | (1.73) | (0.29) | (0.74) |
| South | -0.482** | -2.637** | -0.007 | -0.129 | -0.414 | -0.836* |
| | (1.97) | (2.16) | (0.02) | (0.43) | (1.16) | (1.75) |
| West | -0.335 | -1.509 | 0.600 | 0.536 | 0.678 | 0.533 |
| | (0.90) | (1.59) | (1.38) | (1.40) | (1.55) | (0.95) |
| City Pop. 100,000+ | -0.160 | -0.335 | -0.392 | -0.381 | -0.347 | -0.187 |
| | (0.63) | (0.60) | (1.26) | (1.33) | (1.03) | (0.50) |
| City Pop. 10,000-99,999 | -0.036 | 0.625 | 0.203 | 0.285 | -0.130 | -0.186 |
| | (0.13) | (1.05) | (0.61) | (0.95) | (0.37) | (0.46) |
| Year of Birth | 0.094* | 0.301 | -0.002 | 0.006 | | |
| | (1.83) | (1.27) | (0.09) | (0.28) | | |
| Born Jan-Mar | -0.269 | 0.589 | 0.387 | 0.212 | 0.127 | 0.267 |
| | (1.20) | (0.98) | (1.18) | (0.74) | (0.37) | (0.67) |
| Born Apr-Jun | -0.677*** | -1.698*** | 0.445 | 0.038 | 0.252 | 0.492 |
| | (3.00) | (2.92) | (1.38) | (0.13) | (0.76) | (1.27) |
| Born Jul-Sep | -0.414* | 0.038 | -0.120 | -0.120 | 0.194 | 0.259 |
| | (1.89) | (0.06) | (0.37) | (0.43) | (0.59) | (0.67) |
| Family Size | 0.078 | -1.247 | 0.029 | 0.050 | 0.003 | -0.116 |
| | (0.47) | (0.94) | (0.14) | (0.28) | (0.01) | (0.40) |
| White Collar | 0.343 | -0.056 | -0.077 | -0.143 | 0.613 | 0.795* |
| | (1.27) | (0.09) | (0.21) | (0.44) | (1.62) | (1.85) |
| Farmer | 0.231 | -0.404 | 0.062 | -0.030 | -0.679* | -0.314 |
| | (0.90) | (0.57) | (0.16) | (0.09) | (1.78) | (0.69) |
| Skilled | -0.142 | -1.093** | -0.244 | -0.268 | 0.083 | 0.333 |
| | (0.59) | (1.98) | (0.78) | (0.95) | (0.25) | (0.88) |
| Mother's Age | 0.018 | -0.004 | | | | |
| | (1.20) | (0.11) | | | | |
| Household Mortality Rate | -1.305* | -5.894*** | | | | |
| | (1.81) | (3.03) | | | | |
| Mother Literate | -0.106 | -0.165 | -0.038 | -0.034 | -0.017 | -0.064 |
| | (1.26) | (0.56) | (0.23) | (0.23) | (0.11) | (0.34) |
| Father Literate | 0.669* | 0.653 | 0.416 | 0.365 | -0.528 | -0.539 |
| | (1.90) | (0.45) | (0.69) | (0.71) | (0.90) | (0.77) |
| Mother Speaks English | -0.173 | -0.411 | 0.986 | 1.344 | -0.528 | 0.085 |
| | (0.44) | (0.67) | (0.93) | (1.35) | (0.56) | (0.08) |
| Father Speaks English | 0.158 | 0.024 | 1.444 | -0.438 | 0.788 | 1.752 |
| | (0.39) | (0.04) | (1.07) | (0.36) | (0.66) | (1.23) |
| Mother Native Born | 0.333 | 1.566 | -0.641 | -0.551 | -0.589 | -0.602 |
| | (1.05) | (0.94) | (1.50) | (1.39) | (1.31) | (1.23) |

Table 3. OLS Regression on Age at Death: With Full Set of Controls (continued)

| | (1) 1900-10, Age 70-95 | (2) 1900-10, Age 70-95 | (3) 1920-30, Age 50-75 | (4) 1920-30, Age 50-75 | (5) 1920, Age 50-85 | (6) 1920, Age 50-85 |
|---|---|---|---|---|---|---|
| Father Native Born | -0.595** | -1.938 | -0.685* | -0.531 | -0.904** | -1.018** |
| | (2.01) | (1.15) | (1.70) | (1.43) | (2.16) | (2.23) |
| Black | -0.004 | -0.025 | -0.553 | -0.334 | -0.281 | -0.119 |
| | (0.76) | (1.26) | (1.09) | (0.96) | (0.72) | (0.20) |
| Birth Order | -0.103 | 1.246 | -0.105 | -0.088 | -0.088 | -0.023 |
| | (0.61) | (0.95) | (0.49) | (0.48) | (0.35) | (0.08) |
| Dwelling Size | 0.364 | 9.471 | -0.016 | -0.109 | -0.286 | -0.254 |
| | (1.25) | (0.87) | (0.09) | (0.65) | (1.54) | (1.26) |
| Home Owned | 0.057 | 0.588 | -0.083 | -0.118 | 0.298 | 0.429 |
| | (0.25) | (1.05) | (0.35) | (0.57) | (1.24) | (1.52) |
| Mortgage Paid off | 0.091 | -0.049 | | | | |
| | (0.37) | (0.07) | | | | |
| State Mort Rate at Birth | | -0.151 | -0.019 | | | -0.125 |
| | | (1.06) | (0.36) | | | (1.28) |
| Born 1916 | | | | | -0.474 | -0.081 |
| | | | | | (1.19) | (0.17) |
| Born 1917 | | | | | -0.239 | 0.181 |
| | | | | | (0.59) | (0.37) |
| Born 1918 | | | | | -0.554 | 0.160 |
| | | | | | (1.35) | (0.24) |
| Born 1919 | | | | | -0.418 | -0.305 |
| | | | | | (1.01) | (0.63) |
| Constant | 79.932*** | 85.764*** | 64.111*** | 64.985*** | 73.217*** | 73.657*** |
| | (90.94) | (18.09) | (34.20) | (51.00) | (55.63) | (34.65) |
| Observations | 6964 | 1049 | 3473 | 4498 | 6147 | 4551 |
| Adjusted R-squared | 0.01 | 0.03 | 0.00 | 0.00 | 0.00 | 0.00 |

Absolute value of t statistics in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 4. OLS Regression on BMI & Age at Death, 1920, Age 50-85

| | (1) BMI in Army | (2) BMI in Army | (3) BMI in Army | (4) Age at Death | (5) Age at Death | (6) Age at Death |
|---|---|---|---|---|---|---|
| Midwest | -0.379* | -0.240 | -0.244 | 0.381 | 0.273 | -1.408 |
| | (1.81) | (1.10) | (0.79) | (0.59) | (0.41) | (1.50) |
| South | -0.525** | -0.414 | -0.580 | -0.520 | -0.749 | -1.698 |
| | (2.20) | (1.58) | (1.63) | (0.71) | (0.93) | (1.56) |
| West | -0.816** | -0.593* | -0.531 | 1.057 | 1.178 | -0.685 |
| | (2.47) | (1.74) | (1.20) | (1.04) | (1.13) | (0.51) |
| City Pop. 100,000+ | 0.471** | 0.133 | 0.202 | -0.578 | -0.775 | -0.496 |
| | (2.21) | (0.54) | (0.75) | (0.88) | (1.04) | (0.60) |
| City Pop. 10,000-99,999 | -0.056 | -0.342 | -0.358 | 0.107 | 0.017 | 0.157 |
| | (0.23) | (1.29) | (1.22) | (0.14) | (0.02) | (0.17) |
| Born 1916 | -0.099 | 0.147 | 0.177 | -1.194 | -0.723 | -0.351 |
| | (0.29) | (0.41) | (0.41) | (1.14) | (0.66) | (0.27) |
| Born 1917 | -0.402 | -0.240 | -0.213 | -1.516 | -1.271 | -0.557 |
| | (1.25) | (0.70) | (0.52) | (1.53) | (1.21) | (0.45) |
| Born 1918 | -0.467 | -0.237 | 0.016 | -2.232** | -1.689 | 0.981 |
| | (1.52) | (0.69) | (0.03) | (2.37) | (1.60) | (0.61) |
| Born 1919 | -0.582** | -0.358 | -0.416 | -1.772** | -1.265 | -1.071 |
| | (1.99) | (1.04) | (1.06) | (1.98) | (1.21) | (0.90) |
| Born Jan-Mar | 0.312 | 0.337 | 0.226 | 0.685 | 0.505 | 0.750 |
| | (1.24) | (1.30) | (0.77) | (0.89) | (0.64) | (0.83) |
| Born Apr-Jun | 0.197 | 0.132 | 0.083 | 0.935 | 0.719 | 0.826 |
| | (0.80) | (0.52) | (0.29) | (1.24) | (0.93) | (0.93) |
| Born Jul-Sep | 0.441* | 0.403* | 0.369 | 0.630 | 0.601 | 1.066 |
| | (1.87) | (1.67) | (1.33) | (0.87) | (0.82) | (1.26) |
| Family Size | -0.028 | 0.044 | 0.116 | -0.135 | 0.458 | 0.273 |
| | (0.63) | (0.21) | (0.48) | (0.99) | (0.73) | (0.37) |
| White Collar | | 0.107 | 0.357 | | 0.531 | 1.090 |
| | | (0.39) | (1.16) | | (0.64) | (1.16) |
| Farmer | | -0.221 | -0.422 | | -0.040 | -0.093 |
| | | (0.76) | (1.21) | | (0.04) | (0.09) |
| Skilled | | 0.381 | 0.665** | | 0.205 | 0.914 |
| | | (1.51) | (2.40) | | (0.27) | (1.08) |
| Mother Literate | | -0.020 | -0.121 | | 0.160 | 0.076 |
| | | (0.14) | (0.74) | | (0.37) | (0.15) |
| Father Literate | | -0.121 | -0.366 | | -1.511 | -1.321 |
| | | (0.24) | (0.65) | | (0.98) | (0.77) |
| Mother Speaks English | | -0.228 | -0.317 | | -0.198 | 0.651 |
| | | (0.28) | (0.35) | | (0.08) | (0.24) |
| Father Speaks English | | -0.611 | -0.113 | | 8.233** | 7.158** |
| | | (0.57) | (0.10) | | (2.54) | (2.10) |

Table 4. OLS Regression on BMI & Age at Death, 1920, Age 50-85 (continued)

| | (1)<br>BMI in Army | (2)<br>BMI in Army | (3)<br>BMI in Army | (4)<br>Age at Death | (5)<br>Age at Death | (6)<br>Age at Death |
|---|---|---|---|---|---|---|
| Mother Native Born | | 0.054 | 0.185 | | -0.406 | 0.290 |
| | | (0.17) | (0.54) | | (0.41) | (0.28) |
| Father Native Born | | -0.838*** | -0.989*** | | -2.118** | -2.521** |
| | | (2.71) | (2.98) | | (2.24) | (2.47) |
| Black | | 1.329** | 1.622** | | 1.043 | 2.077 |
| | | (2.46) | (2.48) | | (0.63) | (1.04) |
| Birth Order | | -0.107 | -0.143 | | -0.651 | -0.511 |
| | | (0.51) | (0.58) | | (1.01) | (0.67) |
| Dwelling Size | | 0.081 | 0.057 | | -0.318 | -0.335 |
| | | (0.59) | (0.41) | | (0.77) | (0.80) |
| Home Owned | | 0.088 | 0.279 | | 0.531 | 0.739 |
| | | (0.47) | (1.31) | | (0.94) | (1.14) |
| BMI at Enlistment in Army | | | | 2.391*** | 2.420*** | 2.266** |
| | | | | (3.04) | (3.05) | (2.57) |
| $(\text{BMI at Enlistment in Army})^2$ | | | | -0.051*** | -0.053*** | -0.050*** |
| | | | | (3.23) | (3.27) | (2.81) |
| State Mort Rate at Birth | | | -0.038 | | | -0.534** |
| | | | (0.53) | | | (2.43) |
| Constant | 23.217*** | 24.474*** | 24.749*** | 45.948*** | 39.014*** | 48.423*** |
| | (55.80) | (22.78) | (15.95) | (4.73) | (3.75) | (4.02) |
| Observations | 1218 | 1170 | 933 | 1218 | 1170 | 933 |
| Adjusted R-squared | 0.01 | 0.02 | 0.03 | 0.01 | 0.02 | 0.02 |

Absolute value of t statistics in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 5. OLS Regression on Age at Death for 1915-1919 Births.

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | Excluding State-Level Mortality | | Including State-Level Mortality | |
| | Age 50-69 at Death | Age 70-84 at Death | Age 50-69 at Death | Age 70-84 at Death |
| Midwest | 0.018 | -0.460** | -0.138 | -0.591** |
| | (0.07) | (2.51) | (0.36) | (2.27) |
| South | 0.006 | -0.245 | 0.500 | -0.457 |
| | (0.02) | (1.13) | (1.20) | (1.52) |
| West | 0.119 | -0.039 | 0.376 | -0.051 |
| | (0.30) | (0.15) | (0.72) | (0.15) |
| City Pop. 100,000+ | 0.092 | 0.023 | 0.205 | 0.225 |
| | (0.30) | (0.11) | (0.62) | (1.01) |
| City Pop. 10,000-99,999 | 0.551* | -0.326 | 0.762** | -0.353 |
| | (1.70) | (1.52) | (2.09) | (1.44) |
| Born 1918 | -0.287 | 0.027 | -0.099 | 0.162 |
| | (1.07) | (0.14) | (0.20) | (0.51) |
| Born 1919 | -0.782*** | 0.181 | -0.726** | 0.240 |
| | (3.02) | (1.04) | (2.40) | (1.18) |
| Born Jan-Mar | 0.259 | -0.374* | 0.263 | -0.542** |
| | (0.83) | (1.76) | (0.73) | (2.20) |
| Born Apr-Jun | 0.162 | 0.023 | 0.184 | -0.280 |
| | (0.53) | (0.11) | (0.52) | (1.16) |
| Born Jul-Sep | 0.149 | -0.075 | 0.189 | -0.154 |
| | (0.49) | (0.36) | (0.54) | (0.64) |

Table 5. OLS Regression on Age at Death for 1915-1919 Births (continued).

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | Excluding State-Level Mortality | | Including State-Level Mortality | |
| | Age 50-69 at Death | Age 70-84 at Death | Age 50-69 at Death | Age 70-84 at Death |
| White Collar | -0.140 | 0.018 | -0.139 | -0.040 |
| | (0.40) | (0.08) | (0.35) | (0.15) |
| Skilled | 0.171 | -0.062 | 0.292 | 0.009 |
| | (0.57) | (0.30) | (0.87) | (0.04) |
| Farmer | 0.295 | -0.006 | 0.270 | 0.241 |
| | (0.83) | (0.03) | (0.63) | (0.87) |
| Mother's Age | -0.007 | 0.145* | 0.040 | 0.171 |
| | (0.09) | (1.66) | (0.48) | (1.64) |
| $(\text{Mother's Age})^2$ | 0.001 | -0.002 | -0.000 | -0.003 |
| | (0.46) | (1.53) | (0.11) | (1.54) |
| Birth Order | -0.073 | -0.031 | -0.136* | -0.029 |
| | (1.18) | (0.70) | (1.92) | (0.56) |
| Home Owned | -0.289 | 0.205 | -0.158 | 0.230 |
| | (1.08) | (1.15) | (0.51) | (1.13) |
| Mortgage Paid off | 0.151 | -0.186 | 0.366 | 0.021 |
| | (0.45) | (0.84) | (0.92) | (0.08) |
| Dwelling Size | -0.160 | -0.233** | -0.049 | -0.204 |
| | (1.00) | (1.96) | (0.28) | (1.64) |
| State Mort Rate at Birth | | | -0.007 | -0.008 |
| | | | (0.08) | (0.13) |
| Constant | 62.198*** | 76.017*** | 61.479*** | 75.763*** |
| | (50.02) | (58.69) | (31.71) | (42.40) |
| Observations | 2497 | 3534 | 1841 | 2673 |
| Adjusted R-squared | 0.00 | 0.00 | 0.00 | 0.00 |

Absolute value of t statistics in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

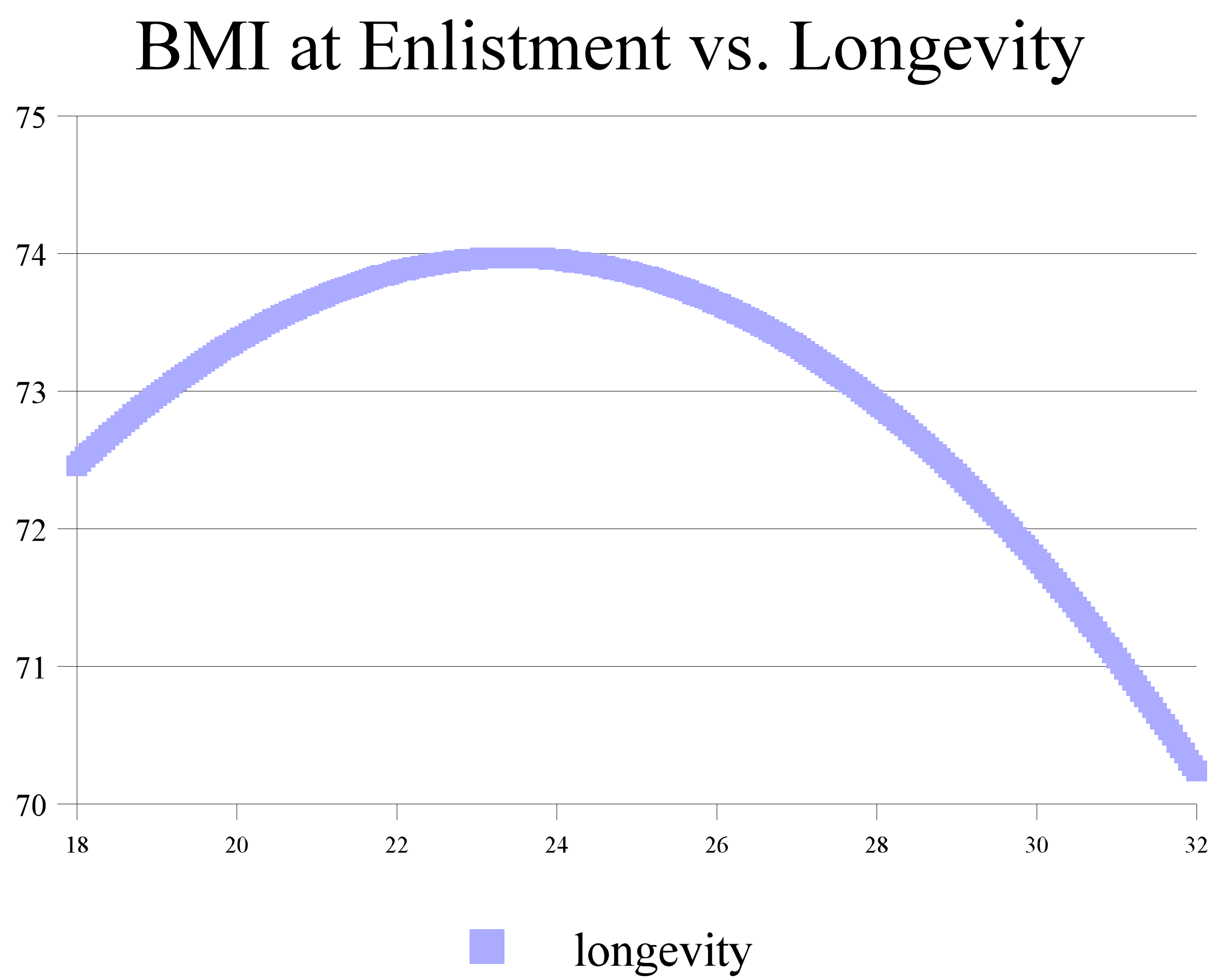


**Figure 1**. Relationship Between BMI at Enlistment and Longevity, Males Initially Observed in the 1920 U.S. Census (When Under Age 5) Who Survived to Age 50 and Died by Age 85. Source: Calculated using coefficients from Table 4, Column 4.