**Sources of Increasing Differential Mortality among the Aged by Socioeconomic Status**

Barry P. Bosworth, Gary Burtless, and Kan Zhang
The Brookings Institution

One of the most basic indicators of well-being is life expectancy. A large empirical literature has established that there are significant differences in life expectancy between people with high and low socioeconomic status (SES) as measured by indicators such as income and educational attainment. Recent studies also show that the differential has widened in the United States and most high-income European countries since the mid-20th century, reversing a long trend toward greater equality. The trend is visible even after excluding the effects of changes in childhood mortality. Some studies suggest that all of the recent increase in life expectancy is concentrated among those with above-average SES, and that life expectancy is constant or even declining for those with lower status. This paper attempts to identify some of the sources of the increasing disparity using data on causes of death in the Health and Retirement Study (HRS).

Evidence of a widening of the difference in mortality by SES has been found in numerous large scale empirical studies. Meara, Richards and Cutler (2008) and Olshansky and others (2012) analyze death certificate data, using educational attainment as the measure of SES, to show a sharp rise in inequality. Waldron (2007) uses Social Security records containing information on career earnings and age at death to establish a similar pattern for men. Singh and Siahpush (2006) provide additional evidence on changes in differential mortality using county-level information from the decennial censuses. They construct county-level indexes of SES linked to death records by location. Increased life span inequality is also evident in other high-income countries (Mackenbach 2003).

Efforts to explain the rising mortality differential by SES have attracted increased attention in recent years. One of the largest studies is by Steenland and others (2002) who

-1-

reported a comparison of cause-specific death rates by educational level in two surveys sponsored by the American Cancer Society. The surveys, covering deaths in 1959-72 and 1982-96, showed an inverse correlation between schooling and age-adjusted death rates among those 45 and older for all causes of death except breast cancer. More important, the authors found increasing mortality differences by educational attainment across the two surveys, and the growing differences were particularly marked for coronary heart disease among men. However, results for a variety of cancers were more mixed. Cutler and others (2010) examined the role of behavioral risk factors in accounting for the increased education differential. They found that smoking, obesity, high cholesterol and hypertension played only a small role. Their results were particularly striking given that smoking and obesity are the two leading behavioral causes of death and are, of course, linked to SES. In sum, the source of the increased mortality differential remains uncertain and controversial.

Some earlier studies have made use of the HRS. Hoffman (2005) shows that SES differences in overall mortality risk persist as people age, but they come to be dominated by health conditions at older ages when SES plays a decreasing independent role. A study by Feinglass and others (2007) used data from the HRS between 1992 and 2002 to explore the role of income, wealth, and education as alternative measures of SES in estimating the relationship to mortality. They found that after inclusion of initial health status, only income remained significant as a determinant of the mortality rate. However, the use of concurrent income and wealth measures may have contaminated their results through reverse causation. Very poor health both reduces current income and leads to early death.

We use the longitudinal data from the HRS to estimate the relationship between SES and the different causes of death as they relate to older persons. From the previous work we know that the pattern of increasing mortality differentials among older Americans can be seen in the HRS sample. The HRS interviews and associated data files contain a great deal of information on SES, work history, health status, causes of death, and other items that are crucial to identifying the links between SES, on the one hand, and causes of poor health and death, on the other.

Our paper relies primarily on data from the HRS. The files include data on more than 30,000 respondents over the age of 50 who entered the sample between 1992 and 2010. Recorded deaths totaled about 12,000 by 2012. The circumstances of those deaths are recorded in an exit interview with a relative or other knowledgeable person. Part of the interview covers

the cause of death grouped into 11 broad categories. Roughly a quarter of the deaths are attributable to various cancers, and more than one-third can be traced to cardiovascular problems. About two-thirds of the respondents have records that are linked to their Social Security earnings files. The link enables us to construct measures of career earnings to use in combination with education as indicators of SES. Thus, our measures of SES reflect status in a period that is usually well before the age at death. As a result, we can tabulate deaths by cause, and compare the classification with data on the individual's SES as measured by educational attainment and/or career earnings.

Second, we analyze changes in the distribution of deaths by cause to determine whether excess deaths as a result of specific causes are linked to our measure(s) of SES. The basic approach involves use of logistic regression to estimate the probability of death by different causes as a function of age, birth cohort, our measure of SES, and an interaction of birth cohort and SES. The interaction term allows us to estimate the change in differential mortality. We also include a number of behavioral indicators, such as smoking and alcohol consumption, as well as indicators of parents' lifespans to determine if inclusion of these variables reduces the magnitude and significance of the interaction term. The study thus provides new information on the growing inequality of mortality risk by SES and may point to some potential causes of the trend. The HRS provides information allowing us to use both lifetime income and educational attainment as determinants of SES. These are crucial advantages compared with other data sets containing information on individual mortality and its cause.

The remainder of the paper is organized as follows. In the next section we provide a detailed description of the SES and mortality data available in our analysis file. We replicate earlier findings showing that there has been an increase in the mortality gradient by SES among birth cohorts represented in the HRS sample. The next section assesses the impact of SES on the specific causes of death reflected in the HRS mortality records. We use two statistical methods to evaluate this relationship, but both show a similar pattern. Our findings can be summarized briefly. We find an increase in mortality differential by SES, whether sample members' SES is measured by average earnings or educational attainment, but the specific causes of death contributing to widening mortality differences are less clear. According to our estimates, differentials have widened for many specific causes of death.

## I.      Data and description of overall mortality trends

The basic structure of the data set is summarized in Table 1. We begin with the first five birth cohort samples of the HRS, which contain a total of 31,875 respondents.[2] The cohorts span birth years from 1890 to 1965.[3] The third line shows the date at which each cohort was introduced in the study. Respondents were resurveyed every two years through 2010. Each cohort sample is selected from a target population that includes the noninstitutional population of the contiguous United States. While the institutionalized population was excluded from the initial survey frame, selected individuals are followed in later survey waves if they entered an institution. The AHEAD and CODA samples differ from the other target populations in being considerably older when they were brought into the study. As a result, we do not observe deaths for these two older age cohorts until they are in their 70s. In contrast, the HRS and younger cohorts were in their 50s when they entered the study. About two-thirds of the respondents granted permission to link their Social Security earnings and benefit records to their interview responses (lines 4 and 5). Our analysis file includes Social Security administrative data that was last updated in 2008.

The HRS records include a total of 11,937 deaths, but we exclude mortality data from 2011 and 2012 because the information for those two years is still incomplete. Thus, our base sample includes 11,689 recorded deaths over the period from 1992 to 2010. The information on an individual's death and its cause is based on an exit interview with a relative or proxy respondent. Our analysis file excludes information about surviving sample members for years after they ceased participating in the study. We constructed a data set to measure annual mortality rates under the following assumptions. Persons with a death record are included for every year from their entry into the study up through their year of death. All other respondents are included from their year of entry up to 2010. The resulting analysis sample contains a total of 166,843 male person-years and 232,177 female person-years of observations

Research on disparities in mortality by SES has relied on two primary measures of SES: educational attainment and income. Educational attainment is recorded in nearly all socio-

---

[2] We exclude respondents in the Mid-Baby Boom cohort (born between years 1954 and 1959) because they don't enter the survey until 2010, which is our last observation year.

[3] The preponderance of the data is drawn from Version M of the RAND HRS, which incorporates the final release of 2010. Actual birth years span a wider range than the target populations because the sample also includes spouses of enrolled sample members in the target age groups.

economic surveys and surprisingly is included in most death records. However, most household surveys include only a measure of current income, and such a measure may be subject to large transient influences. Thus, most studies of the relationship between mortality and SES have relied on educational attainment to distinguish between people with high and low social and economic status. For some purposes, of course, it may be important to measure status using an income-based concept. The link between HRS interview and mortality data, on the one hand, and Social Security earnings records, on the other, makes it possible to construct a measure of average or career earnings that is more representative of lifetime resources. We use a procedure similar to Waldron (2007) and construct career earnings as an average of nonzero annual earnings for ages 41 to 50.[4]

However, the limited work histories of older married women up through the mid-1990s makes career earnings a problematic indicator of their socioeconomic status. Therefore, we rely on a construct of combined husband-and-wife earnings as the primary income-based measure of married individuals' SES. We define household earnings for individuals with a spouse as the sum of the two spouses' combined career earnings divided by the square root of two.[5] In our analysis, we differentiate between high and low SES groups based on whether individuals' career earnings fall above or below the median. Earnings are ranked separately for men and women and relative only to persons of the same gender who are born in the same birth year or in the two adjacent birth years.[6]

---

[4] Nominal earnings in each year were divided by the SSA average wage index with a base of 2005, eliminating the secular growth in wages. In addition, a large proportion of reported wages prior to 1980 were truncated at the taxable wage maximum. However, the earnings records contained information on the quarter in which the worker reached the taxable ceiling. We used that information to impute a wage for workers at the taxable ceiling on the assumption that earnings were constant throughout the year. In our sample, 50 percent of HRS respondents who agreed to provide access to Social Security records had positive earnings in all 10 years, and 14 percent reported zero earnings in all of the ten-year period.

[5] Among married respondents who granted permission to link their Social Security earnings records to the HRS data files, 89% were matched to spouses who *also* granted access to their earnings records. In such cases, where one spouse had no earnings between ages 41 and 50, we defined household earnings to simply be the other spouse's earnings divided by square root of two. However, for the remaining 11% of matched husband-wife households where one spouse did not grant permission to access their earnings records, we were unable to compute household career earnings.

[6] Because of the small number of observations in the earliest and latest birth years, we combined birth years 1890-1909, and birth years after 1956 into two larger groups to compute earnings ranks.

Some simple measures of relative mortality—the annual mortality rate for a group with a specific characteristic divided by the overall mortality rate for the wider population—are displayed in Table 2. The sample is divided by gender and age, with results shown separately for sample members between ages 50-74 and those age 75 and older. Notice that race is a strong distinguishing characteristic only for blacks below age 75. Mortality differs substantially by educational attainment. Mortality rates of respondents who are college-educated are half those of respondents who have not completed high school. Differences by quartiles of household income are equally marked. Note, however, that there is a substantial narrowing of mortality differences by education and income after age 74. This is consistent with previous research showing that SES is less important for predicting the probability of death for older individuals: There is some compression of the mortality risk for the older age group.[7]

When we restrict our sample to respondents who agreed to provide Social Security earnings records there is a sizeable reduction in sample size, especially for the older population enrolled as part of the AHEAD sample. Not surprisingly, the older population is a major source of information about reported deaths. To increase the size of the sample available for analysis, we also generated predictions of career earnings, separately for women and men and using a log-linear regression, based on independent variables reported in the HRS surveys—race,, educational attainment, disability, birth year, and marital status. Coefficient estimates of our career earnings regression are reported in appendix Table A1. We used the predicted values from the log-linear regressions to construct a measure of predicted household earnings and earnings ranks using the same procedures outlined above for actual earnings.[8] As shown at the bottom of table 2, predicted household earnings yield differences in mortality rates for the full sample that are very similar to those for the subsample with actual earnings. However it is important to note that the predicted earnings variable is largely a transformation of educational attainment. In our more formal mortality analysis, we use actual income, predicted income and education as alternative measures of SES.

*Mortality in the HRS sample.* Nearly all available mortality statistics reflect deaths in the total U.S. population, including deaths among people in institutions as well as in the

---

[7] See Brown and others (2012).

[8] In calculating income ranks, we combined birth years 1890-1904 and birth years after 1961. This method of calculation was necessary because of the very small number of respondents born in these years.

noninstitutionalized population. Because the HRS samples were originally drawn from the noninstitutionalized population they may have a different mortality pattern than would be found in a random sample drawn from the entire U.S. population. To investigate this possibility we estimated a parsimonious discrete-time logistic model to summarize observed mortality patterns in our HRS sample:

$$(1)\ log\ (h_{it}\ /\ (1 -\ h_{it}\ )) = \alpha_0 + \beta_1\ Age_{it} + \beta_2\ Birth\ Year_i + \beta_3\ FYr_{it}\ ,$$

where $h_{it}\ = Pr(\ Y_{it} = 1\ /\ Y_{it-1} = 0)$ is the hazard that person $i$ will die in year $t$; and

$FYr_{it}\ =\ 1$ if year $t$ is the first year person $i$ is included in the HRS sample.
$=\ 0$ otherwise.

(We include $FYr_{it}$ in the specification to reflect the fact that respondents are exposed to the risk of dying for less than a full 12 months in the first calendar year of their enrollment.) Equation 1 was estimated separately for women and men and for three birth cohorts in our sample: persons born between 1921 and 1930, between 1931 and 1940, and between 1941 and 1950.

Figure 1 shows the resulting estimates of age-specific mortality rates for HRS sample members who were born in three years, 1925, 1935, and 1945. Results for men are shown in the top panel; those for women are displayed at the bottom. The HRS estimates are indicated with solid dark lines. The lines for the three cohorts do not cover the same span of ages, because members of each birth cohort were observed during a different span of ages between 1992 and 2010, the years covered by our analysis. For both sexes and each birth cohort we see a nonlinear rise in estimated mortality rates at successively higher ages. In the case of men, the lines of the different birth cohorts do not intersect one another, even when there is an overlap in the ages observed in our analysis panel. The younger birth cohorts have lower mortality rates at given ages than the older birth cohorts. Death rates were falling in successive birth cohorts compared with earlier ones. In the case of women the mortality improvement is less clear. The oldest cohort, born in 1925 has lower mortality rates between ages 67 and 74 than the middle cohort, which was born in 1935. On the other hand, women in the youngest cohort, born in 1945, have lower death rates between the ages of 57 and 65 than women in the middle cohort, born in 1935.

The mortality experience of HRS respondents can be compared with that in the broader population. The Social Security's Office of Chief Actuary has published life tables for the Social-Security-area population based on historical and forecast death rates (Bell and Miller

2005). The published tables include estimated and projected death rates for successive years of age from birth to 119 for a handful of birth cohorts. We interpolated the published estimates for selected birth cohorts in order to obtain estimates of age-specific mortality rates for the 1925, 1935, and 1945 birth cohorts. The relevant estimates are displayed as broken lines in Figure 1. There is a tolerably close correspondence between the level and pattern of age-specific mortality rates in the HRS sample and the Actuary's estimates and forecasts. The male mortality estimates for the 1925 cohort are remarkably similar. Our estimates of HRS mortality in the 1935 and 1945 cohorts show a slower increase in death rates at higher ages than was forecast by the Actuary, but mortality rates at younger ages are very close. In the case of women there is less correspondence between HRS mortality rates and those estimated or projected by the Actuary's Office. The Office of the Actuary forecasts small reductions in mortality rates at a given age for women in the three birth cohorts we examine. As can be seen in the bottom panel of Figure 1, however, age-specific mortality rates in the HRS female sample increased in the 1935 cohort compared with the 1925 cohort. Our estimates of female mortality in the 1925 cohort are noticeably lower than the Actuary's estimates for the same cohort. In contrast, our estimates of female mortality in the 1945 and especially in the 1935 cohort are higher than the estimates of the Actuary.

*The impact of low career earnings on mortality.* The specification in equation 1 can be modified in a straightforward way to reflect the effects of socioeconomic disadvantage on the risk of mortality in middle and old age. Table 3 shows logistic regression coefficients of a model that includes an indicator of socioeconomic disadvantage, namely, our determination of whether an individual was in the lower half of the family income distribution based on his or her predicted average earnings between ages 41 and 50. As noted above, this determination is based on combined husband and wife predicted earnings, adjusted for family size, and reflects the person's income position compared with other people of the same gender who were born in or around the same year. The results show that for both men and women the impact of low predicted income is to boost significantly the risk of death. The estimated coefficient on the interaction between age and low predicted income is negative, indicating that the differential declines with age. Based on the evidence in this sample, however, the HRS respondents with low predicted income face higher mortality risk relative to people with high predicted income well into their 90s.

Figure 2 shows implied estimates of the age-specific mortality rate for persons with low and high predicted income born in 1935. The broken lines in the charts refer to estimated mortality rates among men or women with low predicted incomes; the solid lines show estimated death rates among those with high predicted incomes. There is clearly a sizeable gap between the predicted mortality rates of the two groups. The mortality differential caused by low income can be viewed in two ways. First, we can estimate the mortality rate in the low-income group as a multiple of that in the high-income group at the same age. By that standard our estimates imply that the SES mortality differential narrows steadily at successively higher ages. Among men born in 1935, the mortality ratio at age 60 is 2.56. By age 80 the predicted ratio falls to 1.43. Among women born in the same year, the comparable ratios are 2.25 at age 60 and 1.37 at age 80. The estimates also imply that mortality increases markedly at older ages, however. Even though the ratio of mortality rates declines at successively higher ages, the absolute difference between mortality rates in the low- and high-predicted-income populations continues to rise up through about age 80. An alternative measure of the mortality differential is the percentage-point gap between the age-specific death rates in the low- and high-income populations. For HRS sample members born in 1935 the predicted maximum differences are 2.1 percentage points for men (at age 79) and 1.6 percentage points for women (at age 82).

*The trend in mortality differentials across cohorts.* Using a simple measure of socioeconomic status based on differences in predicted midcareer earnings, we find, as did earlier researchers, that mortality differs substantially between populations with high and low predicted incomes. Has the mortality gap widened across successive cohorts? Our ability to detect a widening gap is limited by the range of birth cohorts and ages represented in the HRS sample. To examine the question we divided the HRS birth cohorts into two groups, those born up through 1930 and those born in 1931 and later years. Sample members in the later cohorts are identified as "Late birth." If mortality differences between people with low and high socioeconomic status have widened, we would expect to see wider differentials at a given year of age among HRS respondents who were born after 1930.

In Table 4 we present logistic regression coefficients showing two tests of the hypothesis. Results for men are displayed on the left; results for women are on the right. Both tests use estimates of respondents' midcareer earnings to distinguish between those with above-average and below-average incomes. In the first and third columns we use respondents' *actual* average

nonzero earnings between ages 41 and 50 to determine their rank in the income distribution.  In the second and fourth columns we use our predictions of respondents' earnings to generate predictions of their rank in the income distribution.  Because Social Security earnings records were only obtained from two thirds of the HRS respondents, the results based on actual earnings rely on a smaller number of observations than the results based on predicted earnings. The specification of mortality risk includes individual characteristics that are known or suspected to be linked to differences in death rates, including race, ethnicity, and marital status. The crucial variables in the specification are the two interaction terms with "Late birth."

Though the results differ in detail, both sets of results imply that the mortality differential between high and low SES older Americans was greater in the later birth cohorts compared with the earlier ones. The coefficients on the interaction term of Late birth year and Midcareer income in the bottom half of the income distribution are always positive and statistically significant, implying that age-specific death rates increased among low income recipients in the late birth cohorts compared with the earlier ones. On the other hand, the coefficients on the interaction term of Late birth and income in the top half of the income distribution are negative, implying that the age-specific death rates among high-income recipients fell in the later birth cohorts. When we use respondents' actual midcareer earnings in the estimation, the coefficients are not statistically significant, but when we measure respondents' SES using *predicted* midcareer earnings the estimated drop in high income recipients' mortality rate is highly statistically significant. The results thus indicate that the mortality differential almost certainly increased due to a relative rise in death rates among the low-income elderly and probably increased because of a relative decline in death rates among the high-income elderly.

In Table 5 we investigate the same mortality trends using education as an indicator of HRS respondents' socioeconomic status. We divided respondents into three educational groups, those who failed to obtain a high school diploma, those who completed high school and possibly one or more years of college but did not obtain a college degree, and those who completed college.  Not surprisingly, people who failed to complete high school have elevated mortality rates and those who completed college have below-average mortality.  The critical coefficients for evaluating the change in the SES mortality differential are those on the interaction terms of Late birth and the educational attainment variables. For both men and women these coefficients imply a sizeable increase in mortality differentials in the later birth cohorts compared with the

earlier ones. Men and women who have a college education saw a significant decline in mortality in the Late birth cohorts compared with the earlier cohorts. Women with less than a high school education saw a sizeable increase in mortality relative to that experienced by women with the same level of schooling in the earlier birth cohorts. Men with a high school diploma or some college education but no college degree saw a drop in mortality relative to men in earlier cohorts who had the same level of schooling. Among women with this intermediate level of schooling there was essentially no change in age-adjusted mortality rates in the Late birth cohorts compared with the earlier ones. Thus, all three of our measures of socioeconomic status – based on actual midcareer earnings (for the sample of persons with a Social Security earnings record), predicted midcareer earnings, and observed educational attainment – suggest that mortality differentials between SES groups have increased in more recent birth cohorts compared with cohorts born before 1930.

## II.    Causes of death and their contribution to widening mortality differentials

The HRS data file contains information on the cause of death for the great majority of sample members who died between 1992 and 2010. These data permit us to investigate the role of SES in explaining the incidence and timing of deaths resulting from alternative causes. Are there two or three causes of death that explain the growing divergence in mortality between high and low SES groups? To examine the issue we re-estimated the discrete time logistic model displayed in Table 4 to focus on single causes of death that are recorded in the HRS file. These estimates reflect the risk that respondents will die at a given age of only one cause. Respondents who die of another cause are included in the estimates in calendar years before their death, but in the year of their deaths their information is excluded from the analysis file.[9] Thus, in each of the estimations the dependent variable takes the value 0 in the case of respondents who survive the calendar year and takes the value 1 in the case of respondents who die in the calendar year as a result of a specific cause, for example, cancer.

Table 6 shows logistic regression results for two specific causes of death. The coefficients listed in columns 1 and 3 show estimates of the effects of risk factors for death as a

---

[9] An alternative procedure is to estimate a more elaborate model of mortality and its multiple causes using multinomial logit. This alternative methodology produced results similar enough to those obtained using logistic regression that neither the model nor the estimates will be discussed further here.

result of cancer. The coefficients listed in columns 2 and 4 show estimates of factors that affect the risk of death as a result of heart conditions. Results displayed in the left-hand columns pertain to men enrolled in the HRS; results on the right are for women in the HRS. The measure of SES used in these regressions is whether the respondent's predicted income falls in the top or bottom half of the household income distribution. As expected, the risk of death due to both cancer and health conditions increases with age. At the same time, mortality attributable to both causes is lower in the top half of the predicted income distribution compared with the bottom half of the distribution. The results indicate there was a significant decline in mortality due to cancer and heart conditions for persons in the top half of the predicted income distribution. That is, compared with people in the top half of the income distribution who were born in 1930 and earlier years, people predicted to have high income and born in later years faced a significantly lower risk of death due to either cancer or heart conditions. On the other hand, we find no statistically significant evidence of a change in mortality due to these causes among people in the bottom half of the predicted income distribution. One conclusion we can draw from the results in Table 6 is that older Americans in the top half of the income distribution saw significant declines in their risk of dying from cancer or heart conditions. There is no evidence for such a decline among older Americans in the bottom half of the distribution.

In appendix Table A2 we show condensed results from our analyses of six other causes of death. Among older Americans of both sexes and for every cause of death there is evidence that people predicted to have incomes in the bottom half of the distribution faced an elevated risk of dying. In many, though not all, cases the risk of death among people in the low-predicted-income group was significantly above that among people in the high-predicted-income group. The more interesting question is whether the mortality difference due to a specific cause of death has widened over time. It most cases it has, though in some cases the widening gap is not statistically significant. In most cases the gap between mortality in the high- and low-predicted-income groups widened because high-income Americans saw a decline in mortality rates. In the case of men in the top predicted income group, there were significant declines in mortality due to "allergies, hay fever, sinusitis, and tonsillitis" and to "miscellaneous causes." Among men in the bottom predicted income group, there were significant increases in mortality due to "digestive system issues," "miscellaneous causes," and causes that were not ascertained. Among women it was more common to find a statistically significant increase in mortality in the low-predicted-

income group than it was to find a significant decline in cause-specific mortality in the high-income group. There was a statistically significant rise in mortality among low-predicted-income women as a result of "allergies, hay fever, sinusitis, and tonsillitis" and "digestive system issues."

In sum, both for men and women we find that the mortality differential increased across a wide variety of specific causes of death.  In some cases, notably cancer and heart disease, people with higher predicted incomes saw faster declines in mortality. In other cases, notably digestive system problems among both men and women, and "allergies, hay fever, sinusitis, and tonsillitis" among women, the increase in mortality differentials was mainly due to a sizeable increase in mortality rates among those with low predicted incomes.

## III. Role of behavior, self-reported health, and parental lifespans

*Behavior Risks.* The HRS contains a number of questions aimed at obtaining self-reported measures of behavioral risks.  We included three of these variables (alcohol use, smoking, and physical activity) in our logistic regressions in order to determine their significance as predictors of mortality and to observe the extent to which they might account for the increase in differential mortality. This issue was examined by Cutler and others (2010) for a broader population.  They concluded that behavioral indicators were significant predictors of mortality, but that they did relatively little to explain the variation in mortality by SES or the sources of differential mortality. In Table 7, we expand our base mortality model to include the number of alcoholic drinks consumed in a typical day averaged across all survey waves, whether the individual ever smoked or was smoking at the time of the last interview, and whether the respondent engaged in vigorous physical activity at least three times per week.  We also included a variable indicating whether the respondent's parents died before 65 years of age. While we obtained results using three measures of socio-economic status (actual income, predicted income, and education), the table shows our findings based on actual income. The alternative measures of SES yield similar results.

The findings, shown in column 2, are similar to those reported in the 2010 paper by Cutler and others in that the behavioral variables have high significance as predictors of mortality. However, the inclusion of behavioral variables implies that the underlying mortality risk of the low SES group is increasing for the later cohorts: the interaction terms become positive for the high earners born after 1930 and more positive for their low-earning

-13-

counterparts. On the other hand, the inclusion of behavioral variables does not alter the risk differential for men born before 1930, but it further reduces the risk for women in the high SES group compared to those of lower status.

 *Self-reported health status.* Self-reported health status has been investigated both as a correlate of mortality and as a measure of the quality of life (Cutler and others, 2008). Each wave of the HRS includes a measure of self-reported health status that ranges from a value of 1 for excellent to 5 for poor. The effect of including respondents' self-reported health at the time of their enrollment into the HRS as a predictor of later mortality is shown in columns 3 of Table 7. The estimated coefficient on self-reported health is positive and highly statistically significant, but its inclusion has only a modest impact on the interaction terms between earnings and Late birth as a measure of the change in differential mortality. There is a slight increase in the coefficients on the interaction terms between Late birth and both low and high earnings, implying a rise in mortality risk for all respondents born after 1930. However, the inclusion of initial health status in addition to behavioral indicators appears to reduce the mortality differential between the high and low SES groups born before 1930 for both men and women.

 The inclusion of the behavioral variables and the self-reported health condition measure results in a substantial improvement in the explanatory power of the mortality risk regression, but it does not do much to identify the source of the observed increase in differential mortality. Instead, after accounting for these factors, we find a pattern of increasing mortality risk within the late birth cohorts, for both low- and high-income individuals. Very similar results were obtained using predicted career earnings and education for the basic SES criterion.

## IV. Measures of SES: Education or income?

 Socioeconomic status is usually assessed using income or education or some combination of the two variables. In examinations of the link between SES and mortality, most studies have used education because its measurement is easy and practical in survey contexts. Furthermore, it is included as an element of most American death certificates. In addition, education is often preferred because it is determined in early adulthood and is therefore less likely to be influenced by other determinants of mortality, such as general health status. It is especially likely to be correlated with related aspects of lifestyle and behavior that are useful indicators of SES. In contrast, current income as reported in surveys is potentially influenced by health – a direct

influence on mortality – making it a less independent measure of SES. Moreover, current income is subject to large transient influences that reduce its usefulness as a basic indicator of SES.

Education as an indicator of SES has some limitations, however. Relative to income, there is decreasing variability in years of education. Also, most studies of SES use an individual's education, and that may miss the role of the educational attainment of other family members (especially the spouse) in determining a person's social and economic status. Furthermore, education is not particularly useful as a policy instrument for social programs. For example, the Social Security program determines benefits on the basis of an individual's lifetime earnings, not educational attainment. If there is evidence that mortality differentials are widening and policymakers conclude that a recalibration of the benefit formula is needed to compensate low-income contributors for their relatively small gains in longevity, it is hard to see how measures of widening mortality differentials by educational attainment can be useful in reforming the formula.

The fact that a majority of HRS interview records are linked to Social Security earning records provides us with an unusual opportunity to use "career earnings" as an income-based measure of SES, while avoiding many of the problems caused by measuring SES on the basis of current income. A 10-year average of mid-career earnings comes close to the concept of permanent income. Averaging income over such a long period sharply reduces the role of transitory influences on income. Further, the focus on mid-life income reduces, though it does not eliminate, the potential for reverse causation flowing from health to income.

We compared education and career earnings as measures of SES for a common data sample limited to persons who granted` access to their earning records. For single persons, our measure of income is the individual's average nonzero earnings between ages 41 and 50. For married persons it is the square root of the couple's combined average nonzero earnings between ages 41 and 50. We estimated three versions of the basic mortality equation using actual earnings, predicted earnings, and education as alternative measures of SES. Birth year is classified as early (1930 and earlier years) and late cohorts (born 1931 and later). Career earnings are recorded as above and below the median, and education is classified in three categories: less than high school, high school completion or some college, and college degree holder. Other variables included are age, race, and marital status. The resulting estimation results are reported in Appendix Table A3.Career income and education appear to work equally well as measures of

SES.  The three versions have virtually identical statistics for goodness of fit; and in each case the interaction between birth year and the measure of SES yields the same conclusion of increasing differences in mortality rates for later birth cohorts. The most notable difference is a larger increase in deferential mortality across high and low earners in the regressions based on predicted as opposed to actual earnings. In part, that may be due to the low variance of predicted earnings.

The potential for using career earnings or even the concept of average indexed monthly earnings (AIME) from the social security program, as a proxy measure of SES is more useful than education for evaluating policy responses.  For example, if differences in life expectancy by SES are increasing over time, the benefit formula for retirement benefits could be altered to increase the progressivity of annual benefits so as to offset any widening in the distribution of lifetime benefits.

## V. Conclusion

Mortality data collected in the HRS study cover deaths over a 19-year span for randomly selected samples of older Americans born over a period of more than four decades. A unique advantage of the HRS data file is that it combines extensive survey data on respondents' health, family circumstances, and income in old age with detailed information about the timing and cause of their deaths.  For a large majority of the sample members we also have lifetime Social Security earnings records, which can be used to develop reliable estimates of sample members' economic circumstances in years before their enrollment in the sample. Our tabulations of the mortality experience in the HRS sample suggest that it is similar to the mortality experience estimated and projected by the Social Security Actuary for the broader U.S. population, especially in the case of men. The mortality experience of women in the HRS sample differs from that estimated or projected by the Social Security Actuary in that age-specific mortality rates in the HRS sample did not uniformly decline across successive birth cohorts.

In examining the relationship between mortality experience and social and economic status, we developed three measures of status.  Two were based on the Social Security earnings records that were linked to HRS interview records. For sample members who permitted the records to be linked, we estimated average nonzero earnings received by sample members between ages 41 and 50. To reflect the earnings contribution of a spouse, we combined husband and wife earnings and adjusted the summed earnings to reflect differences in family size between

single earners, on the one hand, and married-couple families, on the other. One of our measures of SES estimated the relative position of HRS sample members in the income distribution by their actual earnings between 41 and 50. The sample for which this measure can be estimated is limited to HRS respondents who provided access to their Social Security earnings records. Our second measure of SES is based on HRS sample members' predicted earnings in mid-career. Finally, we also used HRS respondents' educational attainment to determine SES.

Using all three measures of SES we find a significant widening of mortality differentials between high and low SES groups. That is, the mortality differences between high- and low-income older Americans is larger in younger birth cohorts compared with older ones. We find that mortality differentials widened for a broad variety of causes of death. Two leading causes of death are cancer and heart conditions. For these causes of death the main contributor to wider mortality differences is the fact that higher income older Americans have seen a faster decline in mortality compared with lower income Americans, whose mortality as a result of cancer and health conditions has not declined and may even have increased.

Though the HRS data file has unusually complete information about old-age income, health status, family circumstances, and lifetime earnings, it has limitations for studying the long-term trend in mortality differences by SES. The oldest people enrolled in the HRS sample – that is, the sample members representing the oldest birth cohorts – were already past 70 when they were enrolled. It is therefore impossible to observe the mortality experience of this elderly birth cohort when its members were the same age as younger birth-year cohorts when those younger cohorts were first enrolled in the sample. The most reliable comparisons of SES mortality differentials in the sample are therefore restricted to the overlapping age ranges when both the older and the younger cohorts' mortality experiences can be observed. We cannot, for example, directly compare the mortality experiences of the oldest and youngest birth cohorts at ages 55 or 60. Nonetheless, the wealth of data in the HRS permits an unusually rich analysis of the factors contributing to early mortality and the link between those factors and SES.

**References**

Bell, Felicitie C., and Michael L. Miller. 2005. "Life Tables for the United States Social Security Area." Actuarial Study No. 120. Washington, DC: Office of the Chief Actuary, U.S. Social Security Administration.

Brown, Dustin, Mark Hayward, Jennifer Karras Montez, Robert Hummer, Chi-Tsun Chiu, and Mira Hidajat. 2012. "The Significance of Education for Mortality Compression in the United States." *Demography* 49(3): 819-40.

Cutler, David, Fabian Lange, Ellen Meara, Seth Richards, and Christopher Ruhm. 2010. "Explaining the Rise in Educational Gradients in Mortality," NBER Working Paper 15678 (January). Cambridge, MA: National Bureau of Economic Research.

Feinglass, Joe, and others. "Baseline health, socioeconomic status, and 10-year mortality among older middle-aged Americans: Findings from the Health and Retirement Study, 1992-2002.," *Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*. 62B (4), S209-S217.

Hoffman, Rasmus. 2005. "Do socioeconomic mortality differences decrease with rising age?" *Demographic Research*, 13, 35–62.

Mackenbach JP, and others. 2003. "Widening socioeconomic inequalities in mortality in six Western European countries," *International Journal of Epidemiology*, 32: 830–37.

Meara, Ellen, Seth Richards, and David M. Cutler. 2008. "The Gap Gets Bigger: Changes In Mortality and Life Expectancy, By Education, 1981-2000." *Health Affairs* 27: 350-360.

Olshansky, S. Jay, and others. 2012. "Differences In Life Expectancy Due To Race And Educational Differences Are Widening, And Many May Not Catch Up." *Health Affairs*, 31, no. 9:1803-1818.

Singh, Gopal K., and Mohammad Siahpush. 2006. "Widening Socioeconomic Inequalities in US Life Expectancy, 1980–2000," *International Journal of Epidemiology* 35:969–979.

Steenland, Kyle, J. Henley, and M. Thun, 2002. "All-Cause and Cause-Specific Death Rates by Educational Status for Two Million People in Two American Cancer Society Cohorts, 1959–1996," *American Journal of Epidemiology* 156, no. 1: 11–21.

Waldron, Hilary. 2007. "Trends in mortality differentials and life expectancy for male Social Security-covered workers, by socioeconomic status." *Social Security Bulletin* 67, no. 3:1-28.

**Table 1. Health and Retirement Study, Cohort Characteristics**

| Sample | AHEAD (Aging and Health Dynamics) | CODA (Children of the Depression) | HRS (Health and Retirement Survey) | WB (War Babies) | EBB (Early Baby Boomers) | Total |
|---|---|---|---|---|---|---|
| Size of sample | 8,444 | 2,420 | 13,513 | 2740 | 4758 | 31,875 |
| Target birth years | before 1924 | 1924-1930 | 1931-1941 | 1942-1947 | 1948-1953 | 1890-1965 |
| Date of first enrollment | 1993 | 1998 | 1992 | 1998 | 2004 | 1992-2004 |
| Percent with Social Security earnings file | 45.3 | 70.0 | 79.3 | 70.2 | 42.0 | 63.2 |
| Percentage with benefit file | 41.7 | 65.2 | 44.5 | 30.6 | 10.1 | 39.0 |
| Deaths | 6,705 | 995 | 3,813 | 272 | 152 | 11,937 |

*Source:* Calculated by the authors from the micro-data files of the HRS. The number of deaths are as recorded in Version M of the Rand-HRS file.

**Table 2. Relative Rates of Mortality by Socioeconomic Status, Health and Retirement Study, 1992-2010**

| | | | | Relative Mortality | | |
|---|---|---|---|---|---|---|
| | | | **Women** | | **Men** | |
| Group | | | **50-74** | **75 & over** | **50-74** | **75 & over** |
| | | | | Total Sample | | |
| **Race** | | | 151,206 | 66,902 | 118,749 | 43,913 |
| Black | | | 1.46 | 1.09 | 1.34 | 1.19 |
| Hispanic | | | 0.85 | 0.87 | 0.81 | 0.98 |
| Other | | | 1.22 | 0.98 | 1.15 | 1.06 |
| Non-Hispanic whites | | | 0.90 | 1.00 | 0.96 | 0.97 |
| | | | | | | |
| **Education** | | | 151,171 | 66,900 | 118,700 | 43,913 |
| LT high-school | | | 1.54 | 1.21 | 1.39 | 1.18 |
| High-school graduate | | | 0.96 | 0.89 | 0.97 | 0.98 |
| Some college | | | 0.74 | 0.85 | 0.93 | 0.88 |
| College and above | | | 0.51 | 0.80 | 0.62 | 0.73 |
| | | | | | | |
| **Marital Status** | | | 151,142 | 60,484 | 118,583 | 43,911 |
| Married/Widowed | | | 0.96 | 1.00 | 0.91 | 0.99 |
| Never married | | | 1.32 | 1.15 | 1.21 | 1.24 |
| Separated/divorced | | | 1.13 | 0.97 | 1.43 | 1.07 |
| | | | | | | |
| | | | | Social Security Earnings Sample | | |
| | | | | | | |
| **Actual Household career earnings by quartile** | | | 102,651 | 35,696 | 81,278 | 28,516 |
| | | 1 | 1.35 | 0.96 | 1.39 | 1.03 |
| | | 2 | 1.01 | 0.90 | 0.96 | 0.97 |
| | | 3 | 0.74 | 0.84 | 0.74 | 0.87 |
| | | 4 (top) | 0.54 | 0.81 | 0.66 | 0.77 |
| | | | | | | |
| **Predicted Household career earnings by quartile** | | | 150,787 | 66,867 | 118,426 | 43,881 |
| | | 1 | 1.65 | 1.18 | 1.57 | 1.18 |
| | | 2 | 1.06 | 1.02 | 1.14 | 1.10 |
| | | 3 | 0.85 | 0.95 | 0.75 | 0.92 |
| | | 4 (top) | 0.53 | 0.87 | 0.57 | 0.84 |

*Source:* Computed by the authors from data of the Health and Retirement Study. Relative mortality is the mortality rate of the specific group (*e.g. persons with less than a high school education*) divided by the average mortality rate for all people of the same age group and sex. Rates are based on annual observations over the interval of 1994 to 2009.

**Table 3. Logistic Regression Predicting Respondent Annual Mortality Risk in HRS Sample, 1992-2010**

|  | Men | | Women | |
| --- | --- | --- | --- | --- |
|  | Coef. | Sig. | Coef. | Sig. |
| Intercept | 5.92 |  | -14.39 | ** |
| Age | 0.098 | ** | 0.111 | ** |
| Low predicted income (Yes = 1) | 2.673 | ** | 2.268 | ** |
| Age x Low predicted income | -0.029 | ** | -0.024 | ** |
| Birth year | -0.0086 | ** | 0.0012 |  |
| First year in sample (Yes = 1) | -1.757 | ** | -2.450 | ** |
| *No. of observations* | 166,343 | | 231,560 | |
| *Pseudo R²* | 0.032 | | 0.034 | |

*Source:* Authors' logistic regression estimates from the HRS micro-data files as described in text. Dependent variable indicates respondent's death during the calendar year. The indicator of "low income" used in the regression reflects combined predictions of husband and wife earnings between ages 41 and 50, adjusted for family size.

*\*\* : p<0.01*
*\* : p<0.05*

**Table 4.  Logistic Regression Predicting Respondent Annual Mortality Risk and Its Relationship to Midcareer Earnings in HRS Sample, 1992-2010**

| | Men | | | | Women | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | (1) Actual earnings | | (2) Predicted earnings | | (3) Actual earnings | | (4) Predicted earnings | |
| | Coef. | Sig. | Coef. | Sig. | Coef. | Sig | Coef. | Sig. |
| Intercept | -10.20 | ** | -9.39 | ** | -11.32 | ** | -10.76 | ** |
| Age | 0.10 | ** | 0.088 | ** | 0.11 | ** | 0.10 | ** |
| Midcareer income in top half  (Yes = 1) | -0.23 | ** | -0.32 | ** | -0.13 | ** | -0.21 | ** |
| Midcareer income in top half X Late birth (Yes = 1) | -0.07 | | -0.31 | ** | -0.15 | | -0.16 | ** |
| Midcareer income in bottom half X Late birth (Yes = 1) | 0.17 | ** | 0.14 | ** | 0.36 | ** | 0.31 | ** |
| **Race / ethnicity** | | | | | | | | |
| Black (Yes = 1) | 0.27 | ** | 0.02 | ** | 0.14 | ** | 0.11 | ** |
| Hispanic (Yes = 1) | -0.17 | * | -0.35 | ** | -0.15 | * | -0.24 | ** |
| White or other (Yes = 1) | *Reference group* | | | | *Reference group* | | | |
| **Marital status** | | | | | | | | |
| Married or widowed (Yes = 1) | -0.31 | ** | -0.11 | * | -0.12 | ** | -0.15 | * |
| Separated or divorced (Yes = 1) | 0.16 | | 0.12 | | 0.02 | | -0.08 | |
| Never married (Yes = 1) | *Reference group* | | | | *Reference group* | | | |
| First year in sample (Yes = 1) | -1.72 | ** | -1.70 | ** | -2.20 | ** | -2.43 | ** |
| *No. of observations* | 112,143 | | 166,343 | | 146,293 | | 231,560 | |
| *Pseudo R²* | 0.029 | | 0.034 | | 0.027 | | 0.0342 | |

*Source:* Authors' logistic regression estimates from the HRS micro-data files as described in text. Dependent variable indicates respondent's death during the calendar year. The indicator of "low income" used in the regression reflects combined husband and wife earnings between ages 41 and 50, adjusted for family size.

*Note:* The results displayed under "Actual earnings" were obtained based on the observed earnings of HRS respondents between ages 41 and 50. Results listed under "Predicted earnings" were obtained using a larger sample and based on respondents' predicted earnings between ages 41 and 50.

*\*\* : p<0.01*
*\* : p<0.05*

**Table 5. Logistic Regression Predicting Respondent Annual Mortality Risk and Its Relationship to Educational Attainment in HRS Sample, 1992-2010**

|  | (1) Men | | (2) Women | |
|---|---|---|---|---|
|  | Coef. | Sig. | Coef. | Sig. |
| Intercept | -9.181 | ** | -10.734 | ** |
| Age | 0.086 | ** | 0.098 | ** |
| Less than high school  (Yes = 1) | 0.152 | ** | 0.222 | ** |
| College degree or more (Yes = 1) | -0.308 | ** | -0.075 | |
| Less than high school X Late birth (Yes = 1) | 0.028 | | 0.322 | ** |
| High school or some college X Late birth (Yes = 1) | -0.119 | * | 0.031 | |
| College degree or more X Late birth (Yes = 1) | -0.169 | * | -0.408 | ** |
| **Race / ethnicity** | | | | |
| Black (Yes = 1) | 0.148 | ** | 0.148 | ** |
| Hispanic (Yes = 1) | -0.220 | ** | -0.248 | ** |
| White or other (Yes = 1) | *Reference group* | | | |
| **Marital status** | | | | |
| Married or widowed (Yes = 1) | -0.346 | ** | -0.237 | ** |
| Separated or divorced (Yes = 1) | 0.048 | | -0.008 | |
| Never married (Yes = 1) | *Reference group* | | | |
| First year in sample (Yes = 1) | -1.718 | ** | -2.450 | ** |
| *No. of observations* | 166,495 | | 231,960 | |
| *Pseudo R^2* | 0.032 | | 0.034 | |

   *Source:* Authors' logistic regression estimates from the HRS micro-data files as described in text. Dependent variable indicates respondent's death during the calendar year. The indicator of "low income" used in the regression reflects combined husband and wife earnings between ages 41 and 50, adjusted for family size.

*** : $p<0.01$*

 *** : $p<0.05$*

**Table 6.  Logistic Regression Predicting Respondent Annual Mortality Risk Due to Cancer and Heart Conditions, 1992-2010**

|  | Men | | | | Women | | | |
|---|---|---|---|---|---|---|---|---|
|  | (1) | | (2) Heart conditions | | (3) | | (4) Heart conditions | |
|  | Cancer | | conditions | | Cancer | | conditions | |
|  | Coef. | Sig. | Coef. | Sig. | Coef. | Sig. | Coef. | Sig. |
| Intercept | -9.41 | ** | -10.74 | ** | -8.84 | ** | -12.23 | ** |
| Age | 0.066 | ** | 0.090 | ** | 0.052 | ** | 0.107 | ** |
| Midcareer income in top half  (Yes = 1) | -0.23 | ** | -0.30 | ** | -0.18 | * | -0.21 | ** |
| Midcareer income in top half X Late birth (Yes = 1) | -0.37 | ** | -0.62 | ** | -0.42 | ** | -0.55 | ** |
| Midcareer income in bottom half X Late birth (Yes = 1) | 0.071 | | 0.009 | | -0.016 | | 0.029 | |
| **Race / ethnicity** | | | | | | | | |
| Black (Yes = 1) | 0.065 | | -0.017 | | 0.045 | | 0.060 | |
| Hispanic (Yes = 1) | -0.61 | ** | -0.31 | ** | -0.50 | ** | -0.29 | ** |
| White or other (Yes = 1) | *Reference group* | | | | *Reference group* | | | |
| **Marital status** | | | | | | | | |
| Married or widowed (Yes = 1) | 0.11 | | 0.03 | | 0.13 | | -0.23 | * |
| Separated or divorced (Yes = 1) | 0.14 | | 0.30 | * | 0.19 | | -0.17 | |
| Never married (Yes = 1) | *Reference group* | | | | *Reference group* | | | |
| First Year in Sample (Yes = 1) | -2.24 | ** | -1.96 | ** | -2.73 | ** | -3.30 | ** |
| *No. of observations* | 162,023 | | 162,603 | | 226,646 | | 227,522 | |
| *Pseudo R$^2$* | 0.006 | | 0.013 | | 0.003 | | 0.015 | |

*Source:* Authors' logistic regression estimates from the HRS micro-data files as described in text. Dependent variable indicates respondent's death during the calendar year as a result of the cause of death listed. The indicator of "low income" used in the regression reflects combined predicted husband and wife earnings between ages 41 and 50, adjusted for family size.

*Note:* The estimates shown under columns 1 and 3 exclude observations of deaths attributable to a cause other than cancer; the estimates shown under columns 2 and 4 exclude observations of deaths attributable to a cause other than heart conditions. See text.

*\*\* : p<0.01*
*\* : p<0.05*

## Table 7.  Logistic Regression Predicting Respondent Annual Mortality Risk With Behavioral and Health Indicators

| | Men | | | | | |
|---|---|---|---|---|---|---|
| | Base Model | | Behavior | | Behavior and Health | |
| | Coef. | Sig. | Coef. | Sig. | Coef. | Sig. |
| Intercept | -10.20 | ** | -13.49 | ** | -14.68 | ** |
| Age | 0.098 | ** | 0.113 | ** | 0.116 | ** |
| High Earnings | -0.232 | ** | -0.228 | ** | -0.174 | ** |
| High Earnings * Late Birth | -0.068 | | 0.234 | ** | 0.395 | ** |
| Low Earnings * Late Birth | 0.169 | ** | 0.517 | ** | 0.596 | ** |
| Black | 0.266 | ** | 0.252 | ** | 0.167 | ** |
| Hispanic | -0.169 | * | -0.103 | | -0.142 | |
| Married/Widowed | -0.314 | ** | -0.096 | | -0.075 | |
| Separated/Divorced | 0.159 | | 0.050 | | 0.044 | |
| First year in sample | -1.720 | ** | -1.918 | ** | -1.972 | ** |
| **Behavior** | | | | | | |
| <1 Drink/Day | | | 1.402 | ** | 1.412 | ** |
| 1-2 Drinks/Day | | | 2.193 | ** | 2.176 | ** |
| 3-4 Drinks/Day | | | 2.590 | ** | 2.627 | ** |
| 5+ Drinks/ Day | | | 0.953 | * | 0.911 | * |
| Smoked at time of last interview | | | 0.633 | ** | 0.596 | ** |
| Ever Smoked | | | 0.243 | ** | 0.225 | ** |
| Vigorous physical activity 3+/week | | | -0.660 | ** | -0.373 | ** |
| **Parents' longevity** | | | | | | |
| Father: Early Death | | | -0.059 | | -0.071 | |
| Mother: Early Death | | | 0.082 | | 0.074 | |
| **Health** | | | | | | |
| Self Reported 1st Wave Health | | | | | 0.30 | ** |
| *No. of observations* | 112,143 | | 111,102 | | 111,085 | |
| *Pseudo R²* | 0.029 | | 0.043 | | 0.046 | |

(Continued on next page.)

## Table 7.  Logistic Regression Predicting Respondent Annual Mortality Risk With Behavioral and Health Indicators (continued)
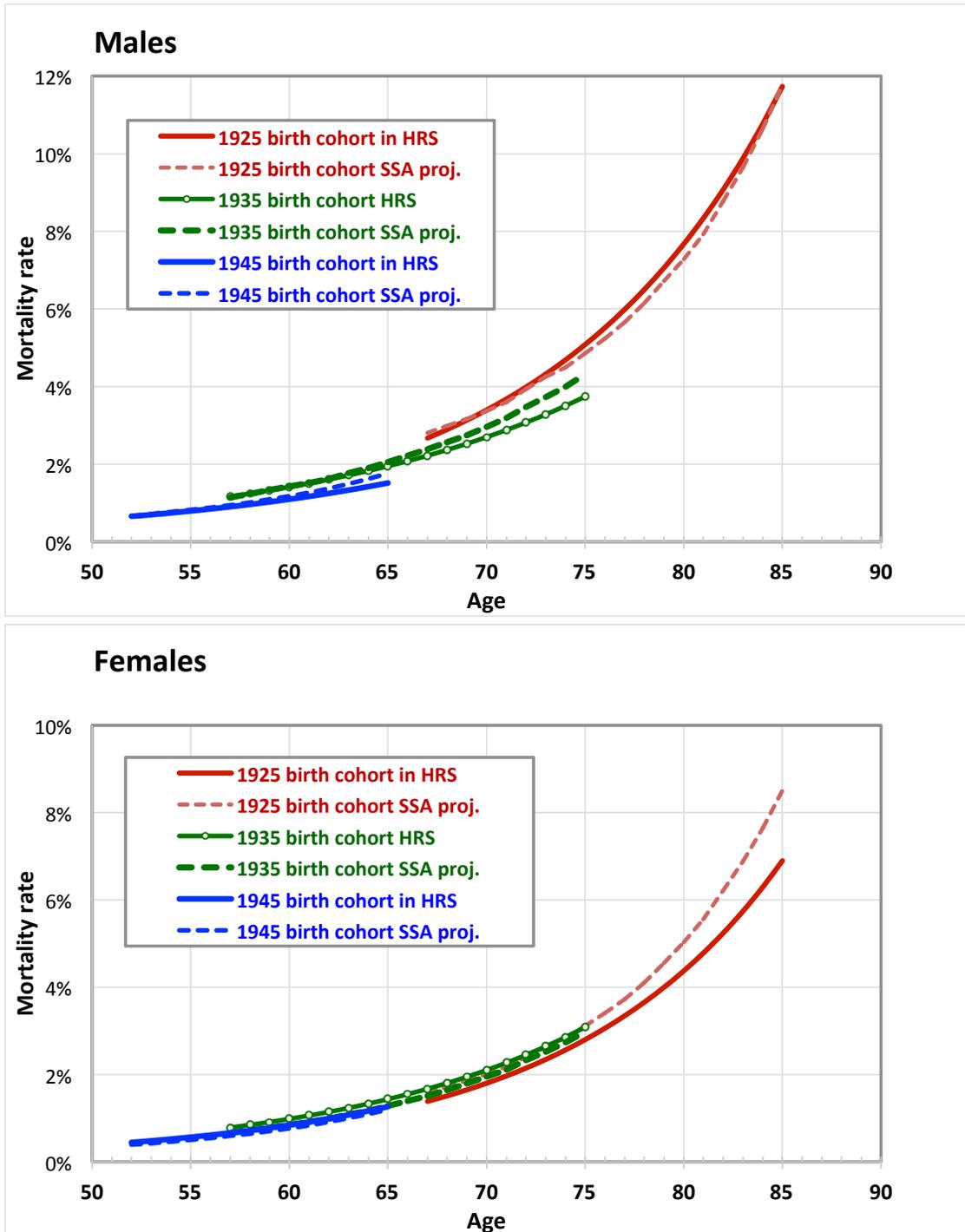
| | Women | | | | | |
|---|---|---|---|---|---|---|
| | Base Model | | Behavior | | Behavior and Health | |
| | Coef. | Sig. | Coef. | Sig. | Coef. | Sig. |
| Intercept | -11.32 | ** | -15.13 | ** | -16.26 | ** |
| Age | 0.105 | ** | 0.125 | ** | 0.127 | ** |
| High Earnings | -0.134 | ** | -0.160 | ** | -0.121 | * |
| High Earnings * Late Birth | -0.147 | | 0.359 | ** | 0.533 | ** |
| Low Earnings * Late Birth | 0.361 | ** | 0.950 | ** | 1.037 | ** |
| Black | 0.143 | ** | 0.216 | ** | 0.081 | |
| Hispanic | -0.154 | | -0.010 | | -0.166 | |
| Married/Widowed | -0.123 | | 0.181 | | 0.101 | |
| Separated/Divorced | 0.017 | | 0.094 | | -0.033 | |
| First year in sample | -2.202 | ** | -2.423 | ** | -2.481 | ** |
| **Behavior** | | | | | | |
| <1 Drink/Day | | | 1.780 | ** | 1.794 | ** |
| 1-2 Drinks/Day | | | 2.576 | ** | 2.559 | ** |
| 3-4 Drinks/Day | | | 2.960 | ** | 3.041 | ** |
| 5+ Drinks/ Day | | | 2.539 | * | 2.552 | * |
| Smoked at time of last interview | | | 0.522 | ** | 0.475 | ** |
| Ever Smoked | | | 0.184 | ** | 0.201 | ** |
| Vigorous physical activity 3+/week | | | 0.735 | ** | -0.394 | ** |
| **Parents' longevity** | | | | | | |
| Father: Early Death | | | 0.052 | | 0.012 | |
| Mother: Early Death | | | 0.018 | | -0.014 | |
| **Health** | | | | | | |
| Self-reported 1st wave health | | | | | 0.32 | ** |
| *No. of observations* | 146,293 | | 145,207 | | 145,207 | |
| *Pseudo R²* | 0.029 | | 0.042 | | 0.044 | |

*Source:* Authors' logistic regression estimates from the HRS micro-data files as described in text. Dependent variable indicates respondent's death during the calendar year. The "career earnings" estimates used in the regression reflect actual combined husband and wife earnings between ages 41 and 50 adjusted for family size.

*\*\* : p<0.01*

*\* : p<0.05*

**Figure 1. Mortality Rates in HRS Sample and Estimated by the Social Security Actuary in Three Birth Cohorts**



**Males**

Legend:
- 1925 birth cohort in HRS
- 1925 birth cohort SSA proj.
- 1935 birth cohort HRS
- 1935 birth cohort SSA proj.
- 1945 birth cohort in HRS
- 1945 birth cohort SSA proj.

**Females**

Legend:
- 1925 birth cohort in HRS
- 1925 birth cohort SSA proj.
- 1935 birth cohort HRS
- 1935 birth cohort SSA proj.
- 1945 birth cohort in HRS
- 1945 birth cohort SSA proj.

*Source:* Authors' tabulations of HRS data as explained in text and Office of the Actuary (2005).

**Figure 2. Predicted Probability of Death by Age, Persons with High and Low Career Income Born in 1935**
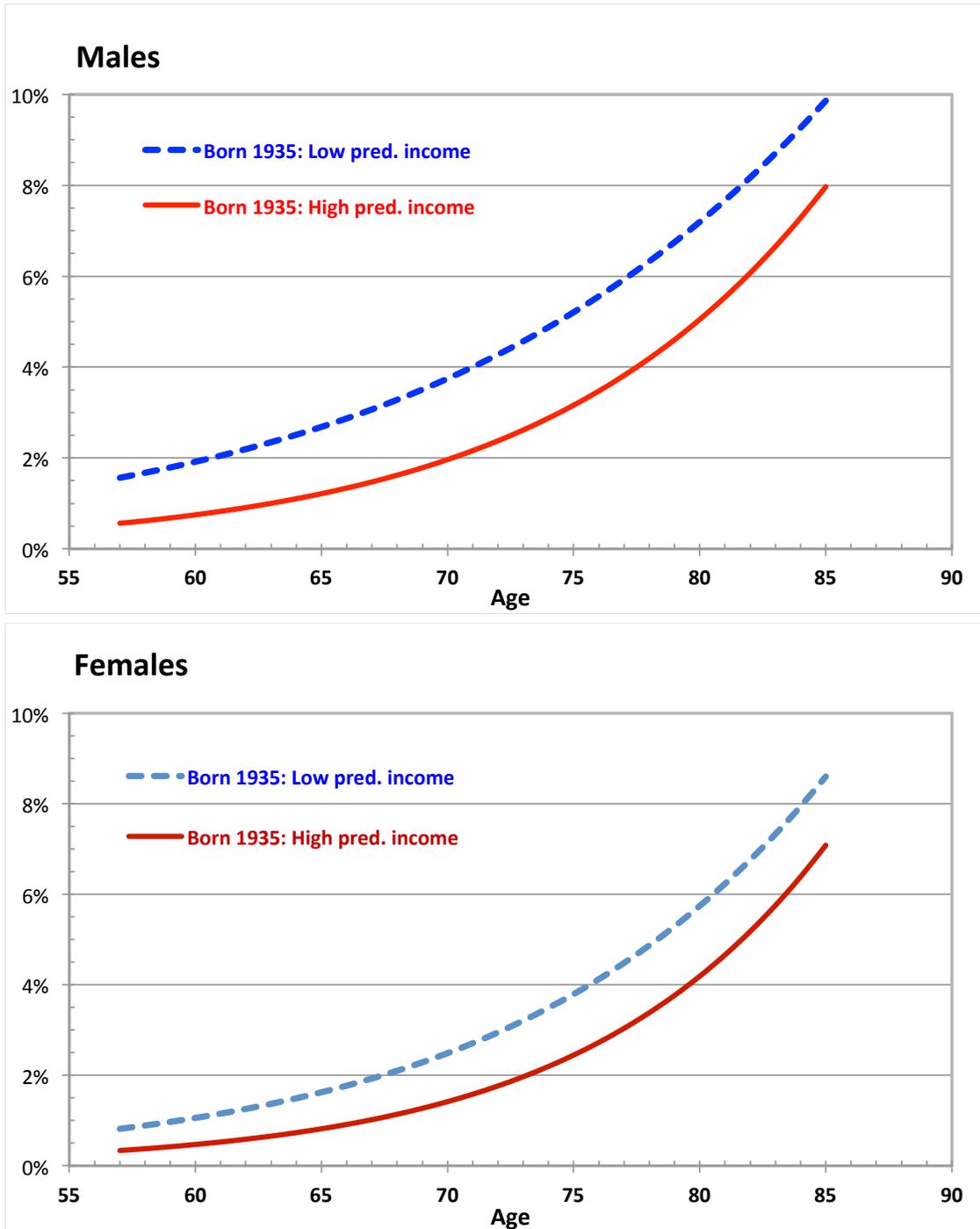


*Source:* Authors' tabulations of HRS data as explained in text.

## Table A1a. Linear Predictors of Career Earnings, Individuals

Thousands of 2005 dollars

| Coefficient | Men | | Women | |
|---|---|---|---|---|
| | *Value* | *t-value* | *Value* | *t-value* |
| Intercept | 33.39 | 20.19 | 11.03 | 12.67 |
| **Education** | | | | |
| Less than high school | -7.17 | -7.56 | -4.61 | -9.18 |
| High school graduate | *Reference group* | | *Reference group* | |
| Some college | 3.72 | 3.56 | 4.18 | 8.01 |
| College and above | 21.00 | 20.40 | 15.98 | 27.92 |
| **Birth year** | 0.32 | 9.68 | 0.17 | 10.46 |
| **Race** | | | | |
| Black | -11.63 | -10.62 | 0.23 | 0.41 |
| Hispanic | -14.69 | -10.95 | -4.07 | -5.54 |
| Other | -16.24 | -5.37 | -0.23 | -0.15 |
| White | *Reference group* | | *Reference group* | |
| **Never disabled** | 11.58 | 11.44 | 5.13 | 9.34 |
| **Marital Status** | | | | |
| Ever married/ widowed | *Reference group* | | *Reference group* | |
| Never married | -19.30 | -9.97 | 9.46 | 9.49 |
| Separated/divorced | -10.09 | -8.99 | 3.15 | 5.78 |
| *$R^2$* | 0.16 | | 0.17 | |
| *No. of observations* | 8,982 | | 9,935 | |

*Source:* Authors' regression estimates from the HRS micro-data files. Dependent variable is the average of nonzero earnings between ages 41-50.

## Table A1b. Log-Linear Predictors of Career Earnings, Individuals

2005 dollars

| | Men | | Women | |
|---|---|---|---|---|
| *Coefficient* | *Value* | *t-value* | *Value* | *t-value* |
| Intercept | 10.34 | 276.16 | 9.08 | 173.06 |
| **Education** | | | | |
| Less than high school | -0.216 | -10.07 | -0.346 | -11.45 |
| High school graduate | *Reference group* | | *Reference group* | |
| Some college | 0.006 | 0.26 | 0.170 | 5.41 |
| College and above | 0.249 | 10.69 | 0.505 | 14.65 |
| **Birth year** | 0.003 | 4.62 | 0.007 | 7.45 |
| **Race** | | | | |
| Black | -0.349 | -14.07 | 0.063 | 1.92 |
| Hispanic | -0.384 | -12.63 | -0.194 | -4.38 |
| Other | -0.355 | -5.19 | 0.002 | 0.03 |
| White | *Reference group* | | *Reference group* | |
| **Never disabled** | 0.360 | 15.74 | 0.282 | 8.51 |
| **Marital Status** | | | | |
| Ever married/ widowed | *Reference group* | | *Reference group* | |
| Never married | -0.601 | -13.70 | 0.397 | 6.62 |
| Separated/divorced | -0.259 | -10.19 | 0.233 | 7.09 |
| $R^2$ | 0.16 | | 0.17 | |
| *No. of observations* | 8,982 | | 9,935 | |

*Source:* Authors' regression estimates from the HRS micro-data files. Dependent variable is the logarithm of average of nonzero earnings between ages 41-50. The results of these regressions were used to impute predicted average nonzero earnings for men and women in the HRS sample for purposes of calculating their predicted position in the income distribution.

**Table A2.  Logistic Regression Predicting Respondent Annual Mortality Risk Due to Six Causes of Death, 1992-2010**

| | Men | | | | Women | | | |
|---|---|---|---|---|---|---|---|---|
| | Coef. | Sig. | Coef. | Sig. | Coef. | Sig. | Coef. | Sig. |
| | Allergies, etc. | | Digestive system | | Allergies, etc. | | Digestive system | |
| Midcareer income in top half  (Yes = 1) | -0.597 | ** | -0.369 | * | -0.176 | | -0.285 | * |
| Midcareer income in top half X Late birth (Yes = 1) | -0.443 | * | -0.085 | | -0.324 | | -0.143 | |
| Midcareer income in bottom half X Late birth (Yes = 1) | -0.205 | | 0.494 | * | 0.407 | * | 0.643 | ** |
| | Endocrine, metabolic | | Miscellane-ous | | Endocrine, metabolic | | Miscellane-ous | |
| Midcareer income in top half  (Yes = 1) | -0.316 | | -0.316 | * | -0.228 | | -0.081 | |
| Midcareer income in top half X Late birth (Yes = 1) | -0.439 | | -0.770 | * | -0.393 | | -0.215 | |
| Midcareer income in bottom half X Late birth (Yes = 1) | -0.148 | | 0.446 | * | 0.344 | | 0.288 | |
| | Other causes | | Not ascertained | | Other causes | | Not ascertained | |
| Midcareer income in top half  (Yes = 1) | -0.232 | | -0.272 | ** | -0.132 | | -0.417 | ** |
| Midcareer income in top half X Late birth (Yes = 1) | 0.145 | | 0.042 | | -0.136 | | 0.624 | ** |
| Midcareer income in bottom half X Late birth (Yes = 1) | 0.237 | | 0.399 | ** | 0.079 | | 0.705 | ** |

*Source:* Authors' logistic regression estimates from the HRS micro-data files as described in text. Dependent variable indicates respondent's death during the calendar year as a result of the cause of death listed. The indicator of "low income" used in the regression reflects combined predicted husband and wife earnings between ages 41 and 50, adjusted for family size.

*\*\* : p<0.01*
*\* : p<0.05*

**Table A3. Comparison of Logistic Regression Results With Three Measures of Socioeconomic Status, Fixed Sample**

|  | MEN | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Actual Earnings** | | | **Predicted Earnings** | | | **Education** | | |
|  | Coeff. | Sig. |  | Coeff. | Sig. |  | Coeff. | Sig. |
| Intercept | -10.201 | *** | Intercept | -10.356 | *** | Intercept | -10.139 | *** |
| Age | 0.098 | *** | Age | 0.099 | *** | Age | 0.097 | *** |
| Black | 0.266 | *** | Black | 0.083 |  | Black | 0.221 | *** |
| Hispanic | -0.169 | ** | Hispanic | -0.344 | *** | Hispanic | -0.188 | ** |
| Married/Widowed | -0.314 | *** | Married/Widowed | -0.078 |  | Married/Widowed | -0.369 | *** |
| Separated/Divorced | 0.159 |  | Separated/Divorced | 0.211 | * | Separated/Divorced | 0.107 |  |
| First-year dummy | -1.720 | *** | First-year dummy | -1.721 | *** | First-year dummy | -1.745 | *** |
|  |  |  |  |  |  | Less than High School (LTHS) | 0.184 | *** |
| Top half of actual career earnings | -0.232 | *** | Top half of predicted career earnings | -0.353 | *** | College Grad | -0.357 | *** |
| Top half actual earnings & Born after 1930 | -0.068 |  | Top half predicted earnings & Born after 1930 | -0.159 | ** | College Grad & Born after 1930 | 0.031 |  |
|  |  |  |  |  |  | (HS grad or some college) & Born after 1930 | 0.050 |  |
| Bottom half actual earnings & Born after 1930 | 0.169 | *** | Bottom half predicted earnings & Born after 1930 | 0.294 | *** | LTHS & Born after 1930 | 0.131 | * |
| *No. of observations* | *112,143* |  |  | *112,119* |  |  | *112,136* |  |
| *Pseudo R²* | *0.029* |  |  | *0.030* |  |  | *0.029* |  |

(Continued on next page.)

**Table A3. Comparison of Logistic Regression Results With Three Measures of Socioeconomic Status, Fixed Sample**
(continued)

| | WOMEN | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Actual Earnings** | Coeff. | Sig. | **Predicted Earnings** | Coeff. | Sig. | **Education** | Coeff. | Sig. | |
| Intercept | -11.325 | *** | Intercept | -11.286 | *** | Intercept | -11.258 | *** | |
| Age | 0.105 | *** | Age | 0.105 | *** | Age | 0.103 | *** | |
| Black | 0.143 | *** | Black | 0.061 | | Black | 0.111 | ** | |
| Hispanic | -0.154 | * | Hispanic | -0.236 | *** | Hispanic | -0.229 | *** | |
| Married/Widowed | -0.123 | | Married/Widowed | -0.068 | | Married/Widowed | -0.194 | * | |
| Separated/Divorced | 0.017 | | Separated/Divorced | -0.013 | | Separated/Divorced | 0.050 | | |
| First-year dummy | -2.202 | *** | First-year dummy | -2.205 | *** | First-year dummy | -2.225 | *** | |
| | | | | | | Less than High School (LTHS) | 0.269 | *** | |
| Top half of actual career earnings | -0.134 | *** | Top half of predicted career earnings | -0.191 | *** | College Grad | 0.004 | | |
| Top half actual earnings & Born after 1930 | -0.147 | * | Top half predicted earnings & Born after 1930 | -0.135 | * | College  Grad & Born after 1930 | -0.411 | *** | |
| | | | | | | (HS grad or some college) & Born after 1930 | 0.089 | | |
| Bottom half actual earnings & Born after 1930 | 0.361 | *** | Bottom half predicted earnings & Born after 1930 | 0.365 | *** | LTHS & Born after 1930 | 0.355 | *** | |
| *No. of observations* | *146,293* | | | *146,092* | | | *146,276* | | |
| *Pseudo R$^2$* | *0.027* | | | *0.027* | | | *0.027* | | |

*Source:* Logistic regression of annual mortality risk for a sample of individuals in the Health and Retirement Study that have available Social Security earnings records.

*\* : p<0.1    \*\* : p<0.05 \*\*\* : p<0.01*