Increasing Differential Mortality by Lifetime Earnings in the United States

Julian P. Cristia
Inter-American Development Bank
jcristia@iadb.org

February 2008

Abstract
Despite widespread interest in differential mortality, little is known about how these differentials, for the nonretired population, have evolved in the last twenty years. This study aims to fill this gap by providing estimates of trends in differential mortality by measures of lifetime earnings in the 1983 to 2003 period. A unique data set constructed matching data from the Survey of Income and Program Participation to earnings, benefits, and mortality data from reliable administrative sources is used. Results indicate a consistent increase in differential mortality by lifetime earnings for both men and women and across age groups.

The author wishes to acknowledge comments and suggestions by Paul Cullinan, Tom DeLeire, Arlene Holen, Howard Iams, Joseph Kile, Joyce Manchester, Noah Meyerson, John Sabelhaus, Hilary Waldron, Lina Walker, and G. Thomas Woodward as well as seminar participants at the Congressional Budget Office and the 2007 Population Association of America Annual Meeting.

I. Introduction

Since Kitagawa and Hauser (1973) seminal's work, a large and growing body of research has emerged focusing on the extent, causes and trends of differential mortality in the United States. The research effort that ensued has been unique not only in its depth but also in the fact that it encompassed work from researchers from diverse fields such as demography, public health, epidemiology, economics, sociology, and psychology. This overlapping interest across disciplines highlights the fact that the topic attractiveness stems for a number of different reasons. First, health inequality may arise due to differences in health behavior or access to health services across groups which can be prevented or attenuated through public intervention (Gepkens and Gunning-Schepers, 1996; Cooper et al., 2002). Second, welfare measures solely based on economic variables, such as Gross Domestic Product, give an incomplete and potentially biased representation of the standard of living (Deaton and Paxson, 1998; Becker, Philipson and Soares, 2005; Steckel, 2008). Third, if education and income have a causal impact on mortality then education or redistribution policies can have an impact on health (Evans and Snyder, 2002; Munes-Llera, 2005). Finally, differential mortality have the potential of undoing the progressivity built in the Social Security benefit formula as well as in other entitlement programs such as Medicare (Congressional Budget Office, 2006; Bhattacharya and Lakdawalla, 2006).

In particular, the evolution of these differentials over time has drawn significant attention related to the important reasons mentioned. In a recent study, Waldron (2007) used data from Social Security records and finds evidence on a markedly increase differential mortality for the 65 and older male population in the 1972 to 2001 period.[1] Historically, researchers (e.g. Antonovsky, 1967) have found greater differentials are younger ages. Hence, it is of great interest to know what has happened with differential mortality for the non-retired population.

Four previous studies (Duleep, 1989; Feldman et al., 1989; Pappas et al. 1993; Preston and Elo, 1995) have provided fairly consistent evidence of increasing differential mortality by education from 1960 to the mid 1980s for men, though results for women are more mixed. It should be noted that the robustness of this result, however, is not as strong as one might first think. Even though each of these studies used a different data set to measure the post-1960 socioeconomic status mortality differential, each used as their base point the Kitagawa-Hauser (1973) study. If the Kitagawa-Hauser mortality differential by income were underestimated, all four studies would show a spurious increase in the differential.[2] Unlike these studies Schalick, et al. (2000) computed death rates for 1967 and 1986 using data from the National Mortality Followback Survey and the National Health Interview Survey and provided evidence of increasing differential mortality by income groups for individuals aged 35 to 64 years old.[3] Still, due to paucity of data, the question of how mortality differentials for the nonretired population have evolved in the last twenty years remains unanswered.

This paper aims to answer this question by providing differential mortality estimates by measures of lifetime earnings for the period 1983 to 2003. These measures

[1] Individuals in the study are classified by their average relative earnings while aged 45 to 55.

[2] In the Kitagawa-Hauser (1973) study, 1960 Census records where matched to death certificates. The estimated differentials in mortality may be biased due to a substantial rate of nonmatches (more than 20 percent in each race-sex group), and the possibility that the matching rate was different across socioeconomic characteristics.

[3] Nevertheless, there are problems in classifying individuals by one-year income groups as it will be discussed below.

are constructed using long averages of past earnings. For individuals older than 53, earnings from age 41 to 50 are used to capture the years when the person was closest attached to the labor market. For younger individuals, averages ranging from 5 to 10 years were computed without including the immediate preceding 3 years (e.g. for individuals aged 43, earnings from age 31 to 40 are used).

The study uses a unique data set constructed matching extensive demographics data from the 1984, 1993, 1996 and 2001 panels of the Survey of Income and Program Participation (SIPP) to earnings, benefits and mortality data from the Social Security Administration (SSA) and earnings data from the Internal Revenue Services (IRS). The earnings data is particularly rich as it contains the history of Social Security covered earnings as well as information for this variable from tax income returns from the Internal Revenue Services (IRS) for the period 1978 to 2003.[4] The resulting sample used for the analysis contains roughly 130,000 individuals aged 35 to 75 for which the mortality window ranges from 3 to 21 years yielding a total of approximately 1.2 million person-year observations. The uniqueness of this data set, as compared to other data sets used to study differential mortality in the United States, lies in its sample size, representativeness, time period covered, comparability across time and the fact that it includes high-quality longitudinal individual level earnings data.

As in this study, Waldron (2007) used Social Security data and hence shared some of these advantages (sample sizes, comparability across time and time period covered). However, that study counted with limited demographic information (no survey data) and only capped Social Security earnings for covered jobs (no IRS data). Furthermore, while

[4] Earnings data from IRS, as opposed to Social Security sources, is uncapped and includes earnings from both Social Security covered and uncovered jobs.

Waldron (2007) used Tobit regressions to impute above the Social Security taxable maximum, the current study used additional information in the data set to infer in which quarter the individual hit the taxable maximum for years before 1978. Using the quarters of coverage information to estimate earnings above the taxable maximum is vastly preferable to using other methods, based on strong assumptions, such as tobit, for estimating above the taxable maximum.[5]

Typically studies of trends in differential mortality have classified individuals by education (Duleep, 1989; Feldman et al., 1989; Pappas et al. 1993; Preston and Elo, 1995) or current income groups (Schalick et al., 2000). There are two problems associated with measuring differences in mortality rates in a given year across groups defined by income in the previous year. First, this strategy suffers from reverse causation: individuals who experience health shocks (which increase their mortality probability) may drop out from the labor market and simultaneously suffer a drop in income. As a result, this approach will overstate the true correlation between permanent income and mortality. Second, yearly income is a noisy measure of permanent income. Taking into account only this effect, we should expect that estimates of differential mortality by income in a specific year will underestimate the extent of differential mortality by permanent income.

This paper aims to tackle these two problems by classifying individuals using average lagged earnings and excluding from the average computation years immediately preceding when mortality is ascertained. In this way, the problem of reverse causation is at least partially addressed by distancing the measurement of earnings from the measurement of mortality. The problem of attenuation bias due to noisy yearly data is

[5] This procedure was first used by Duleep (1986).

tackled by computing long averages of yearly earnings.[6] Moreover, the choice of using measures of lifetime earnings instead of education as the classifying dimension is based in the following reasons. First, earnings data in this study is from high-quality administrative sources whereas the education information was reported by SIPP respondents. Given that it has been shown that there exists substantial error in reported education data (Black et al., 2003) changes in differential mortality by this variable could be due to changes in the extent of measurement error. Second, during the period under analysis, 1984 to 2003, there has been a significant increase in income inequality (Gottschalk and Danziger, 2005); hence it is interesting to know whether this increase in economic inequality has been accompanied by an increase in health inequality using the same underlying variable to classify individuals. Finally, results from this paper can be used as an input in studies of progressivity of public programs such as Social Security and Medicare which typically used lifetime earnings measures to classify individuals.

The breadth of the data set used in the study also allows one to contribute to the literature by providing estimates on the extent of differential mortality by race, ethnicity, education, marital status, disability status and particularly by lifetime earnings quintiles. Regarding the results by lifetime earnings, it contributes to a small but growing literature which is described next. Duleep (1986) matched Social Security earnings data to mortality records to predict the death probability in a five-year window (1973 to 1978) using a five-year average of earnings (1968 to 1972). Menchik (1993) used the National Longitudinal Survey of Mature Men and constructed a measured of average earnings up

[6] The idea of distancing the measurement of earnings from the measurement of mortality to tackle the problem of reverse causation was first introduced by Duleep (1980), who used the longitudinal Social Security earnings data to explore how the income-mortality relationship is affected as progressively earlier measures of earnings are related to mortality. Also, Duleep (1986) also explored the effect on the income-mortality relationship of using several years of earnings as opposed to a single year

to age 61 to use as a control while probing for the effect of poverty on mortality. McDonough et al. (1997) employed data from the Panel Study of Income Dynamics to construct ten-year panels in which income is averaged over the first five years and mortality status is ascertained over the subsequent five years. More recently, studies by Waldron (2007) and Duggan et al. (2007) have used large samples from Social Security Administration (SSA) records to provide very precise estimates of mortality differentials by lifetime earnings.

Findings regarding the extent of differential mortality by lifetime earnings can be summarized as follows. First, there are large differentials in age-adjusted mortality rates across individuals in different quintiles of the individual lifetime earnings distribution (e.g., men ages 35 to 49 in the bottom quintile have age-adjusted mortality rates 6.4 times larger than those in the top quintile). Second, controlling for race, Hispanic origin, marital status, and education only slightly reduces these differentials. Third, differentials for men are slightly larger for individual compared with household lifetime earnings, but the opposite is true for women.[7] Fourth, men and women have similar differentials when average household lifetime earnings are used to sort individuals into quintiles. Finally, differentials decrease markedly with age.

With respect to trends in differential mortality by lifetime earnings, there is substantive evidence pointing toward an increase in differential mortality in the period 1983 to 2003. For example, in the period 1983 to 1997 men ages 35 to 49 in the bottom lifetime earnings quintile had mortality 5.9 (1.8 for women) higher than those in the top quintile; in the period 1998 to 2003 this ratio increased to 8.3 (4.8 for women). This

[7] "Household lifetime earnings" refers to the average lifetime earnings of the individual and his or her spouse (if he or she is or was married).

increase in differential mortality is also found for all other age-sex groups, when sorting individuals by household earnings and even when using alternative measures of lifetime earnings.

It is worth noting that in the data set used, the quality of the earnings information increased over time; this increase in quality of information could bias the results toward finding increasing differential mortality. The increase in data quality is due to the fact that, for the period 1951 to 1977, only Social Security earnings are available.[8]

Taking these concerns into account, several robustness checks were performed to explore whether the uncovered pattern expresses a true phenomenon rather than a data artifact. First, as a way to tackle the problem of using years of noncovered employment trends in differential mortality were computed dropping from the calculation of the lifetime earnings measure years with zero earnings. Second, the analysis was repeated using a two-year average of earnings in ages A-3 and A-4 (where A is the person's age) in order to use earnings only from IRS sources. In both cases, the same patterns of increasing differential mortality were observed, giving support to the view of a real increase in mortality differentials across lifetime earnings groups in the last 20 years.

II. Data and Sample Construction[9]

This study uses data from the 1984, 1993, 1996, and 2001 panels of the Survey of Income and Program Participation (SIPP), matched to several files administered by the

---

[8] The coverage of Social Security increase markedly in the early 1950s but it slowed significantly starting in 1957; at that point about 80 percent of the total earnings in the economy corresponded to jobs covered by Social Security (Committee on Ways and Means, *2004 Green Book*). For this reason, this paper uses earnings only for the period 1957 to 2003.

[9] An Appendix containing more detailed information about the data used as well as the methods to construct the sample are available from the author upon request.

Social Security Administration (SSA) containing information on earnings, disability, and mortality. The SIPP provides information for a representative sample of the U.S. non-institutional population. It contains information about cash and noncash income, taxes, assets, liabilities, demographics, labor force status, and participation in government transfer programs.[10] Information from SSA files include information on: yearly Social Security taxable earnings for the period 1951 to 2003 (including a variable which can be used to proxy in which quarter the individual hit the taxable maximum), federal income taxable earnings fro the period 1978 to 2003, Social Security benefits, and month and year of death.

To create the sample for this study (the Mortality sample), I construct a panel data set in which the unit of observation is a person-year, containing basic demographic and economic variables from the SIPP. For time-varying variables (education, marital status, and spouse links), monthly information from the SIPP is used to construct yearly observations. Observations in which the person was 24 years of age or younger are dropped. Second, information on Social Security annual earnings from 1951 to 2003, federal income taxable earnings from 1978 to 2003, disability status, and year of death is attached to the sample. Third, the resulting intermediate data set is "aged" forward, completing years outside the SIPP window with information from the last year of available SIPP data up to the year 2003 (or up to the year of death if earlier).[11] Last, only observations for individuals ages 35 to 75, born in 1909 or later, are kept.[12] The resulting

[10] For more details on the SIPP, see www.bls.census.gov/sipp/index.html.
[11] That is, for an individual with SIPP data in 1984, 1985, and 1986, additional yearly observations for 1987 onward are created using the variable values from 1986.
[12] Individuals younger than 35 are dropped because it is necessary to observe their earnings at ages while they were potentially attached to the labor market to construct the measures of lifetime earnings. Those born before 1908 are dropped because there is no earnings data for ages 48 and younger for them. Finally,

data set is a panel data where the unit of observation is a person-year. It includes yearly observations for individuals since the year they first entered the SIPP until 2003 (or until their death year, if they died before 2003).[13]

To construct the measures of lifetime earnings used in the study the following steps are taken. First, total annual earnings for the years 1957 to 2003 are obtained. Second, measures of lifetime earnings are constructed using five- to ten-year averages of past indexed earnings. Last, quintiles of lifetime earnings within sex, five-year age, and five-year cohort groups are computed. Each of these steps is described next.

For the period 1957 to 1977, Social Security taxable earnings from the SER are used. For individuals with capped earnings, a procedure was followed to impute earnings above the taxable maximum which used information about in which quarter the individual hit the taxable maximum. For years 1980 to 2003, annual earnings were computed using earnings data from federal income tax sources (IRS). For years 1978 to 1979, due to presumably low quality of the IRS data, total earnings variable was set, for individuals with capped earnings, as the weighted average between total earnings in 1977 and 1980. All earnings were indexed using the Personal Consumption Expenditure deflator).

When constructing this measure, the goal was to approximate the permanent earnings level of the individual while he or she had the closest attachment to the labor market. Also, in order to mitigate the problem of reverse causation, the measure does not include earnings received in the three years preceding when mortality was ascertained.

---

given the cohort restriction imposed in the sample, individuals older than 75 are eliminated from the sample to ensure that the sample contains individuals in the same age range across time.

[13] Sample statistics are presented in subsection II.3.

Taking these issues into consideration, for individuals ages 53 and older, the permanent earnings measure was constructed as the ten-year average earnings from ages 41 to 50. For younger individuals of age *A*, the measure is constructed by averaging earnings between age *A*-3 and the maximum of 28 and *A*-12. For example, for an individual age 35, earnings between ages 28 and 32 will be used; for a 45-year-old person, the measure averages earnings between ages 33 and 42.

Finally, to avoid interactions between earnings levels and sex, age and cohort, I sort all individuals alive in a year into quintiles of the lifetime earnings distribution computed *within* sex, five-year age and five-year cohort groups. Thus, results in the study show differences in mortality rates by lifetime earnings when individuals are compared with others of the same sex and similar age and year of birth.

The Mortality sample constructed for this study constitutes a unique data set for exploring the relationship between lifetime earnings and mortality. However, the way that it was constructed (pooling SIPP panels, matching them to SSA records and filling years forward) may raise doubts about the representativeness of the sample. The question is whether the results are representative of a certain period of time (the period of time when mortality was ascertained). Because individuals in the Mortality data set enter the sample when they first are interviewed in the SIPP and remain in the sample until the year 2003 (or until they die), the sample contains observations for the years 1983 to 2003, but its composition is tilted toward later years.

To tackle this problem, and to make the sample representative for the period 1983 to 2003, population counts by age, sex, race, Hispanic origin, and year were obtained from U.S. Census intercensal estimates.[14] The same age restriction used for constructing the Mortality sample was applied to the Census data (only individuals ages 35 to 75 were kept). Next, weights were constructed to match to the Census data, the distribution of observations in the Mortality sample by sex, five-year age, race, Hispanic origin, and five-calendar-year groups. Table 1 shows that the age, sex, and race distributions in the unweighted Mortality sample are similar to the Census counterparts. However, the distributions by year are quite different Finally, comparing Columns 3 and 4, we see that when the Mortality sample is reweighted, the distributions by age, sex, race, Hispanic origin, and year match closely the distributions in the Census data. Therefore, the constructed weights are used for all results presented in the remainder of the paper.[15]

III. The Extent of Differential Mortality by Lifetime Earnings

This section presents estimates of the extent of differential mortality by lifetime earnings. In the first subsection, mortality ratios are reported for groups defined by race, Hispanic origin, education, marital status, disability status, and lifetime earnings quintiles. The ratios, computed separately for sex, represent the relationship between the mortality rates for each group (compared with the whole population) after the rates have been adjusted for differences in the age distribution between the group and the whole population. The second subsection focuses on differences in mortality rates by lifetime earnings, using logistic regressions to adjust for different sets of covariates. The final subsection presents robustness checks of the main findings in this section.

[14] The Census estimated counts were obtained at http://www.census.gov/popest/estimates.php.

[15] Additionally, the distribution by other covariates of the Mortality sample is quite similar to the SIPP distribution for different years across the 1983 to 2003 period. Also, to gauge the quality of the mortality data, death rates by age and sex were compared to those from the Human Mortality Database and found to be very similar.

III.1. Mortality Ratios

The mortality ratio for a certain subpopulation in certain age group (e.g. black men ages 35 to 49) is computed in the following way:

$$Mortality\ Ratio_{BLACK\ MEN} = \frac{\sum_{a=35,\ldots,49} weight_a * mortality\ rates\ of\ black\ men_a}{\sum_{a=35,\ldots,49} weight_a * mortality\ rates\ of\ all\ men_a}$$

where *mortality rates of black men* $_a$ is the one-year age-specific mortality rate for black men age a, *mortality rates of all men*$_a$ is the one-year age-specific mortality rate for all men age a and $weight_a$ corresponds to the fraction of men age a from all men in this age group in the sample.

The numerator is the age-adjusted one-year mortality rate for black men and the denominator is the average mortality rate for all men in the sample. A ratio of 1 for a certain group indicates that, once we adjust for differences in the age distribution, the group has the same mortality rate as all individuals in that age-sex group. A ratio higher than 1 (e.g., 1.5) means that the group has a higher age-adjusted mortality rate than all individuals of the same sex in the sample (50 percent higher).

Table 2 presents mortality ratios for men by age groups. Ratios by race, Hispanic origin, education, marital status, and Social Security Disability Insurance (DI) status replicate the general patterns documented in previous studies on differential mortality. Focusing on individuals of ages 35 to 75 we see that blacks have a 48 percent higher age-adjusted mortality rate (compared with all men), Hispanics a 6 percent lower rate, and college graduates have a 38 percent lower mortality rate.[16] Being never married, separated/divorced, or widowed is associated with a 51 percent to 57 percent higher mortality rate. Individuals who have ever received DI have a 270 percent larger mortality risk.[17] Comparing Columns 3 to 5 of Table 2, we see evidence of the well documented pattern of mortality differentials decreasing with age (though more slowly for differentials by education).

The bottom panel of Table 2 presents mortality ratios by lifetime earnings quintiles computed within sex, five-year age, and five-year cohort groups. For quintiles computed by individual or household lifetime earnings we observe similar patterns, although the gradient is slightly stronger when using individual lifetime earnings. Overall, there is a strong relationship between these measures of lifetime earnings and mortality. Individuals ages 35 to 49 in the bottom lifetime earnings quintile have a 125 percent higher mortality rate, while those in the top have a 65 percent lower rate. The decrease in mortality differentials by age group is very strong. The ratio of age-adjusted mortality rates of bottom to top lifetime earnings quintiles for men ages 35 to 49 is 6.4 (2.25/0.35), it drops to 2.7 for men ages 50 to 64 and to only 1.5 for men ages 65 to 75.[18]

Table 3 presents mortality ratios for women. Overall, the patterns of differential mortality by race, Hispanic origin, education, marital, and DI status found for men are also present for women except for certain differences. First, Hispanic women have adjusted mortality rates that are generally closer to those of all women compared to mortality differences between Hispanic men and all men. Second, mortality rate

[16] In this subsection, for brevity, mortality rates refer to age-adjusted mortality rates.
[17] Mortality ratios for individuals currently on DI are not computed for age groups 35 to 75 and 65 to 75 because individuals on DI have their status updated to Social Security retirees when they turn 65.
[18] Still, given that overall death rates increase very rapidly with age, decreasing relative mortality rates with age can be accompanied by increases in the differences in absolute mortality rates.

differences for all women across marital status are less pronounced than those for men especially for women ages 65 to 75. Third, the mortality "penalty" for being on DI or having ever been on this program is higher than for men but still the patterns are similar.

To compare estimates of differential mortality by lifetime earnings between men and women, we can focus on the bottom panels of Tables 3 and 4. Although the gradient is steeper for men than for women when using individual lifetime earnings, it is strikingly similar when using household lifetime earnings.[19] The former result should be expected, given the higher attachment to the labor market for men (which suggests that men are relatively better "sorted" when individual lifetime earnings are used). However, the latter result is an interesting finding that deserves further exploration in future work.

III.2. Logistic Results

In this subsection, differences in mortality risks are estimated using a discrete-time logistic model, a particular type of a survival model. In the analysis, mortality is measured at the year level and then, there are a large number of ties in the data (i.e. two or more events showing in the data as happening in the same point in time). Allison (1995) showed that the discrete-time logistic model is equivalent to the discrete-time proportional odds model proposed by Cox (1972) when there are many ties in the data. The discrete-time logistic model allows the estimation of the effect of time-varying covariates on the hazard, requirement which is essential for this study. Finally, the fact that the data set includes multiple observations (years) for an individual does not make necessary to apply some type of clustering adjustment because the independence of

[19] For example, the ratio of age-adjusted mortality rates for the bottom to top quintiles of individual lifetime earnings is just 2.4 for women ages 35 to 49 compared with 6.4 for men in that age group. The analogous ratios using household earnings are 4.0 and 5.2 for women and men, respectively.

observations still holds because of the factoring of the likelihood function allows each term to be treated as independent (Allison, 1995).[20]

Logistic models are run where the dependent variable is an indicator that equals 1 if the individual died in the next year, and the key independent variable is the quintile of individual lifetime earnings to which the individual is assigned. Odds ratios are estimated relative to individuals in the bottom quintile.

As expected, given that lifetime earnings quintiles are computed within sex, five-year age and five-year cohort groups, results from running models with no covariates are very similar to those when age and cohort are added linearly or as single year dummies as controls. Given this, Figure 1 presents odds ratios from specifications with just three sets of controls: a) age and cohort, b) age, cohort, race, and marital status, c) age, cohort, race, marital status, and education. Results with age, cohort, and race are not presented because they are very similar to those when only age and cohort are added as controls.

All the patterns are robust when race, marital status, and education controls are added. Figure 1 shows that for men the degree of differential mortality by lifetime earnings slightly decreases when we control for these factors, but for women it slightly increases (when adding only race and marital status) or remains virtually unchanged (when adding all the mentioned controls).[21]

It is difficult to compare these results to those from previous studies that used cross-section income measures instead of multi-year averages, because of differences in income concept used (earnings from employment versus income from all sources, individual earnings versus household earnings, and earnings in categories of levels versus

[20] This condition would be violated if one individual could have more than one event; still, this is not the case in this analysis because the event under consideration is death.
[21] Complete regression results are available from the author upon request.

quintiles), age groupings, and time periods used. However, it is interesting to note that although in this study mortality differentials are only slightly affected when adjusting for other covariates, in the study by Sorlie et al. (1995), which used income in a year data from the Current Population survey matched to National Death Index records, differentials were significantly reduced when adjusting for covariates. For example, for men ages 45 to 64, the mortality ratio between those in the top and bottom income groups (more than $50,000 and less than $5,000, respectively) was 0.32 when no covariates were used and was 0.66 when covariates were added to the model.[22]

Finally, I explore how the findings on differential mortality are robust when changes are made in the way the lifetime earnings measure is constructed. In particular, results obtained using two alternative measures are presented. In the first alternative, only years with positive earnings are included in the computation of the average to check whether years with zero earnings due to data problems (between 1957 to 1977 earnings from jobs noncovered by Social Security are not included) or to temporary withdrawals from the labor market affect the results. In the second alternative, zero earnings years are included but the six years prior to when mortality is ascertained are excluded to better handle the problem of reverse causation.[23] In general, previous results are quite robust to these alternative measures of lifetime earnings.[24]

IV. Trends in Differential Mortality by Lifetime Earnings

[22] Still, the set of covariates added was not identical in both studies; employment status was added as a covariate in Sorlie et al. (1995) but not in the current study.
[23] For example, for an individual age 40, earnings between ages 28 and 34 are used (for a 50-year old, ages 38 to 44 are averaged).
[24] Full results are available from the author upon request.

IV.1. Main Results

This section presents evidence about changes in differential mortality by lifetime earnings for the period 1983 to 2003. For each of the six age-sex groups used in the study, observations are divided into two groups defined by time period: 1983 to 1997 (Early sample) and 1998 to 2003 (Late sample). The cut-off year to define the groups was selected to create two samples of roughly the same size

Figure 2 presents logistic estimates of the one-year probability of dying by individual lifetime earnings quintiles for the Early sample (solid line) and the Late sample (dotted line). The graphs present substantive and consistent evidence of increasing differential mortality for all age-sex groups. For a better sense of the changes in magnitude, Table 4 presents estimates of the mortality odds ratios of the top quintile relative to the bottom for the six age-sex groups. A decrease in these ratios expresses that mortality rates for the top quintile has decreased faster than the bottom quintile widening the mortality inequality across these groups. The top panel presents results when individuals are sorted using own earnings and the bottom panel when household earnings is used as the classifier.

For men, the drop in the top-to-bottom ratio is highest for those in the 50 to 64 age group as it decreases from 0.47 to 0.21 (a 56% reduction). Though the estimates for the other two age groups are not statistically significantly different across time periods, still the magnitude of the drop in this ratio is substantial (29% for individuals aged 35 to 49 and 12% for those aged 65 to 75).[25] For women, the drop in the ratio is statistically significant and substantial for those younger than 65. The drop for women in the 35 to 49

[25] A more dramatic way to illustrate the increase in the differential for men aged 35 to 49 is to pose the changes in terms of the increase in the mortality ratio of bottom to top quintiles: from 5.8 to an astounding 8.3.

age group is 62% whereas for those aged 50 to 64 a 43%. Results when using household lifetime earnings as the classifier, presented in the bottom panel, also point towards increases in differential mortality across the board. Taken together, these estimates provide evidence of a substantial increase in the mortality inequality across sexes and age groups for the time period analyzed.

The evidence of widening mortality when individuals are sorted into quintiles of the lifetime earnings distribution may reflect an increase in the dispersion of the distribution of lifetime earnings (with a constant relationship between earnings and mortality) or, alternatively, an increase in the slope of the earnings-mortality gradient. To shed some light on this issue, Table 5 presents average lifetime earnings by age group, sex, and quintiles for the Early and Late samples. The table shows that, although the distribution of lifetime earnings for women has become more dispersed, the distribution for men has remained quite stable for those ages 35 to 64 (though top earners have gained in this period) and has widened for those ages 65 to 75. Given that differentials have increased for all age-sex groups and that the increase has not been limited to the mortality ratios of the top to bottom quintiles, it seems that the mortality-earnings gradient is becoming steeper.

IV.2. Robustness Checks

This subsection presents several robustness checks to gauge the reliability of the evidence found on increasing differential mortality in the last 20 years. The basic motivation for this exercise stems from the fact that the quality of the earnings data is increasing over time and, as noted before, this can create an artificial increase in the correlation between earnings and mortality.

The results of the robustness checks are presented in Table 6. The top panel of this table presents odds mortality ratios of the top quintile relative to the bottom by age group, time period and alternative average lagged earnings measures. The first line replicates results from Table 4 (i.e. results obtained using the basic measure of individual lifetime earnings). The second line presents results when including only positive earnings years; the third line shows ratios when at least the six years prior to the mortality window are excluded. The fourth line presents results when averaging the third and fourth year before when mortality is ascertained (e.g., for a person age 50, earnings for ages 46 and 47 are averaged). The advantage of this last measure is that it uses only earnings data from IRS sources (uncapped and including noncovered Social Security jobs).

Comparing columns 2 to 3, 4 to 5 and 6 to 7, we see that for all earnings measures there is evidence of increasing differential mortality (that is, the ratio of the top to bottom quintile mortality is decreasing over time). Similarly, the bottom panel of Table 6 shows that the evidence of increasing differential mortality by lifetime earnings is also robust across alternative measures of lifetime earnings for women.

V. Conclusions

This paper estimates the extent and trends of differential mortality by lifetime earnings using a very large panel data set containing information on mortality, earnings history, and demographic and economic characteristics. Measures of lifetime earnings are constructed to deal with the problems of reverse causation and noise in yearly earnings data present in estimates of differential mortality by previous year income. Summarizing the results, the study found a strong negative relationship between one-year mortality and lifetime earnings, robust when controlling for usual covariates, weaker for women than for men, and decreasing with age. Most importantly, the evidence in the paper points towards a substantial increase in differential mortality by lifetime earnings in the period 1983 to 2003, especially for the nonretired population.

When interpreting these results it is important to note the following caveat. Given the post-1964 expansion of transfer programs, a reasonable supposition is that such programs siphoned off from the labor force chronically ill persons with a higher than average probability of death. Such a phenomenon would increase the inverse relationship between earnings and mortality. Still, during the period under analysis, the prevalence of individuals on the Social Security Disability Insurance (DI) program did not increase dramatically (from 1980 to 2000, it increased from 2.1% to 3.0% of the 20 to 64 population according to figures from the Trustees Report, 2005). Moreover, if the underlying explanation for the evidence on increasing differential mortality is due to the increase in the DI enrollment, we should observe a significant attenuation of the rise in differentials when using the alternative lifetime earnings measures presented above. However, Table 6 shows that the increase in differentials is as strong for the baseline measure as with the alternative ones.

These findings raise concerns by themselves but even more when they are placed in the context of previous results in the literature of differential mortality in the United States. Waldron (2007) found compelling evidence of increases in differential mortality for men aged 65 and older using mortality data covering the period 1972 to 2001. Four previous studies (Duleep, 1989; Feldman et al., 1989; Pappas et al. 1993; Preston and Elo, 1995) found quite consistent evidence of increasing differential mortality for the period 1960 to the mid 1980s for men though for women they were somewhat mixed. Taken together, these results suggest an increase in differential mortality for almost the last five decades.

The evidence on widening inequality in the United States mirrors similar results reported in studies for western European countries. In particular, Mackenbach et al. (2003), using data from national longitudinal sources in six Western European countries (Finland, Sweden, Norway, Denmark, England/Wales, and Italy), found that relative inequalities in total mortality have increased in these countries between the early 1980s to the early 1990s. Similarly, Bronnum-Hansen and Baadsgaard (2007) explored trends in differential mortality by education in Denmark in the period 1981 to 2005 and found that social inequality in life expectancy has widened during the period.

The empirical results presented in this study on increasing differential mortality raise a number of important questions —in particular: What are the causes and consequences of increasing differential mortality by lifetime earnings? With respect to causes, the explanations that have been put forward to explain differential mortality can be used to check whether they can explain the rise in this correlation. In particular, a potential explanation for increasing differential mortality by lifetime earnings could be that the correlation between poor lifestyle habits (such as smoking, poor diet, and lack of exercise) and low lifetime earnings have increased over time.[26] Another explanation

[26] Evidence towards this hypothesis may be mixed. Zhang and Wang (2004) reported that for most demographic groups the relationship between body mass index and socio-economic status has weakened in the 1971 to 2000 period. Kant and Graubard found no evidence of an increase in the relationship between

could be that recent advances in medical treatments are more readily available to high earners than to low earners compared to the past.

Increasing differential mortality also has implications for the long-term budget outlook. First, if the “life-expectancy premium” for high-earners is increasing over time, this may worsen the budgetary pressures facing the U.S. Social Security system, given that high earners receiving larger benefits will collect them (on average) for a longer period of time (Diamond and Orszag, 2004) . Second, studies that established the progressivity of Social Security have used historical data on the correlation between earnings and mortality in order to account for the effect of differential mortality on progressivity measures. However, if differential mortality by lifetime earnings continues to increase over time, then we should expect that, holding other factors constant, the progressivity of the system will diminish.

The increase in income inequality in the United States during the last quarter of the 20th century has attracted a great deal of attention, both in the academic literature and in policy circles. However, a subject of equal or greater concern should be that the country is also becoming increasingly unequal in terms of health.

---

certain healthier diet profiles and socio-economic characteristics. Finally, the association between smoking and education has become significantly stronger 1965 to 1993 period (Garfinkel, 1997).

References

Allison PD. Survival analysis using the SAS system: A practical guide. Cary, NC: SAS Institute; 1995.

Antonovsky A. Social class, life expectancy and overall mortality. The Milbank Memorial Fund Quarterly 1967;45; 31-73.

Becker GS, Philipson TJ, Soares RR. The quantity and quality of life and the evolution of world inequality. American Economic Review 2005;95(1); 277-291.

Bhattacharya J, Lakdawalla D. Does Medicare benefit the poor? Journal of Public Economics 2006;90; 277-292.

Black D, Sanders S, Taylor L. Measurement of higher education in the census and current population survey. Journal of the American Statistical Association 2003;98(463); 545-554.

Bronnum-Hansen H, Baadsgaard M. Increasing social inequality in life expectancy in Denmark. European Journal of Public Health 2007;17(6); 585-586.

Committee on Ways and Means U.S. House of Representatives. 2004 green book. WMCP: 108-6. Washington, D.C.; 2004.

Congressional Budget Office. Is Social Security progressive? CBO Economic and Budget Issue Brief. Washington, D.C.; 2006.

Cox DR. Regression models and life-tables. Journal of the Royal Statistical Society (Series B) 1972;34(2); 187-220.

Deaton AS, Paxson CH. Aging and inequality in income and health. American Economic Review 1998;88(2); 248-53.

Diamond P, Orszag P. Saving Social Security: a balanced approach. Washington, D.C.: Brookings Institution Press; 2004.

Duggan J, Gillingham R, Greenlees J. Mortality and lifetime income: Evidence from the U.S. Social Security records. Working Paper number 07/15. International Monetary Fund, Washington, D.C.; 2007.

Duleep HO. Measuring the effect of income on mortality using longitudinal administrative record data. Journal of Human Resources 1986;21(2); 238-51.

Duleep HO. Measuring socioeconomic differentials over time. Demography 1989;26(2); 345-51.

Evans WN, Snyder SE. The impact of income on mortality: evidence from the social security notch. Working Paper number 9197. National Bureau of Economic Research, Cambridge, M.A.; 2002.

Feldman JJ, Makuc DM, Kleinman JC, Cornoni-Huntley J. National trends in educational differentials in mortality. American Journal of Epidemiology 1989;129(5); 919-33.

Garfinkel L. Trends in cigarette smoking in the United States. Preventive Medicine 1997;26(4); 447-450.

Gepkens A and Gunning-Schepers L. Interventions to reduce socioeconomic health differences. The European Journal of Public Health 1996;6(3); 218-226.

Gottschalk P, Danziger S. Inequality of wage rates, earnings and family income in the United States, 1975-2002. Review of Income and Wealth 2005;51(2); 231-254.

Human Mortality Database. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at www.mortality.org. Accessed November 28, 2006.

Kant AK, Graubard BI. Secular trends in the association of socio-economic position with self-reported dietary attributes and biomarkers in the US population: National Health and Nutrition Examination Survey (NHANES) 1971-1975 to NHANES 1999-2002. Public Health Nutrition 2007;10(2); 158-167.

Kitagawa E, Hauser P. Differential mortality in the United States: A study in socioeconomic epidemiology. Cambridge, MA: Harvard University Press; 1973.

Lleras-Muney A. The relationship between education and adult mortality in the United States. Review of Economic Studies 2005;72(1); 189-221.

Mackenbach JP, Bos V, Andersen O, Cardano M, Costa G, Harding S, Reid A, Hemstrom O, Valkonen T, Kunst AE. Widening socioeconomic inequalities in mortality in six Western European countries. International Journal of Epidemiology 2003;32(5); 830-837.

McDonough P, Duncan GJ, Williams D, House J. Income dynamics and adult mortality in the United States, 1972 through 1989. American Journal of Public Health 1997;87(9); 1476-83.

Menchik PL. Economic status as a determinant of mortality among black and white older men: Does poverty kill? Population Studies 1993;47(3); 427-36.

Pappas G, Queen S, Hadden W, Fisher G. The increasing disparity in mortality between socioeconomic groups in the United States, 1960 and 1986. New England Journal of Medicine 1993;329(2); 103-9.

Preston SH, Elo IT. Are educational differentials in adult mortality increasing in the United States? Journal of Aging and Health 1995;7(4); 476-96.

Schalick LM, Hadden WC, Pamuk E, Navarro V, Pappas G. The widening gap in death rates among income groups in the United States from 1967 to 1986. International Journal of Health Services 2000;30(1); 13-26.

Sorlie PD, Backlund E, Keller JB. US mortality by economic, demographic and social characteristics: The National Longitudinal Mortality Study. American Journal of Public Health 1995;85(7); 949-56.

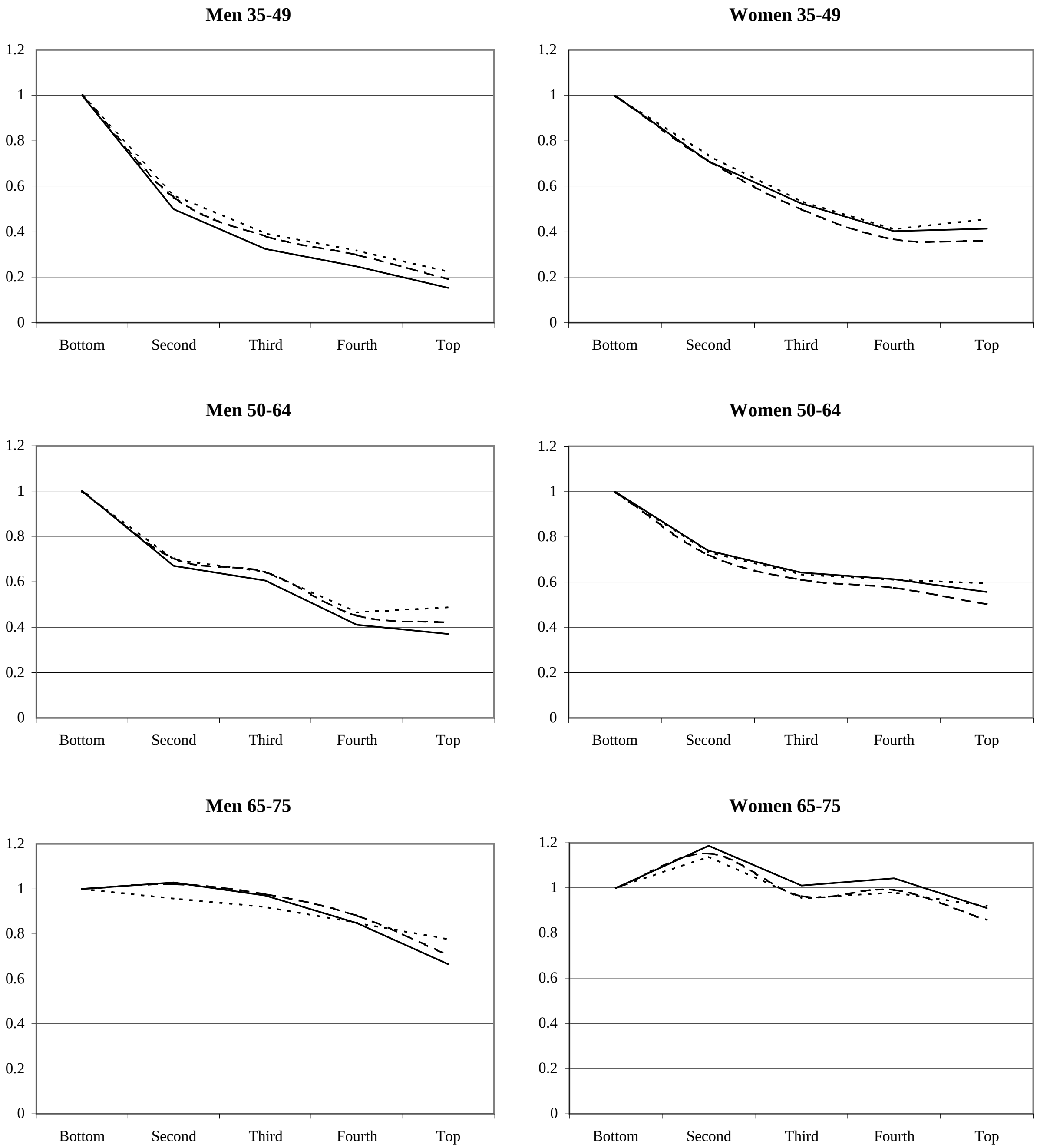
Steckel RH. Biological measures of the standard of living. Journal of Economic Perspectives 2008;22(1); 129-152.

Social Security Administration Board of Trustees. The 2005 annual report of the board of trustees of the federal old-age and survivors insurance and disability insurance trust funds. Washington, D.C.; 2005. Available at www.ssa.gov/OACT/TR/TR05/tr05.pdf.

Waldron H. Trends in mortality differentials and life expectancy for male Social Security-covered workers, by average relative earnings. Working Paper number 108. Office of Research, Evaluation, and Statistics, Office of Policy, Social Security Administration, Washington, D.C.; 2007.

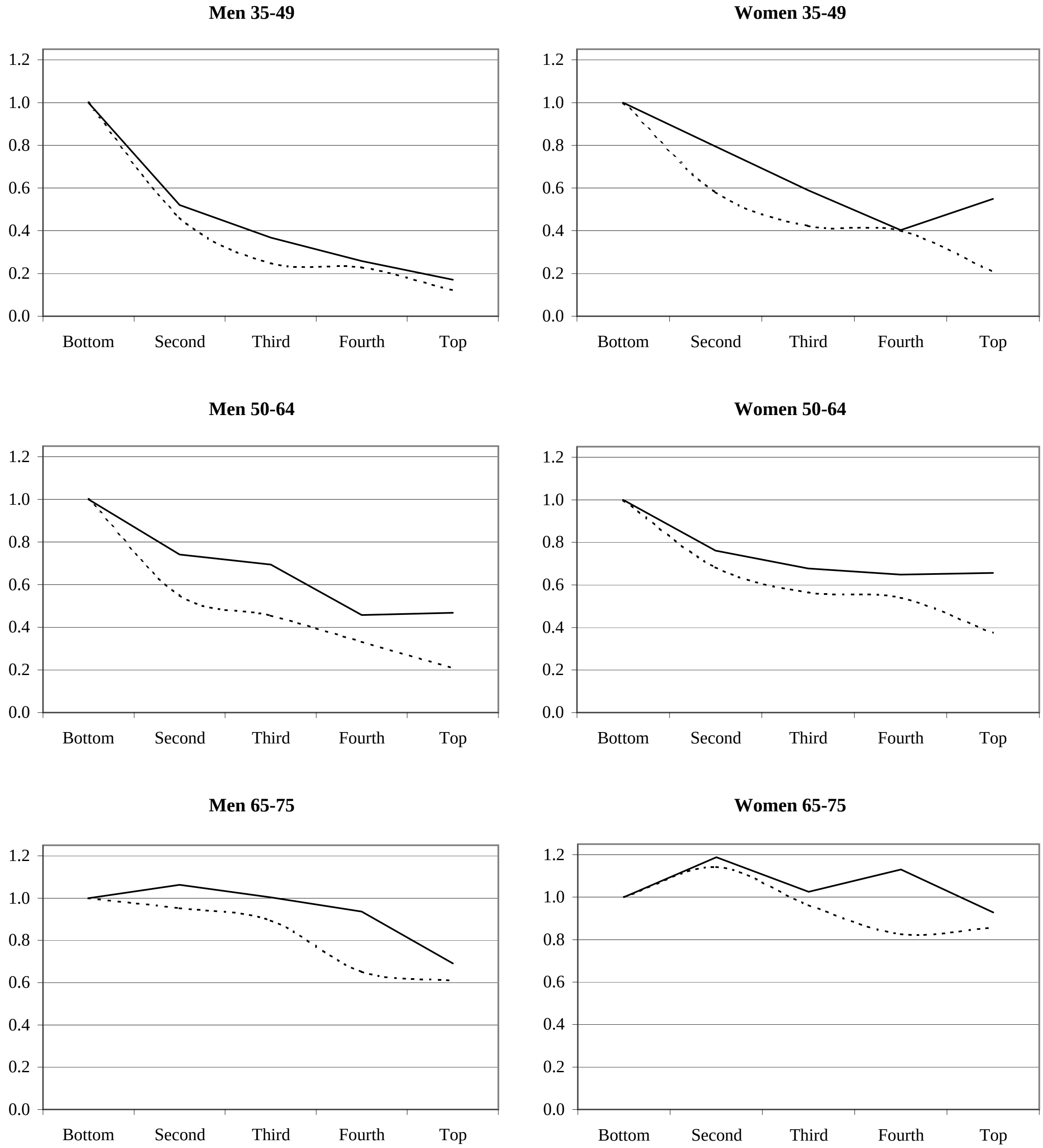
Zhang Q, Wang YF. Trends in the association between obesity and socioeconomic status in US adults: 1971 to 2000. Obesity Research 2004;12(10); 1622-1632.

Figure 1. Adjusted Odds Ratios of One-Year Mortality by Individual Lifetime Earnings Quintiles



Note: Adjusted odds ratios are obtained from logistic regressions of one-year mortality indicators on individual lifetime earnings quintiles adjusting for age and birth year (solid line); age, birth year, race, and marital status (dashed line); and age, birth year, race, marital status, and education (dotted line).

Figure 2. Adjusted Odds Ratios of One-Year Mortality by Individual Lifetime Earnings Quintiles
1983-1997 versus 1998-2003



Note: Adjusted odds ratios for the period 1983 to 1997 (solid line) and the period 1998 to 2003 (dotted line) are obtained from logistic regressions of one-year mortality indicators on individual lifetime earnings quintiles, adjusting for age and birth year.

Table 1. Comparison of the Mortality Sample's Descriptive Statistics and Census Data

	Mortality Sample		
	Unweighted	Weighted	Census
% Male	47.4	47.8	47.7
Average Age	51.5	51.4	51.4
% Age 35-49	49.3	49.9	49.9
% Age 50-64	33.9	32.7	32.7
% Age 65-75	16.8	17.4	17.4
Race and Ethnicity			
% White	86.3	85.8	85.7
% Black	10.2	10.5	10.5
% Other Race	3.5	3.7	3.8
% Hispanic	6.7	7.5	7.4
Observation year			
Average	1996.6	1993.7	1993.7
% Year 1983-1987	9.2	20.3	20.3
% Year 1988-1992	11.5	22.2	22.3
% Year 1993-1997	27.4	24.7	24.6
% Year 1998-2002	42.5	27.1	27.1
% Year 2003	9.4	5.7	5.7

Note: Census male, age, and race statistics correspond to average yearly statistics weighted by population counts for each year in the period 1983 to 2003. The weights used in the Mortality Sample Weighted column were constructed to match the sample distribution by sex, five-year age group, race, Hispanic origin, and five-calendar-year group to the U.S. Census counts in the period 1983 to 2003.

Table 2. Mortality Ratios, Men

Age group		35-75	35-49	50-64	65-75
All		1.00	1.00	1.00	1.00
Race and Ethnicity					
White		0.96	0.90	0.95	0.98
Black		1.48	1.74	1.58	1.35
Other Race		0.82	1.13	0.79	0.76
Hispanic		0.94	0.98	0.93	0.93
Education					
Less than High School		1.32	1.56	1.36	1.23
High School		1.02	1.11	1.05	0.98
Some College		0.91	0.97	0.89	0.90
College		0.62	0.55	0.64	0.62
Marital Status					
Never Married		1.57	1.95	1.66	1.42
Married		0.86	0.72	0.85	0.90
Separated/Divorced		1.51	1.56	1.46	1.53
Widowed		1.53	1.53	1.93	1.26
Disability Insurance					
Currently on DI		-	8.22	12.90	-
Ever on DI		3.69	8.24	4.18	2.15
Lifetime Earnings Quintiles					
Top	Own	0.64	0.35	0.61	0.74
	Household	0.77	0.40	0.73	0.90
Fourth	Own	0.80	0.56	0.68	0.94
	Household	0.85	0.54	0.82	0.96
Third	Own	1.00	0.73	0.99	1.08
	Household	0.90	0.83	0.79	0.99
Second	Own	1.12	1.13	1.10	1.14
	Household	1.07	1.16	1.07	1.05
Bottom	Own	1.44	2.25	1.63	1.10
	Household	1.41	2.07	1.60	1.10

Note: The mortality ratio for a group is computed by dividing the weigthed average of the one-year age-specific mortality rate for the group, where the weights correspond to the fraction of men in the sample in that age, by the male mortality rate in the sample. DI corresponds to Social Security Disability Insurance. Mortality ratios for individuals currently on DI are not computed for age groups 35 to 75 and 65 to 75 because individuals on DI have their status updated to Social Security retirees when they turn 65.

Table 3. Mortality Ratios, Women

Age group		35-75	35-49	50-64	65-75
All		1.00	1.00	1.00	1.00
Race and Ethnicity					
White		0.95	0.93	0.93	0.96
Black		1.48	1.53	1.58	1.42
Other Race		0.92	0.89	1.01	0.88
Hispanic		1.03	0.92	0.99	1.07
Education					
Less than High School		1.37	1.61	1.48	1.26
High School		0.93	1.12	0.89	0.91
Some College		0.81	0.78	0.82	0.81
College		0.65	0.58	0.64	0.68
Marital Status					
Never Married		1.39	1.92	1.60	1.16
Married		0.81	0.75	0.81	0.83
Separated/Divorced		1.29	1.35	1.32	1.26
Widowed		1.29	1.53	1.44	1.16
Disability Insurance					
Currently on DI		-	10.12	16.24	-
Ever on DI		4.10	10.86	4.54	2.51
Lifetime Earnings Quintiles					
Top	Individual	0.84	0.68	0.79	0.90
	Household	0.74	0.49	0.71	0.81
Fourth	Individual	0.93	0.66	0.86	1.03
	Household	0.87	0.75	0.76	0.96
Third	Individual	0.95	0.86	0.90	1.00
	Household	1.00	0.79	0.92	1.09
Second	Individual	1.09	1.17	1.03	1.11
	Household	1.02	1.04	1.09	0.99
Bottom	Individual	1.21	1.65	1.41	1.01
	Household	1.36	1.96	1.53	1.15

Note: The mortality ratio for a group is computed by dividing the weigthed average of the one-year age-specific mortality rate for the group, where the weights correspond to the fraction of women in the sample in that age, by the female mortality rate in the sample. DI corresponds to Social Security Disability Insurance. Mortality ratios for individuals currently on DI are not computed for age groups 35 to 75 and 65 to 75 because individuals on DI have their status updated to Social Security retirees when they turn 65.

Table 4. Trends in Differential Mortality by Lifetime Earnings.
Estimated Odds Ratios of One-Year Mortality. Top Relative to Bottom Quintile

Own Earnings

		Odds Ratios		Reduction in Ratio	
		1983-1997	1998-2003	Percentage point	Percent
Men	35-49	0.17	0.12	0.05	29%
	50-64	0.47	0.21 **	0.26	56%
	65-75	0.69	0.61	0.08	12%
Women	35-49	0.55	0.21 **	0.34	62%
	50-64	0.66	0.38 **	0.28	43%
	65-75	0.93	0.86	0.07	8%

* Significantly different from the 1983-1997 period estimates at the 5 percent level.
** Significantly different from the 1983-1997 period estimates at the 1 percent level.

Household Earnings

		Odds Ratios		Reduction in Ratio	
		1983-1997	1998-2003	Percentage point	Percent
Men	35-49	0.23	0.14	0.09	40%
	50-64	0.57	0.26 **	0.26	45%
	65-75	0.88	0.66 *	0.22	25%
Women	35-49	0.32	0.13 *	0.19	59%
	50-64	0.60	0.23 **	0.37	61%
	65-75	0.73	0.65	0.08	11%

* Significantly different from the 1983-1997 period estimates at the 5 percent level.
** Significantly different from the 1983-1997 period estimates at the 1 percent level.

Table 5. Average Individual Lifetime Earnings by Sex, Age Group, Period, and Lifetime Earnings Quintiles, 1983-1997 versus 1998-2003

	Men					
	35-49		50-64		65-75	
	1983-1997	1998-2003	1983-1997	1998-2003	1983-1997	1998-2003
Average Individual Lifetime Earnings						
Top	76,216	85,207	92,722	110,112	69,695	83,795
Fourth	44,281	44,073	49,913	53,633	39,313	46,862
Third	31,913	30,653	36,367	37,636	29,893	34,852
Second	20,349	19,079	22,373	22,796	16,201	20,325
Bottom	5,940	5,225	5,378	5,706	1,793	3,303

	Women					
	35-49		50-64		65-75	
	1983-1997	1998-2003	1983-1997	1998-2003	1983-1997	1998-2003
Average Individual Lifetime Earnings						
Top	40,033	49,733	35,777	48,327	25,461	30,634
Fourth	20,641	24,794	17,769	24,761	10,809	14,152
Third	11,082	14,259	8,672	13,949	3,352	5,819
Second	4,109	6,057	2,707	5,536	457	1,169
Bottom	372	746	106	425	0	3

Note: Quintiles of individual lifetime earnings in a certain year are computed within five-year age, sex, and five-year cohorts for individuals alive that year.

Table 6. Trends in Differential Mortality by Individual Lifetime Earnings
Estimated Odds Ratios of One-Year Mortality. Top Relative to Bottom Quintile
Using Alternative Average Lagged-Earnings Measures

	Men					
	35-49		50-64		65-75	
	1983-1997	1998-2003	1983-1997	1998-2003	1983-1997	1998-2003
Alternative Average Lagged-Earnings Measures						
Basic measure (includes zero earnings years and excludes at least three years prior to mortality window)	0.17	0.12	0.47	0.21	0.69	0.61
Basic measure but including only positive earnings years	0.20	0.14	0.45	0.24	0.59	0.50
Basic measure but excluding at least six years prior to mortality window	0.23	0.12	0.49	0.22	0.69	0.60
Two-year average of years (age-3) and (age-4). Example: for an individual age 40, this is the average of earnings at ages 36 and 37. Sample: Ages 35-60	0.20	0.12	0.27	0.13	-	-

	Women					
	35-49		50-64		65-75	
	1983-1997	1998-2003	1983-1997	1998-2003	1983-1997	1998-2003
Alternative Average Lagged-Earnings Measures						
Basic measure (includes zero earnings years and excludes at least three years prior to mortality window)	0.55	0.21	0.66	0.38	0.93	0.86
Basic measure but including only positive earnings years	0.52	0.22	0.82	0.50	0.79	0.75
Basic measure but excluding at least six years prior to mortality window	0.70	0.28	0.66	0.38	0.94	0.86
Two-year average of years (age-3) and (age-4). Example: for an individual age 40, this is the average of earnings at ages 36 and 37. Sample: Ages 35-60	0.47	0.18	0.36	0.24	-	-