Earnings and health: Comparing earnings measures from single-year self-reported to multi-year administrative records

Rehkopf, David H[1]
Glymour, M Maria[2, 3]

[1] Robert Wood Johnson Foundation Health and Society Scholar, University of California Berkeley and San Francisco
[2] Robert Wood Johnson Foundation Health and Society Scholar, Columbia University, New York
[3] Department of Society, Human Development and Health, Harvard School of Public Health, Boston

Draft copy, comments welcome, please do not cite or quote.

This version prepared for presentation at the Population Association of America (2006) Annual Meeting, New York, NY, March 29-31

Please direct any comments on this draft to: David Rehkopf, UCSF Center for Health and Community, 3333 California Street, Campus Box 0844, San Francisco, CA 94118 (David.Rehkopf@ucsf.edu)

Acknowledgments

The authors thank the Robert Wood Johnson Foundation Health & Society Scholars program for its financial support. MG was supported by a grant from the National Institute of Aging (AG023399).

Abstract

Although the association between earnings and health has been well documented, difficulties in measuring earnings have introduced substantial uncertainty into estimates of the causal effect of earnings on health. The most common earnings measures are based on single-year self-reports, usually assessing earnings in the year prior to the health assessment. Using such a single-year, self-reported earnings measure may incorrectly estimate the causal effects of earnings on health due to bias from measurement error, short-term fluctuations in earnings, reporting bias or reverse causation. We use a unique data set including 40 years of Social Security recorded earnings (SSE) available for a subset of participants (n=5,951) in the Health and Retirement Study (HRS) to assess the importance of these sources of bias, focusing on prevalent disability, diabetes, stroke, heart disease, cancer, and depression. We find evidence that the association between reporting zero earnings and health may be in part due to reverse causation from health to earnings. Specifically, introducing increasingly long lag periods between the time of reporting zero earnings and the health assessment leads to greater attenuation of the association between the two variables, until the association declines to almost null with 5-15 year lags. In contrast, among individuals with positive earnings, we find evidence that the impact of health on earnings, although substantial, diminishes quickly with increasing lag periods between the earnings and health assessments. The reverse causation bias appears to be generally quite modest with a 5-year lag period between earnings assessments and health. For most outcomes, earnings predict health even with 8+ year lag periods between earnings and the health assessment. Importantly, although we find evidence that health events compromise earnings for a period of time, we also show that

individuals with major health events report lower earnings even prior to the event onset relative to comparable individuals who do not suffer such an onset, i.e. poverty predicts risk of disease onset, and disease onset lowers earnings. Additionally, although many limitations of self-reported earnings have been documented, the associations between a single-year self-report of earnings and health outcomes are generally stronger than a single-year administrative report. Averaging SSE in 1961, 1971, 1981, and 1991 predicted health outcomes in 1992 more strongly (odds ratios farther from the null) than any single-year administrative earnings and similarly to a single year self-report of earnings in 1991. We also present evidence that, when current self-reported earnings are used as a control variable in analyses in which earnings are a potential confounder, adding additional multi-year and administrative measures of earnings to the model does not substantially change the regression coefficients for the effects of interest. We also show that adjustment for wealth, a potentially more comprehensive indicator of financial resources, does influence coefficients in some cases (e.g., for race). This provides mixed comfort to epidemiologists using data sets in which self-reported current earnings is the only available indicator of financial resources.

Abbreviations
SRE= Self-reported earnings
SSE= Social Security recorded earnings
HRS= Health and Retirement Study

Introduction

Overview

Overwhelming evidence links income and health, but causal inference in this area is routinely hampered by concerns about the measurement of income. Obtaining accurate retrospective self-reports over the lifetime is difficult and access to databases with wage data based on administrative records is limited. As a result, the income measure most commonly used in health surveys (e.g. National Health and Nutrition Examination Surveys, National Health Interview Survey, Behavioral Risk Factor Surveillance System) is self-reported total income in the single year preceding the survey administration. Self-reported single year measures have repeatedly been shown to be associated with a variety of health outcomes (Backlund, Sorlie and Johnson 1999; Ecob and Davey Smith 1999; Kawachi, Kennedy and Wilkinson 1999). However, using a single year self-reported income measure may be inadequate for identifying the effects of income on health for several reasons. The etiology of most health outcomes and the mechanisms through which income is hypothesized to impact health occur over substantially longer time-periods than a single year. Furthermore, evidence from research on socioeconomic position at different points in the life course suggests that income at early time periods or specific ages may be especially important for some health outcomes (Davey Smith et al. 1998; Lynch et al. 1994). In addition, individuals may effectively smooth earnings over multiple years; thus average or permanent income may be more predictive of health than any single year's value. Individuals may be unable or unwilling to accurately self-report earnings, potentially introducing measurement error or reporting bias. Finally, earnings from the immediately prior year may be especially vulnerable to reverse causation bias because health cataclysms in the prior year may strongly influence both prior year earnings and health in the survey year. The longer the time lag between assessment of earnings and subsequent assessment of health, the less plausible reverse causation is to explain the earnings-health association. Thus, income measures based on single-year self-reports may either under- or over-estimate the effects of income on health. The goal of the analyses we present is to contrast the adequacy of a single year cross-sectional self-reported measure of earnings for predicting disease prevalence with objectively measured earnings over the life course.

To attempt to better understand the effects of income on health, we use a unique data source for participants in the Health and Retirement Study (HRS) that links biennial health reports from 1992 onwards with social security earnings (SSE) records from 1951-1991 (the figure summarizes the years of exposure and outcome data we use). In the analyses below, we seek to address several limitations of prior research on earnings and health, in some cases by testing whether theoretically plausible sources of bias do in fact appear to substantially influence results. First, we will test whether reporting any (as opposed to zero) SSE is associated with better health, examining the association between zero-earnings and health while introducing increasingly long (up to 30 years) lag periods between the SSE and health measures. We next compare how single-year self-reported earnings (SRE) and SSE predict disease prevalence, again introducing increasingly long lag periods between the earnings and health assessments. Combining the retrospective data on SSE available in HRS with the longitudinal nature of the health assessments (biennially from 1992), we are able to examine the association between earnings and health with up to an 11 year lag period between SRE and health, up to a 30 year lag period between SSE and the health assessments. To help verify these results, we present results on how SSE change with onset of a major health condition, using retrospective reports of year of onset of a health condition. For those individuals who reported that they suffered major health condition with onset prior to the baseline HRS assessment, we examine how their SSE trajectories changed in the year of onset of this condition. Third, we will examine how earnings and average earnings at different ages predict later life prevalence of health conditions. Finally, for analyses in which earnings are not the primary predictor of interest, we estimate the extent of residual confounding that persists after adjusting for the easiest to obtain measure of earnings, self-reported earnings for a single year. To assess this, we compare estimated odds ratios relating race, education, gender, and physical activity to health outcomes when adjusted for single year SRE to those obtained when adjusting for the more comprehensive life course earnings data available in HRS.

Below we review the motivation for each of these analyses, and how the results can help us interpret prior research and rule out competing hypotheses to explain the associations between earnings and health.

Threats to causal inference: the impact of health on income

Although the association between earnings or income and health is extremely well established, one of the most important arguments against inferring that earnings influence health is the possibility that transient or chronic poor health influences income, sometimes called “health selection bias” (Kitagawa and Hauser 1973; Smith 1999). This may occur if illness prompts individuals to decrease work hours or change jobs (reducing earnings) or to leave the labor force entirely (reducing earnings to zero). Establishing whether either or both of these processes operate substantially in practice is important in assessing the validity of prior research on earnings and health.

In order to ameliorate the bias due to the impact of health on earnings, one common strategy is to collect measures of income that are a number of years prior to the health-assessment, using longitudinal designs, data linkage or retrospective earnings recall. This is most appropriate for chronic diseases and acute events for which onset reflects an underlying chronic condition. If there is a substantial impact of disease onset on earnings, we expect that introducing longer lag periods between income and health will reduce the source of the association between the two variables and thus attenuate the coefficients if there is substantial reverse causation. This is likely to be especially relevant for middle aged adults, among whom major health events are likely to have occurred recently and who generally seek an income through employment. On the other hand, we expect longer lag periods to introduce only modest attenuation if the relationship is primarily causal from earnings to health. Although epidemiologists debate the pathways linking socioeconomic conditions to health, only a few pathways posit very short etiologic periods between exposure to socioeconomic disadvantage and onset of major health conditions. The processes by which earnings-related social deprivations change human physiology to increase vulnerability to major health conditions such as stroke, heart disease, or diabetes, do not usually operate over only a year or two, but rather develop over many years and even decades. Major health conditions, even those with apparently acute onset such as myocardial infarction, are generally preceded by a decades or longer period of developing, largely asymptomatic, disease, such as carotid atherosclerosis, and such asymptomatic conditions are in turn the consequence of long-term behavioral, environmental, or physiological exposures, such as diet or systemic inflammation (Berenson 2002). Earnings a year before, five years before, or ten years before the health assessment are therefore likely to have similar ability to predict health outcomes if the causal direction is from earnings to health; similarly, averages over longer time periods should be more strongly associated with health outcomes. Modest attenuation with increasing lag periods may reflect a cumulative effect of exposure to long-term earnings deprivation. Substantial attenuation in the earnings-health association as increasing lag periods are introduced is, however, more consistent with the hypothesis that reverse causation from health to earnings could largely account for the original cross-sectional association. This argument depends strongly on the specific health condition under consideration. Depression, automobile accidents, and birth outcomes may be more strongly influenced by recent or current income. In addition, the separation of long term and proximal effects of earnings or social class more generally has been the subject of numerous investigations in life course research.

Despite the theoretical arguments for a sizable impact of health on income, McDonough found very little difference in predicting all-cause mortality between a one year lag and a five year lag (1997). The health literature has generally assumed that the etiology of most diseases would be such that 5 years prior to the clinical detection of disease there would be few symptoms leading to changes in work that would impact earnings. However, a five year lag such as that used by McDonough may still capture effects of subtle preclinical symptoms of illness on income. Primarily because of a lack of data, there has been little work to quantify the different associations of income and health with more extended lag periods.

In analyses of the effects of income on health, it is often unclear whether it is preferable to exclude individuals not in the labor force from analyses or treat these individuals as having zero earnings. If the effects of illness on earnings are primarily mediated by leaving the labor force, including zero-earnings individuals in analyses will introduce especially large biases due to health selection, while restricting analyses to positive earners may substantially circumvent this bias. Previous analyses, however, have not explicitly contrasted the attenuation of coefficients with increasing lag periods for the effect estimates of being in/out of the labor force and the effect estimates for positive earnings.

Self report vs. administrative earnings data

As noted by Bound *et al*., investigators using self-reported earnings data tend to make "strong – and exceedingly convenient" assumptions about error in the measurement of income (Bound et al. 1994). Self-reported cross sectional earnings are normally taken as accurate, and the error with respect to predicting health outcomes has not been previously investigated. The primary reason for this is that earnings data that are not based on self-report have rarely been available for use due to the limited linkage of earnings data with demographic and health data. Exceptions are the IRS-SSA Income and Wealth Study which had statistics of earnings from the 1969 Social Security Administration Summary (SSA) Earnings file which matched earnings data from 1969 with mortality from 1970 to 1976 (Caldwell and Diamond 1979), the 1973 CPS-IRS-SSA exact match study (Kliss and Scheuren 1978), which included income data from both SSA earnings as well as single year (1972) internal revenue service individual income tax returns, and the 1978 CPS-Social Security Earnings Records match which contained 1950-1978 social security earnings (Bound and Krueger 1991). Evidence has been mixed, but these studies suggest that self-reported earnings data provide a reasonable approximation to actual earnings data as determined through administrative files. Using a special supplement to the 1977 Current Population Survey (CPS) that included both worker and employer reports of wages, there were virtually no differences between these sources of data (Mellow and Sider 1983). Using data from the Panel Study of Income Dynamics (PSID), Bound *et al.* also found that self-report earnings data were fairly accurate (Bound et al. 1994). These results stood in contrast to analyses using a multiple imputation approach to correct for errors measurement error of income in the PSID, which found that errors are substantial and methods should be used to account for these errors (Brownstone and Valletta 1996).

The general conclusion drawn from the studies that directly contrast self-report and administratively recorded earnings is that measurement of wages through self-report is fairly accurate but imperfect. While these studies allowed estimation of the accuracy of income data, there has generally been no health data or only a linkage with mortality data. The relevance of bias in the report of earnings with respect to impact on health outcomes has thus rarely been investigated. While the earnings studies have been important for researchers, we are also interested in understanding the practical implications of earnings measurement error for researchers interested in health outcomes as their dependent variable. That is, does the mismeasurement of income through single-year self-report lead to erroneous conclusions about the associations of earnings and health outcomes?

While in prior analyses and the analyses presented here social security earnings are considered as the gold standard, it is important to note that social security earnings are not without limitations (Bound and Krueger 1991; Haider and Solon 2006). Workers who were self-employed, received a large portion of their income from tips, or who spent part of the year in uncovered employment likely have social security earnings which underestimate true earnings. Haider and Solon note that the number of jobs with social security earnings data is over 80% of jobs since 1957 and over 85% for most of the sample period (Haider and Solon 2006). Between 1951 and 1956 this percentage ranged from 66 to 79%, but these years are prior to the earliest SSE assessments we use in the current study. Second, social security earnings are measured only up to the maximum dollar amount eligible for social security tax. Individuals who earn more than this cap appear in the earnings file as if they earned the maximum value. In the HRS linked social security administrative data the fraction of individuals who reached this cap varies from 62% of the sample in 1965 to 9% of the sample in 1991. Although this is a potentially major limitation of using SSE as an indicator of total earnings, most research on health indicates diminishing returns to increasing income (Backlund, Sorlie and Johnson 1996). Thus for predicting health, it is likely that variations in the lower end of the earnings spectrum are much more important to capture accurately than variations at high levels of earnings.

How biased is current income for predicting health?

Both permanent and transitory changes in income over time have been well documented (Bane and Ellwood 1986; Gottschalk 1982), but the extent to which this instability leads to errors that underestimate the relation between single year measures of earnings and health outcomes is not clear.

Using HRS data linked to SSE data, Haider and Solon investigate the extent to which earnings in a single year mismeasure lifetime earnings due to differential trajectories of earnings over the life course that differ systematically by income level or other demographic characteristics (Haider and Solon 2006). Their findings indicate that current income does not follow a “textbook errors-in variables model” for describing lifetime income. This error varies depending on age, where there will be less of a discrepancy in single year earnings between workers with high and low earnings at younger ages, with a more rapid growth in earnings for the high earners. There will, however, be a point at a later age where the decreasing steepness (and on average decline) in single year earnings for the highest earning group in particular will lead to a difference in single year earnings that approximates the difference in average lifetime earnings (Haider and Solon 2006). So while there is a potential for at some ages to have current earnings differences approximate lifetime earnings differences, it is unknown whether this type of error when present will impact the association of single year earnings with health outcomes.

In addition to bias in estimating income effects, the extent to which current earnings inaccurately reflects lifetime earnings is relevant in studies for which earnings are considered a potential confounder (common cause of the exposure and outcome of interest). Many studies seek to control for the effects of income in regressions focused on the causal impact of other factors, but have only single-year self-report measures of earnings. The extent to which a single measure of self-reported income is an adequate control for income in the investigation of the effects of independent variables such as physical activity or race has been the subject of extensive debate in the medical literature (Braveman et al. 2005; Kaufman, Cooper and McGee 1997; Krieger, Williams and Moss 1997). Much of the recent work on measures of income presumes that a single measure will miss important aspects of income, and this view is reasonably based upon evidence that shows disease-specific independent effects of socioeconomic position at different points in the life course. Nevertheless, many plausible sources of bias do not turn out to be major problems in reality, so empirical assessment is invaluable. There has to our knowledge not been a comparison of effect estimates when comparing a very comprehensive measure of life course earnings available with social security earnings data as compared to a single self-report measure of current income.

Competing hypotheses addressed in the current analyses

Our first set of analyses speaks to distinguishing between health selection and earnings causation hypotheses. We address whether individuals who report zero SSE in one year are at increased risk of illness in subsequent years. The most common reason for an individual included in the SSE database to report zero SSE in a given year is because she or he left the labor force in that year. However, SSE reports do not apply for certain jobs or for under-the-table earnings. We have controlled for occupational class as well as years of education in our analysis to control for this potential confounder. In all cases, our earnings measures are temporally prior to the health assessments, but we examine increasing lag periods for the measurement of zero-earnings ranging from 1-31 years prior to the reporting of the health outcome. This allows us to assess the number of years prior to report that health may be impacting earnings or earnings may be impacting health. If the coefficient declines substantially with increasing lag periods between earnings and health assessment, we consider this most consistent with a "health selection" as opposed to "earnings causation" hypothesis. Similarly, we examine the relationship between health at baseline and log earnings for individuals with positive SSE for years prior, introducing increasing lag periods up to 31 years.

The next analyses extend the prior work by using successive health examinations to introduce a time lag between earnings and health. We examine how self-reported earnings and SSE in 1991 predict health outcomes assessed in 1992, 1994, 1996, 1998, 2000 and 2002. We expect that health assessments immediately after income assessments are substantially more vulnerable to causation from health to income, whereas when there is a longer period of time between the assessment of income and the onset of disease this suggests primarily income to health effects.

For HRS enrollees who reported prevalent major health conditions at study enrollment, we use the self-reported year of onset of this condition to directly estimate how onset affected SSE trajectory, and the number of years after onset before earnings recovered to pre-onset levels.

The concern implicit in the health selection arguments addressed above is that the earnings coefficient is overestimated due to reverse causation, but another set of issues hypothesizes that the earnings coefficient might underestimate the effect of earnings on health because of measurement error or inappropriate etiologic periods. To address the etiologic period concern, we examine how SSE at ages 35, 45, or 55 predicts disease outcomes. With respect to attenuation due to measurement error, we also examine differences between how self-reported and objective earnings data predict health outcomes. We consider whether the data are more consistent with either of two competing hypotheses: a)Administratively-reported income data will more accurately predict health because it is not subject to self-report bias, b) self-report income data will capture economic circumstances better than single-year administrative-reported earnings because it will more accurately capture permanent income, either due to 1) respondents biasing reporting to more accurately reflect lifetime earnings or, 2) capturing current earnings at a time when they approximate lifetime earnings. Similarly, self-reported earnings data will capture economic circumstances as well as administrative-reported earnings averaged over multiple points in the lifecourse.

This question is relevant to better understanding results from the hundreds of studies that use current income as a proxy for lifetime earnings, as well as recommend measures for future income-health investigations.

Finally, we will examine the extent of residual confounding by financial resources that may persist after adjusting for the standard measure of income: self-reported single-year earnings. Extensive prior thinking about social resources beyond income, and evidence about measurement error in single-year self-report earnings measures both suggest that there would be extensive residual confounding when using a single measure of current self-reported income. On the other hand, when confounding is induced by a common cause of the exposure and the outcome, the magnitude of bias is proportional to the product of the relationship between the confounder and the exposure and the confounder and the outcome. In many circumstances, control for even an imperfect measure of a confounder eliminates most of the bias. We assess this by examining how regression coefficients for commonly considered risk factors for health conditions (e.g. physical activity) are modified as increasingly comprehensive measures of financial resources are included in the statistical model as covariates.

Methods

sample

Data are from Health and Retirement Study (HRS) participants enrolled in 1992, an ongoing national probability sample of US residents born 1931-1941 and their spouses. The current analysis included only age-eligible individuals. The Survey Research Center at the University of Michigan provides detailed documentation on the HRS sampling design and selection and validation of cognitive and health measures in HRS (Heeringa and Connor 1995; Ofstedal, McAuley and Herzog 2002; Wallace and Herzog 1995).

From the 9,750 age-eligible HRS sample members interviewed in 1992, we made the following exclusions: 2,459 (25.2%) did not release Social security earnings (SSE) data or could not be matched to the Social Security Administration database; 1,256 (12.9%) released earnings data but had fewer than 10 years of positive earnings reported during ages 25-60; outcome measures in 1992 were not available for 84 (0.9%). Thus, a maximum of 5,951 sample members were included in the analysis data set; characteristics of this group compared to the 3,799 HRS respondents not in the analysis set are presented in Table 1. Additional sample members were excluded from selected analyses because of incomplete data for a particular year of earnings. For models comparing 1961, 1971, 1981, 1991 SSE and 1991 self-reported income, 3,187 individuals were excluded due to missing SSE in one of the four years and 134 respondents were excluded due to missing self-reported income. The oldest sample members (born in 1931) were aged 60 by 1991, the last year of SSE data, but the youngest sample members (born in 1941) were only 50. As a result, sample sizes diminish for earnings estimates above age 50, to a minimum sample of 1,040 for earnings at age 59.

measures

Social security earnings (SSE) were based on annual earnings from 1951 to 1991 released by the Social Security Administration, adjusted to 1991 dollars using the CPI-U. Respondents report zero earnings in many years, and it was unclear whether zero-earnings years should be excluded from the analyses or treated as low-income years. We compare results using these two approaches. To assess the role of measurement error or short term instability in earnings estimates, we average earnings over 5 year periods: 1967-1971; 1977-1981; 1987-1991. To assess the role of earnings in multiple calendar years, we average earnings in 1961, 1971, 1981, and 1991. Earnings at ages 25 to 60 were calculated based on birth year only. For example, SSE in calendar year 1965 was used as the earnings at age 25 for respondents born in 1940.

We also use a measure of self-reported individual earnings in 1991 and self-reported household wealth (total assets) in 1991. For both of these variables, we used the standard RAND coding of the HRS data set.

ADL limitations were based on five activities: bathing, dressing, eating, bed, walking. Instrumental ADL limitations were based on three activities: Dependence in any of these either basic or instrumental activities was considered impairment. Depression was assessed with an 8-item version of the CESD, and a score above 3 was considered depressed. Disease diagnoses for stroke, heart disease, diabetes, and cancer were based on self-report during the 1992 interview, with all questions phrased as “Has a doctor ever told you that you had…” Participants who reported certain health conditions (stroke, myocardial infarction, cancer, diabetes, or broken bones) at enrollment in 1992 were asked the age of onset of this condition. We used this to calculate the year of onset (based on the participant's birthyear). We also focused on year of onset for major health conditions we considered especially likely to have immediate consequences for labor market participation: stroke, myocardial infarction, or cancer. The selection of these conditions was based on the conditions for which the HRS questionnaire addressed age of onset.

All analyses were adjusted for measured covariates temporally prior to earnings that may predict both earnings and subsequent health, including year of birth (linear), sex, race (black vs white), Hispanic ethnicity, mother's education, father's education, father's occupational category (unknown, farming, military, or ranked 0-3 with 0=managerial/professional, 1=sales/clerical, 2=protection services/mechanics/precision production/construction, 3=personal, health, or food services/ machine operators/transport operators), region of birth (Northeast, South, Midwest, West, non-US), and years of education (linear)

Physical activity in 1992 was assessed by self-report "How often do you participate in vigorous physical exercise or sports -- such as aerobics, running, swimming, or bicycling?" with a 5-point frequency scale (3 or more times a week, 1 or 2 times a week, 1 to 3 times a month, less than once a month, or never?).

model link

All analyses were conducted with logistic regression using earnings variables to predict presence or absence of a health condition. Regressions were adjusted to account for the complex survey design and sample weights using SAS proc surveylogistic.

For analyses of how onset of health conditions affected SSE trajectory, we specified growth curves in Stata v 9 (xtreg), using earnings as the dependent variable. We specify the year of onset of the condition as the reference year, and examine how earnings in the years prior to onset and subsequent to onset compare with earnings in that year. We show analyses restricted to only individuals who reported disease onset, and comparing those who reported a major condition to those who reported no such condition at enrollment.

Results

Respondents included in the earnings analyses are compared to age-eligible HRS respondents excluded in Table 1. There were statistically significant differences in a number of characteristics. Among demographic variables, excluded respondents were more likely to be black, Hispanic, or female. Among socioeconomic variables, excluded respondents had lower self-reported earnings in 1991 and had parents who had lower levels of occupation and education. The socioeconomic differences are unsurprising since the primary reason for exclusion was reporting no SS earnings in most years of adulthood. The differences in health outcomes between the excluded and included individuals were minimal, with the exception of impaired activities of daily living, with excluded individuals having a higher prevalence.

In most of the subsequent analyses that we will present, we examine the impact of different years of income independently. Hence, we describe first the Pearson correlations between different measures of log earnings at different periods of time. The correlations between self-reported earnings in 1991 and social security earnings at ages 25, 35, 45, and 55 were 0.19, 0.28, 0.37 and 0.48, respectively (p-values for association were all <0.0001). The correlations between individual year SSE and 5-year average SSE were above 0.75 for all cases examined (Table 2), suggesting modest advantages to using the multi-year averaged values. The correlation between SSE and SRE reported for the same year (1991) was 0.48.

Given the modest association between 1991 SSE and 1991 SRE, we examined the overall distribution of differences of self-reported wages and social security wages. Although median SRE was higher than SSE by $400, there was substantial variability around this median difference: the 75th percentile of the SRE-SSE gap was $4,560 and the 25th percentile was -$1,300. There was a positive and significant correlation between age and a larger gap between types of earnings measures (spearman correlation of 0.049, $p<0.01$). As might be expected given the cap on SSE, there was a negative and significant correlation between a higher SRE as compared to SSE as SSE increased (spearman correlation of -0.082, $p<0.001$). Thus, older individuals and lower SSE individuals averaged a higher absolute gap between SRE and SSE.

zero earnings compared to some earnings with increasing time lags between earnings and health assessments

Table 3 presents odds ratios for the effect of reporting any social security earnings (SSE) in a given year as compared to reporting no SSE for a given year on prevalent conditions in 1992. We examined lags from 1 year (1991 SSE) to 31 years (1961 SSE) (note that rows show 1961,1971, then every year from 1976 to 1991). Men with zero SSE in 1991 had markedly higher odds of ADL/IADL, diabetes, heart disease, stroke and depression in 1992 compared to men with any positive earnings. For each outcome odds ratios declined with increasing lag periods, such that with 15 year lags (1976 earnings), reporting zero earnings was not associated with elevated risk of illness in 1992 for any outcome. The time lags where zero-earnings ceased to predict the health outcome varied by the health condition, with the longest time lag relevant for stroke and depression. There were no significant associations found for cancer. Results were similar for women, except that the lag periods at which reporting zero-earnings appeared to no longer predict health outcomes were in some cases of fewer years.

Comparing SRE and SSE lagged earnings among positive earners and subsequent health

Table 4 presents the odds ratios of disease from 1992 to 2002 (by every other year) associated with an increase in log earnings in the year 1991, comparing associations with self-reported earnings (SRE) and social security earnings (SSE). The purpose of this table is to rule out the possibility that the association between earnings and disease prevalence arises because illness impacts income. If there is a substantial health selection effect, we would expect to see smaller odds ratios when predicting health in the years following income assessment, then convergence to a stable estimate of income's impact on disease. Among men there were generally increases (towards the null) in the odds ratios of disease for 1 to 3 years following the assessment of income, then odds ratios were relatively stable for most outcomes out to 10 years beyond the assessment of income. Results were similar for women. Odds ratios were near the null and not significant for heart disease and cancer. There were also relatively minor differences between using self-reported earnings data and social security earnings data for most outcomes. For diabetes and ADL/IADL, there were differences in the magnitude of the odds ratios during the first three years following the assessment of income, with SRE showing a stronger association.

Another approach to examining whether health conditions affect earnings or reflect earnings or both is to use the date of onset information for HRS participants reporting a major health condition at enrollment. Table 5 shows the difference in earnings between the year prior to diagnosis and years of earnings up to 10 years prior and 10 years after that year, restricting to individuals who were diagnosed with cancer, diabetes, myocardial infarction, broken bones or stroke or, in right side of Table 5, diagnosed with one of three major conditions (cancer, stroke or myocardial infarction). Levels of income for years prior to diagnosis is generally slightly higher than the year prior to diagnosis, but fairly stable (i.e. there are not substantial decreases or increases in income in the years prior to the year prior to diagnosis). In the years following diagnosis, for the first few years income is lower, and then by 5 years later income appears to have recovered, even for onset of major conditions. These results are similar for men and women, but with a slightly faster recovery of income for women. Results are similar when excluded to the diagnosis of only major conditions.

The coefficients for individual years in Table 5 are not statistically significant, but this is likely attributable to the small sample size and intrinsic variability in earnings. To strengthen the analyses, we tested whether there were significant earnings differences in the 5-year period prior to diagnosis or in the 5-year period following diagnosis, compared to individuals who were never diagnosed with a condition. Table 6 presents the results of these analyses. These results show that in the 5-year period prior to diagnosis men's annual earnings were reduced by approximately $543. In the 5-year period after diagnosis, men averaged $1,346 lower annual earnings than others. Thus, income was lower for individuals with disease onset both before and after disease diagnosis, indicating that individuals with lower earnings were more likely to have an onset of disease, and individuals were likely to have lower income after onset. There is a greater impact on earnings for men than women, and for major conditions as compared to all conditions. The largest differences are thus seen for men with or without the onset of a major condition, where income was on average $1462 (939, 1985) lower five years before onset for those with a diagnosis, and $4517 (3841, 5192) lower after the onset of disease.

average earnings by age and year

We next compared how well SSE at different ages predicted prevalence of disease in 1992. Table 7 presents the odds ratios of disease in 1992 by earnings at ages 35, 45, 55 and averages of earnings over ages 31-35, ages 41-45 and ages 51-55. For ADL/IADL and depression, income was a significant predictor at age 45 or an average of ages 41 to 45. For heart disease and stroke, earnings averaged over the years 51 to 55 were significant predictors. No significant relationships were found for diabetes or stroke. In general, little difference was found when averaging over a 5 year period as compared to single year measures, and stronger associations were found at older ages. Among women, significant associations were found only for depression at age 55, and diabetes at age 35. Higher level of income at age 55 was associated with a higher probability of diabetes in 1992.

In Table 8 we present odds ratios associated with incremental increases in earnings. We again restricted these analyses to respondents with positive earnings in all four years considered. We compare self-reported earnings in 1991 (SRE) to social security earnings in 1991, 1981, 1971, and 1961 (all natural log transformed). First, we note that despite the expected limitations in self-reported income, the association with SRE and health outcomes is as strong as or stronger than the association with SSE and the health outcome for nearly every comparison presented. In addition, SRE associations were almost identical to the average of SSE over years 1961, 1971, 1981 and 1991. The exception to this is for depression, where there was a stronger association with average SSE. Odds ratios for the association with disease in 1992 were generally similar across different lag periods. Because of the known instability in year to year earnings, we also examined whether averaging earnings over a 5-year period improved the predictive power of the earnings measures. For example, we averaged earnings from 1987-1991 and compared this to the 1991 only coefficient. This approach offered surprisingly little benefit. Overall, the coefficients were not consistently lower (or higher) and the confidence intervals were not substantially narrower. Results were generally similar for men and women.

residual confounding of single year self-reported earnings

Earnings are often of interest primarily as confounders of other exposure-outcome relationships. In Table 9 we show results from our examination of how adjustment for increasingly comprehensive measures of financial resources affects estimated coefficients for four exposures: race, sex, years of education, and vigorous physical activity. There is an extensive literature on how socioeconomic position might confound regression coefficients for each of these variables. We note that earnings are also generally affected by race and education, so the race and education coefficients conditional upon earnings do not have a clear causal interpretation. Despite this, questions about such adjusted coefficients, and how they would change with improved earnings measures are often of interest in the epidemiologic literature.

The first column in table 9 shows, in models without adjustment for any measure of earnings, statistically significant associations of race, gender, education and physical activity with ADL/IADL, diabetes (not for education), heart disease (physical activity only), stroke (physical activity only), cancer (education only) and depression. The second column of results in table 5 shows the odds ratios of disease prevalence after controlling for self-reported earnings in 1991, the year prior to assessment of disease prevalence. As discussed in the introduction, this is the most common way in which the largest nationally representative health surveys are able to control for earnings within the limitations of the data available. Changes in the associations between each exposure and outcome after adjustment for SRE are specific to each exposure/outcome combination, but in general, earnings adjustment introduces moderate changes to the coefficients.

The third, fourth, fifth and sixth columns show the impact of additionally controlling for social security earnings in 1991, 1981, 1971 or an average of earnings over 1961, 1971, 1981, and 1991 odds ratios for disease in 1992. As compared to controlling for the limited measure of self-reported income in 1991, there was little impact in controlling for these additional measures on any of effect measures.. In the final column in table 9, we present results which also control for wealth in 1991. Unlike including better measures of earnings, adjusting for wealth introduced at least modest attenuations in the estimated odds ratios for the association of black race with disability, diabetes, depression, and stroke.

Discussion

Our analyses focus on understanding the limitations and appropriateness of using a single measure of self-reported earnings to predict the prevalence of health outcomes by contrasting coefficients obtained with this measure to those obtained with objective measures of earnings over the earnings careers of survey participants. The reason for this focus is that self-reported earnings is usually the only measure of earnings available in the nationally representative health surveys that public health and medical researchers use for understanding population prevalence and trends and the etiology of disease. In addition to being used to measure the importance of earnings itself, self-reported income is commonly used to control for the confounding effects of earnings or income when other exposures are of primary interest. Prior to our analyses, the adequacy of this single earnings measure for these purposes has not been understood.

limitations

Before summarizing the implications of our findings for understanding the use of a single self-reported measure of earnings, there are a number of limitations and considerations of generalizability that are important to consider. First, as shown in table 1, there are significant socioeconomic differences between individuals included in our analysis and those not included, due to the lack of social security earnings data for some participants. While this is a cause of concern for the generalizability of our findings, there are three mitigating factors that reduce this concern. First, our analyses control for the variables that predict inclusion in the SSE dataset, including race, parental education and individual education. Second, the disease outcomes were not substantially different between those included and excluded, with the exception of ADL/IADL. Finally, prior analysis of the social security linkage supports the view that differences in those included and not included in the social security earnings data linkage should have a limited impact on generalizability.

A second limitation to our study is the lack of objective health outcomes. While the HRS survey's strength for this analysis is the earnings data over time as well as information on wealth, health outcomes are self-reported. It is possible that there are associations between earnings and the likelihood of reporting health outcomes that contribute towards our findings, rather than associations with actual disease. Despite this concern, initial reports on the health constructs in the HRS data have shown good reliability (Wallace 2001). The reanalysis of this data when objective health outcomes and biomarkers are included in the HRS dataset will be useful for confirming our findings.

While social security earnings data provides a useful objective measure to contrast with self-reported earnings, a third concern is that it is far from a gold standard for measuring earnings. Some occupations or jobs have not required reporting social security earnings, and some of these requirements have varied over time (as discussed in the introduction). In addition, the cap on earnings means that there is an underestimate of earnings for those individuals with the highest earnings. The important effects of these limitations are to a certain extent mitigated by controlling for additional socioeconomic variables in our analyses. In addition, regarding the cap on earnings, a number of studies have shown that there is little change in association between income and health with increasing income at the highest income levels, so this cap likely has no impact on our findings.

Fourth, it is important to note that we used individual measures of income here, when earnings are in reality pooled within households – and the household measure is generally the most relevant level of analysis for examining the impact of material resources on health (Krieger et al. 1997).

Finally, it is important to note that our results are for an older population, where the impact of earnings and income on disease have been consistently demonstrated to be weaker. Theories to explain this attenuation with increasing age have focused on the different role that income plays in the lives of older individuals, selection effects where the poorest individuals who are least healthy have died at earlier ages, and the fact that earnings may not be an appropriate measure of socioeconomic position once individuals are retired. Our results do not speak directly to distinguishing between these hypotheses, although they are generally less consistent with the later explanation.

implications of results

First, we find that reporting zero as compared to no social security earnings is a strong predictor of all of the health outcomes we examined, with the exception of cancer. In examining lag time of assessing absence of earnings up to 30 years prior to the assessment of health, we find that, for most outcomes, individuals who report zero earnings even 10 to 15 years prior to the health assessment have higher odds of a health condition. The marked attenuation in coefficients associated with longer lag periods suggests that inclusion of zero-earners in the assessment of earnings introduces substantial reverse causation bias. Based on this result, our subsequent analyses examined only the strata of individuals who reported at least some earnings for the years and ages we examined.

Second, our results are consistent with there being an increased association with disease onset 1-3 years after the assessment of earnings, with little increase after that time. This is consistent with their being an initial impact of health on earnings, but that after a lag of 5 years and greater the association with income appears stable. For studies with only a single measure of the prevalence of health conditions, using a lag of 5 years for the assessment of income prior to examining disease prevalence would seem to offer promise for reducing the impact of causation from health to income.

Our results also underscore the way in which income is different for those with and without disease onset, both before and after diagnosis. Our data showing that there is generally lower earnings before disease onset, and even lower earnings after onset, quantifying the importance of low earnings for etiology of disease onset, and the impact of a health shock on dramatically decreasing earnings, especially in the case of the onset of major conditions. This has important implications for longitudinal analyses that attempt to isolate the effect of earnings on health by conditioning on baseline health. Since baseline health already reflects the influence of prior earnings, such adjustments may bias the coefficients away from the causal effect of earnings on subsequent health outcomes (Cole and Hernan 2002).

Thirdly, our results shed light on the appropriateness of using self-reported income to proxy lifetime earnings in contrast with using more difficult to obtain earnings data. Self-report income shows nearly identical associations with the health outcomes we examined as compared to an average of social security earnings from 30 years, 20 years, 10 years and current earnings. This is somewhat of an unexpected result given the extent of speculation over the bias in a single self-report measure for capturing lifetime earnings. One explanation is that respondents naturally adjust their self-reported earnings based on lifetime earnings experiences. However, it is unclear if SRE performs better than expected because it truly reflects financial resources well or because it suffers offsetting sources of bias. Nonetheless these results are at least modestly encouraging with respect to the validity of results based on single-year SRE. Future work must determine whether our results also hold for populations under the age of 55 and for other health outcomes.

Finally, results for examining residual confounding for studies examining the impacts of being black vs. white, physical activity and education find that additional objective life-course measures of earnings do not appreciably change findings. We note, however, that the inference here is limited because controlling for the self-reported earnings measure initially did not have a substantial impact on effect estimates. This may be due to the number of other potential confounders we controlled for in all models (e.g. father's and mother's education, father's occupation). However, for outcome and predictor combinations where there was a small change, having additional measures of earnings to control for residual confounding may not be necessary. In contrast, adjusting for wealth did influence coefficients for race, consistent with prior work noting the stark wealth differentials between US blacks and whites (Conley 1999).

conclusions

Efforts to link existing datasets with objective lifetime working histories should continue and be done wherever possible, as they allow much more detailed and useful studies of the dynamics of income and subsequent health effects. In addition, well designed analyses have shown systematic error in using current earnings to approximate lifetime earnings (Haider and Solon 2006). However, our analysis shows that for studies where only current income is available for predicting disease prevalence, this measure estimates similar associations with disease as compared to average of objective earnings data over the life course. Similarly, we find that using current self-reported income as a control variable for lifetime earnings is appropriate. We also find that there are significant lag effects to having no earnings data in given years, and this should be used as a stratification or control variable in studies of health outcomes. In contrast, based on our data, a lag of five years between the assessment of income and health conditions appears to substantially reduce the impact of health on income. Finally, while there are significantly lower earnings for those who have the onset of disease, there is an even more dramatic impact on decreasing earnings in the immediate years after disease occurs. Examination of these questions using objective health outcomes and in younger populations will be useful to see if our results are generalizable beyond the ages and outcomes we examined here.

Figure. Earnings and health data from HRS participants born 1931-1941 used in the current analyses.

Years	1961	…	1990 ·	1991	1992	1994	1996	1998	2000	2002
HRS enrollment & baseline interview					X					
Health					X	X	X	X	X	X
Self-reported earnings				X						
Social Security earnings	X	X	X	X						
Age of Oldest Cohort	30	…	59	60	61	63	65	67	69	71
Age of Youngest Cohort	20	…	49	50	51	53	55	57	59	61

X=data available.

…=all intervening years available

<u>References</u>

Backlund, E., P.D. Sorlie, and N.J. Johnson. 1996. "The shapes of the relationship between income and mortality in the United States: Evidence from the National Longitudinal Mortality Study." *Ann Epidemiol* 6:12-20.

—. 1999. "A comparison of the relationships of education and income with mortality: the national longitudinal mortality study." *Social Science and Medicine* 49:1373-1384.

Bane, M.and D.T. Ellwood. 1986. "Slipping into and out of Poverty: The Dynamics of Spells." *The Journal of Human Resources* 21(1):1-23.

Berenson, G. 2002. "Childhood risk factors predict adult risk associated with subclinical cardiovascular disease. The Bogalusa Heart Study." *Am J Cardiol* 90(10C):3L-7L.

Bound, J., C. Brown, G. Duncan, and W.L. Rodgers. 1994. "Evidence on the validity of cross-sectional and longitudinal labor market data." *Journal of Labor Economics* 12(3):345-368.

Bound, J.and A.B. Krueger. 1991. "The extent of measurement error in longitudinal earnings data: do two wrongs make a right?" *Journal of Labor Economics* 9(1):1-24.

Braveman, P.A., C. Cubbin, S. Egerter, S. Chideya, K.S. Marchi, M. Metzler, and S. Posner. 2005. "Socioeconomic status in health research - One size does not fit all." *Jama-Journal of the American Medical Association* 294(22):2879-2888.

Brownstone, D.and R.G. Valletta. 1996. "Modeling Earnings Measurement Error: A Multiple Imputation Approach." *The Review of Economics and Statistics* 78(4):705-717.

Caldwell, S.and T. Diamond. 1979. "Income Differential in Mortality: Preliminary Results based on IRS-SSA Linked Data." in *Proceedings of the Survey Research Methods Section, American Statistical Association*.

Cole, S.R.and M.A. Hernan. 2002. "Fallibility in estimating direct effects." *Internation Journal of Epidemiology* 31(1):163-165.

Conley, D. 1999. *Being Black, Living in the Red: Race, Wealth, and Social Policy in America*. Berkeley: University of California Press.

Davey Smith, G., C. Hart, D. Blane, and D. Hole. 1998. "Adverse socioeconomic conditions in childhood and cause specific adult mortality: prospective observational study." *BMJ* 316:1631-1635.

Ecob, R.and G. Davey Smith. 1999. "Income and health: what is the nature of the relationship?" *Social Science and Medicine* 48:693-705.

Gottschalk, P. 1982. "Earnings Mobility: Permanent Change or Transitory Fluctuations?" *The Review of Economics and Statistics* 64(3):450-456.

Haider, S.and G. Solon. 2006. "Life-cycle variation in the association between current and lifetime earnings." *American Economic Review* 96(4):1308-1320.

Heeringa, S.G.and J. Connor. 1995. "Technical description of the Health and Retirement Study sample design." Ann Arbor, Michigan: Survey Research Center, University of Michigan.

Kaufman, J., R.S. Cooper, and D.L. McGee. 1997. "Socioeconomic Status and Health in Blacks and Whites: The Problem of Residual Confounding and the Resiliency of Race." *Epidemiology* 8(6):621-628.

Kawachi, I.o., B.P. Kennedy, and R.G. Wilkinson. 1999. *The society and population health reader*. New York: New Press : Distributed by W.W. Norton.
Kitagawa, E.and P. Hauser. 1973. *Differential Mortality in the United States*. Cambridge,MA: Harvard University Press.
Kliss, B.and F.J. Scheuren. 1978. "The 1973 CPS-IRS-SSA Exact Match Study." *Social Security Bulletin* 41(10):14-22.
Krieger, N., D.R. Williams, and N.E. Moss. 1997. "Measuring social class in US public health research: concepts, methodologies, and guidelines." *Annu Rev Public Health* 18:341-378.
Lynch, J.W., G.A. Kaplan, R.D. Cohen, J. Kauhanen, T.W. Wilson, N.L. Smith, and J.T. Salonen. 1994. "Childhood and adult socioeconomic status as predictors of mortality in Finland." *Lancet* 343(8896):524-527.
Mellow, W.and H. Sider. 1983. "Accuracy of Response in Labor Market Surveys: Evidence and Implications." *Journal of Labor Economics* 1(4):331-344.
Ofstedal, M.B., G.F. McAuley, and A.R. Herzog. 2002. "Documentation of cognitive functioning measures in the health and retirement study." Pp. 68. Ann Arbor, Michigan: Survey Research Center, University of Michigan.
Smith, J.P. 1999. "Health Bodies and Thick Wallets: The Dual Relation Between Health and Economic Status." *J of Economic Perspectives* 13(2):145-166.
Wallace, D.N. 2001. "Discriminatory public policies and the New York City tuberculosis epidemic, 1975-1993." *Microbes Infect* 3(6):515-524.
Wallace, R.B.and A. Herzog. 1995. "Overview of the health measures in the health and retirement study." *Journal of Human Resources* 30(Supplement):S84-S107.

Table 1. Comparison of Health and Retirement Study sample members with and without social security earnings data

	Individuals excluded from the analysis		Individuals included in the analysis		
	n/mean	%/s.e.	n/mean	%/s.e.	p-value*
Demographic	3799		5 951		
men	1507	40	3087	52	<0.01
black	710	19	984	17	<0.01
Hispanic	486	13	416	7	<0.01
birth year	1936	3.1	1936	3.2	0.58
Socioeconomic					
household wealth (median,IQR)	91 200	203 100	90 000	175 800	0.84
self-reported income (median, IQR)	7 000	25 500	17 000	28 600	<0.01
education (years)	11.8	3.5	12.2	3	<0.01
mother's education	8.9	3.9	9.4	3.5	<0.01
mother's education unknown	421	11	492	8	<0.001
father's education	8.7	8.9	9	9.1	<0.01
father's education unknown	579	15	674	11	<0.01
father's occupation					
managerial/professional	1026	27	1786	30	
sales/clerical	267	7	442	7	
skilled manual/service	556	15	1031	17	0.56
unskilled labor	644	17	1175	20	
army	28	1	44	1	0.99
farming	651	17	1131	19	<0.01
unknown	1306	34	1517	25	<0.01
region of birth					
northeast	566	15	1010	17	<0.01
midwest	723	19	1536	26	<0.01
south	1327	35	2200	37	<0.01
west	267	7	432	7	0.67
non-United States	423	11	430	7	<0.01
health in 1992					
impared activities of daily living	946	25	1167	20	<0.01
diabetes	416	11	611	10	<0.01
heart disease	432	11	786	13	0.44
stroke	100	3	165	3	0.66
cancer	185	5	335	6	0.67
depression	576	15	796	13	<0.01

* p-values for test of the null hypothesis that the mean values for individuals included in the social security analysis equals the mean for individuals excluded from the social security analysis; calculated from t-tests, chi-square tests

Table 2. Correlations Among Social Security Reported Earnings and Self Reported Earnings

	Social Security Earnings (SSE)			
	1961	**1971**	**1981**	**1991**
SSE average 1967-1971	0.47	0.80	0.50	0.35
SSE average 1977-1981	0.37	0.60	0.77	0.49
SSE average 1987-1991	0.29	0.44	0.59	0.76
SSE average 1961,1971,1981,1991	0.49	0.63	0.69	0.60
1991 Self Reported Earnings	0.16	0.25	0.35	0.48

Sample sizes vary across correlations. All correlations are significant at α=.001 level.

Table 3. Odds ratios of reporting disease in 1992 for individuals reporting no social security earnings in indicated years as compared to individuals reporting some earnings

	ADL/IADL		Diabetes		Heart Disease		Stroke		Cancer		Depression	
	OR	**CI**	**OR**	**CI**	**OR**	**CI**	**OR**	**CI**	**OR**	**CI**	**OR**	**CI**
men												
1961	1.2	(.8 ,1.9)	1.2	(.7 ,1.9)	1.2	(.8 ,1.7)	1.5	(.7 ,3.1)	0.5	(.1 ,2.2)	1.4	(.9 ,2.1)
1971	1.0	(.6 ,1.6)	0.9	(.6 ,1.5)	1.0	(.6 ,1.5)	1.2	(.5 ,2.8)	1.4	(.6 ,3.2)	1.3	(.8 ,2.1)
1976	1.4	(.9 ,2.1)	1.0	(.6 ,1.6)	1.0	(.7 ,1.5)	1.8	(.9 ,3.5)	0.9	(.4 ,2.3)	1.1	(.7 ,1.8)
1977	1.2	(.8 ,1.7)	1.1	(.7 ,1.6)	1.0	(.7 ,1.4)	1.6	(.8 ,3.)	1.2	(.5 ,2.7)	**1.7**	(1.1 ,2.6)
1978	1.3	(.9 ,1.8)	1.1	(.7 ,1.6)	1.1	(.8 ,1.6)	**2.3**	(1.2 ,4.1)	0.6	(.2 ,1.6)	**2.1**	(1.4 ,3.)
1979	1.3	(.9 ,1.9)	0.9	(.6 ,1.4)	1.2	(.8 ,1.7)	**2.2**	(1.3 ,4.)	1.1	(.5 ,2.4)	**1.7**	(1.2 ,2.6)
1980	1.3	(.9 ,1.9)	0.9	(.6 ,1.4)	1.2	(.8 ,1.7)	**2.0**	(1.1 ,3.7)	1.5	(.7 ,2.9)	1.5	(1. ,2.2)
1981	**1.5**	(1.1 ,2.2)	1.1	(.7 ,1.6)	1.2	(.9 ,1.7)	1.4	(.7 ,2.5)	0.9	(.4 ,1.9)	**1.9**	(1.3 ,2.7)
1982	**2.0**	(1.4 ,2.7)	1.2	(.8 ,1.7)	1.3	(1. ,1.8)	**1.8**	(1. ,3.2)	1.1	(.5 ,2.1)	**2.1**	(1.5 ,3.)
1983	**1.9**	(1.4 ,2.6)	1.3	(1. ,1.9)	1.4	(1. ,1.8)	1.6	(.9 ,2.9)	0.7	(.3 ,1.4)	**2.0**	(1.4 ,2.7)
1984	**2.1**	(1.5 ,2.9)	1.4	(1. ,1.9)	1.2	(.9 ,1.7)	1.5	(.9 ,2.6)	1.2	(.7 ,2.2)	**2.3**	(1.7 ,3.2)
1985	**2.2**	(1.6 ,3.)	1.4	(1. ,2.)	1.4	(1. ,1.9)	1.4	(.8 ,2.5)	1.0	(.6 ,1.9)	**2.5**	(1.8 ,3.4)
1986	**2.5**	(1.8 ,3.4)	**1.7**	(1.2 ,2.3)	**1.4**	(1.1 ,1.9)	**2.0**	(1.2 ,3.4)	0.8	(.4 ,1.6)	**2.6**	(1.9 ,3.6)
1987	**2.3**	(1.7 ,3.1)	**1.5**	(1.1 ,2.)	**1.7**	(1.3 ,2.3)	**2.4**	(1.5 ,3.9)	1.4	(.8 ,2.4)	**2.3**	(1.7 ,3.1)
1988	**2.4**	(1.8 ,3.3)	1.4	(1. ,2.)	**1.6**	(1.2 ,2.1)	**2.6**	(1.6 ,4.2)	1.1	(.6 ,1.9)	**2.1**	(1.5 ,2.8)
1989	**2.0**	(1.5 ,2.6)	**1.8**	(1.3 ,2.4)	**1.9**	(1.5 ,2.5)	**2.6**	(1.6 ,4.2)	1.0	(.5 ,1.7)	**2.4**	(1.8 ,3.2)
1990	**2.4**	(1.8 ,3.2)	**1.7**	(1.2 ,2.3)	**1.9**	(1.5 ,2.5)	**2.7**	(1.7 ,4.3)	1.3	(.7 ,2.2)	**2.8**	(2.1 ,3.8)
1991	**2.3**	(1.7 ,3.)	**1.7**	(1.2 ,2.2)	**2.1**	(1.6 ,2.7)	**2.8**	(1.8 ,4.5)	1.3	(.8 ,2.2)	**2.9**	(2.2 ,3.9)
women												
1961	1.0	(.8 ,1.2)	1.1	(.8 ,1.5)	1.1	(.8 ,1.4)	1.2	(.7 ,2.2)	0.9	(.7 ,1.3)	0.9	(.7 ,1.1)
1971	0.9	(.8 ,1.2)	0.9	(.6 ,1.2)	0.8	(.6 ,1.1)	0.9	(.5 ,1.6)	0.7	(.5 ,1.)	0.8	(.6 ,1.)
1976	1.1	(.9 ,1.4)	1.1	(.8 ,1.5)	1.1	(.8 ,1.6)	1.3	(.7 ,2.4)	1.0	(.7 ,1.4)	1.2	(.9 ,1.6)
1977	1.2	(1. ,1.5)	1.4	(1. ,1.9)	1.2	(.8 ,1.6)	1.2	(.6 ,2.4)	0.9	(.6 ,1.2)	1.1	(.9 ,1.5)
1978	1.2	(1. ,1.6)	1.3	(1. ,1.9)	1.1	(.8 ,1.6)	1.1	(.6 ,2.3)	0.8	(.6 ,1.2)	1.2	(.9 ,1.6)
1979	1.2	(1. ,1.6)	**1.6**	(1.1 ,2.2)	1.2	(.9 ,1.7)	1.4	(.7 ,2.6)	0.8	(.6 ,1.2)	1.2	(.9 ,1.6)
1980	1.2	(1. ,1.6)	1.1	(.8 ,1.6)	**1.5**	(1.1 ,2.1)	1.9	(1. ,3.3)	0.8	(.5 ,1.2)	1.0	(.8 ,1.4)
1981	1.2	(.9 ,1.5)	1.0	(.7 ,1.5)	1.4	(1. ,1.9)	**2.2**	(1.2 ,3.9)	0.9	(.6 ,1.3)	1.3	(.9 ,1.7)
1982	1.2	(.9 ,1.6)	1.3	(.9 ,1.8)	**1.4**	(1.1 ,2.)	**2.5**	(1.4 ,4.3)	1.1	(.8 ,1.6)	**1.4**	(1.1 ,1.9)
1983	1.3	(1. ,1.6)	1.3	(.9 ,1.8)	**1.8**	(1.3 ,2.5)	**2.7**	(1.6 ,4.7)	1.1	(.7 ,1.6)	**1.5**	(1.1 ,2.)
1984	1.3	(1. ,1.7)	1.1	(.8 ,1.6)	**1.7**	(1.2 ,2.3)	**2.0**	(1.1 ,3.5)	1.2	(.8 ,1.8)	**1.7**	(1.3 ,2.3)
1985	**1.4**	(1.1 ,1.8)	1.2	(.8 ,1.7)	**2.0**	(1.4 ,2.7)	**2.1**	(1.2 ,3.7)	1.3	(.9 ,1.8)	1.3	(1. ,1.8)
1986	**1.5**	(1.2 ,1.9)	1.4	(1. ,2.)	**1.9**	(1.4 ,2.7)	**2.2**	(1.2 ,4.)	1.3	(.9 ,1.9)	1.3	(1. ,1.7)
1987	1.3	(1. ,1.7)	**1.6**	(1.2 ,2.3)	**2.1**	(1.5 ,2.9)	**2.3**	(1.3 ,4.3)	1.4	(1. ,2.)	**1.4**	(1.1 ,1.8)
1988	**1.4**	(1.1 ,1.8)	**1.7**	(1.2 ,2.4)	**2.2**	(1.6 ,3.)	**2.5**	(1.4 ,4.5)	1.4	(1. ,2.)	1.3	(1. ,1.7)
1989	**1.6**	(1.3 ,2.)	**1.6**	(1.1 ,2.2)	**2.3**	(1.7 ,3.1)	**2.2**	(1.2 ,4.1)	1.3	(.9 ,1.9)	**1.7**	(1.3 ,2.2)
1990	**1.4**	(1.1 ,1.8)	**1.6**	(1.1 ,2.2)	**2.6**	(1.9 ,3.5)	**2.3**	(1.3 ,4.2)	1.2	(.8 ,1.7)	**1.5**	(1.1 ,2.)
1991	**1.7**	(1.4 ,2.2)	1.4	(1. ,1.9)	**2.2**	(1.6 ,2.9)	**2.3**	(1.3 ,4.3)	1.2	(.9 ,1.8)	**1.7**	(1.3 ,2.1)

odds ratios signifcant at alpha = 0.05 are shown in bold type

Confidence intervals (CI) are 95%

models control for years of education, birthyear, race (black vs. white), Hispanic ethnicity, mother's education, father's education, father's occupational category, region of birth (Northeast, West, Midwest, South, non-US)

Table 4. Odds of disease in 1992-2002 with increasing earnings for self-reported earnings (SRE) in 1991 and Social Security Earnings (SSE) in 1991

	ADL/IADL				Diabetes				Heart Disease			
	SRE in 1991		SSE in 1991		SRE in 1991		SSE in 1991		SRE in 1991		SSE in 1991	
	OR	CI	OR	CI	OR	CI	OR	CI	OR	CI	OR	CI
all												
1992	**0.73**	(.67 .8)	**.87**	(.81 .94)	**.83**	(.74 .92)	.95	(.85 1.05)	.91	(.82 1.02)	.92	(.84 1.02)
1994	**0.75**	(.68 .83)	**.77**	(.71 .84)	.97	(.67 1.42)	1.11	(.8 1.56)	.96	(.87 1.07)	.92	(.84 1.01)
1996	**0.83**	(.75 .92)	.88	(.8 .96)	**.85**	(.76 .95)	.95	(.86 1.05)	.92	(.83 1.02)	.96	(.87 1.05)
1998	**0.79**	(.71 .88)	**.83**	(.76 .92)	**.83**	(.74 .93)	.92	(.84 1.02)	.9	(.82 .99)	.91	(.84 1.)
2000	**0.82**	(.73 .91)	**.79**	(.72 .87)	.87	(.79 .97)	.92	(.84 1.01)	.94	(.86 1.04)	.93	(.85 1.02)
2002	0.90	(.81 1.)	.88	(.8 .97)	.88	(.79 .98)	.92	(.84 1.)	.92	(.84 1.01)	.95	(.87 1.04)
men												
1992	**0.65**	(.56 .75)	**.8**	(.71 .89)	**.79**	(.68 .91)	.96	(.83 1.1)	.88	(.77 1.01)	.93	(.82 1.05)
1994	**0.71**	(.61 .83)	**.74**	(.65 .83)	1.11	(.55 2.27)	1.18	(.68 2.06)	.92	(.81 1.05)	.94	(.83 1.05)
1996	**0.79**	(.68 .92)	**.85**	(.79 .91)	.84	(.72 .98)	.97	(.85 1.12)	.87	(.76 .98)	.95	(.84 1.07)
1998	**0.67**	(.58 .77)	**.8**	(.7 .92)	**.8**	(.68 .93)	.91	(.8 1.03)	.89	(.78 1.01)	.97	(.86 1.09)
2000	**0.77**	(.66 .89)	**.8**	(.7 .91)	.83	(.72 .97)	.89	(.79 1.02)	.95	(.83 1.09)	.97	(.85 1.1)
2002	0.83	(.72 .95)	.87	(.76 .99)	**.82**	(.71 .95)	.87	(.77 .98)	.92	(.81 1.04)	.97	(.86 1.09)
women												
1992	**0.79**	(.71 .89)	.92	(.83 1.02)	.87	(.75 1.02)	.93	(.79 1.1)	.98	(.82 1.16)	.91	(.78 1.07)
1994	**0.79**	(.69 .91)	**.78**	(.69 .89)	.88	(.64 1.21)	1.03	(.71 1.49)	1.03	(.87 1.21)	.89	(.77 1.03)
1996	0.86	(.75 1.)	.9	(.79 1.02)	.85	(.73 .98)	.92	(.79 1.06)	1.	(.85 1.18)	.95	(.83 1.1)
1998	0.94	(.81 1.09)	.86	(.75 .99)	.86	(.74 1.)	.93	(.8 1.08)	.92	(.79 1.06)	.84	(.74 .96)
2000	0.85	(.73 .99)	**.78**	(.68 .88)	.91	(.78 1.05)	.95	(.82 1.1)	.94	(.82 1.09)	.89	(.78 1.02)
2002	0.97	(.84 1.13)	.9	(.78 1.03)	.94	(.81 1.09)	.98	(.84 1.14)	.93	(.81 1.07)	.94	(.83 1.06)

	Stroke				Cancer				Depression*			
	SRE in 1991		SSE in 1991		SRE in 1991		SSE in 1991		SRE in 1991		SSE in 1991	
	OR	CI	OR	CI	OR	CI	OR	CI	OR	CI	OR	CI
all												
1992	**0.70**	(0.59 0.83)	**0.73**	(0.62 0.85)	1.07	(0.92 1.23)	1.01	(0.88 1.16)	**0.80**	(0.72 0.89)	**0.84**	(0.77 0.92)
1994	**0.71**	(0.61 0.83)	**0.76**	(0.66 0.88)	1.07	(0.80 1.43)	1.09	(0.82 1.45)				
1996	**0.78**	(0.66 0.92)	0.81	(0.69 0.95)	1.23	(1.08 1.42)	1.09	(0.95 1.26)				
1998	**0.78**	(0.67 0.91)	**0.77**	(0.66 0.90)	1.10	(0.96 1.27)	1.03	(0.91 1.17)	**0.79**	(0.70 0.90)	**0.79**	(0.70 0.88)
2000	**0.82**	(0.71 0.94)	**0.81**	(0.70 0.95)	1.10	(0.97 1.25)	1.04	(0.93 1.17)	0.98	(0.82 1.16)	0.87	(0.76 1.00)
2002	0.88	(0.74 1.04)	0.86	(0.72 1.01)	1.15	(1.02 1.30)	1.05	(0.94 1.17)				
men												
1992	**0.66**	(0.53 0.81)	**0.68**	(0.56 0.83)	0.86	(0.70 1.07)	0.91	(0.72 1.14)	**0.71**	(0.61 0.83)	**0.75**	(0.66 0.85)
1994	**0.70**	(0.57 0.85)	**0.72**	(0.59 0.87)	0.91	(0.62 1.34)	1.02	(0.66 1.56)				
1996	0.81	(0.65 1.01)	0.80	(0.65 1.00)	1.19	(0.96 1.47)	1.02	(0.81 1.28)				
1998	0.82	(0.67 1.01)	0.81	(0.65 1.01)	1.08	(0.88 1.33)	0.98	(0.81 1.17)	**0.78**	(0.63 0.95)	**0.75**	(0.63 0.91)
2000	0.85	(0.70 1.03)	0.91	(0.72 1.15)	1.03	(0.87 1.23)	1.01	(0.84 1.20)	0.86	(0.64 1.15)	0.84	(0.69 1.03)
2002	0.90	(0.73 1.11)	0.92	(0.74 1.16)	1.10	(0.91 1.33)	1.00	(0.85 1.17)				
women												
1992	0.76	(0.57 1.03)	0.84	(0.65 1.08)	1.23	(1.02 1.50)	1.07	(0.89 1.27)	0.88	(0.77 1.01)	0.90	(0.80 1.02)
1994	**0.73**	(0.56 0.94)	0.84	(0.67 1.05)	1.27	(0.84 1.92)	1.17	(0.80 1.71)				
1996	0.77	(0.60 0.98)	0.84	(0.68 1.05)	1.27	(1.06 1.53)	1.13	(0.96 1.34)				
1998	**0.73**	(0.58 0.92)	**0.74**	(0.58 0.93)	1.12	(0.92 1.35)	1.08	(0.91 1.28)	0.81	(0.69 0.96)	**0.79**	(0.68 0.92)
2000	0.79	(0.64 0.97)	**0.75**	(0.62 0.92)	1.15	(0.96 1.38)	1.07	(0.91 1.25)	1.07	(0.87 1.32)	0.89	(0.73 1.07)
2002	0.84	(0.65 1.09)	0.81	(0.64 1.02)	1.20	(1.01 1.42)	1.09	(0.93 1.27)				

*missing odds ratios for depression are due to this data not being available in our HRS analytic dataset

both SRE and SSE are natural log tranformed earnings

Confidence intervals (CI) are 95%

odds ratios signifcant at alpha = 0.05 are shown in bold type

models control for years of education, gender (in non gender stratified model), birthyear, race (black vs. white), Hispanic ethnicity, mother's education, father's education, father's occupational category, region of birth (Northeast, West, Midwest, South, non-US)

Table 5. Among individual with diagnoses, mean differences between earnings in year prior to diagnosis and single years up to 15 years prior and post

	Years prior to (-) or post (+) dx	All diagnoses*				Major Conditions**			
		men (n=244)		women (n=291)		men (n=127)		women (n=172)	
		beta	**CI**	**beta**	**CI**	**beta**	**CI**	**beta**	**CI**
Years prior to diagnosis	-10	730	(-2691 , 4152)	826	(-911 , 2563)	536	(-4224 , 5296)	1844	(-543 , 4231)
	-9	239	(-2888 , 3365)	520	(-1087 , 2127)	72	(-4271 , 4416)	1165	(-1048 , 3379)
	-8	66	(-2777 , 2910)	1005	(-483 , 2494)	-380	(-4324 , 3565)	1661	(-392 , 3714)
	-7	426	(-2150 , 3002)	1050	(-323 , 2423)	-190	(-3758 , 3378)	1878	(-22 , 3778)
	-6	784	(-1539 , 3108)	961	(-304 , 2227)	-504	(-3716 , 2709)	1622	(-131 , 3375)
	-5	361	(-1731 , 2454)	550	(-621 , 1721)	-1263	(-4151 , 1625)	1337	(-286 , 2961)
	-4	933	(-966 , 2832)	184	(-915 , 1282)	166	(-2448 , 2781)	915	(-614 , 2443)
	-3	1229	(-518 , 2976)	60	(-984 , 1103)	282	(-2118 , 2682)	479	(-978 , 1937)
	-2	520	(-1129 , 2169)	-189	(-1198 , 820)	411	(-1850 , 2673)	57	(-1356 , 1470)
	-1 Year Prior to Diagnosis (reference year)								
Years after diagnosis	0	-409	(-2059 , 1240)	-423	(-1432 , 586)	-257	(-2519 , 2005)	-1023	(-2436 , 390)
	+1	-810	(-2558 , 939)	-439	(-1482 , 604)	-1368	(-3769 , 1034)	-780	(-2237 , 677)
	+2	-310	(-2213 , 1592)	-199	(-1297 , 900)	-699	(-3317 , 1920)	-654	(-2182 , 874)
	+3	-194	(-2295 , 1906)	80	(-1092 , 1252)	-163	(-3059 , 2734)	-329	(-1951 , 1294)
	+4	-262	(-2594 , 2070)	220	(-1040 , 1481)	-222	(-3443 , 3000)	244	(-1494 , 1981)
	+5	256	(-2334 , 2847)	618	(-744 , 1981)	575	(-3007 , 4157)	569	(-1300 , 2438)
	+6	1145	(-1751 , 4041)	642	(-851 , 2135)	933	(-3062 , 4929)	146	(-1890 , 2182)
	+7	406	(-2818 , 3631)	732	(-899 , 2363)	782	(-3659 , 5223)	459	(-1750 , 2668)
	+8	555	(-3002 , 4113)	351	(-1415 , 2117)	773	(-4141 , 5687)	-4	(-2392 , 2383)
	+9	1571	(-2307 , 5448)	-254	(-2163 , 1655)	1575	(-3780 , 6930)	-752	(-3319 , 1815)
	+10	1363	(-2851 , 5577)	173	(-1876 , 2221)	1381	(-4427 , 7190)	-72	(-2825 , 2681)

Confidence intervals (CI) are 95%. All models adjusted for age, age-squared, sex, race, Hispanic ethnicity, parental education, father's occupation, and years of education. Coefficients are expressed in dollars.

*All diagnoses include cancer, heart disease, stroke, broken bones, diabetes

**Major conditions include cancer, stroke or heart disease

Table 6. Among individual with diagnosis, mean differences between earnings in year prior to diagnosis and single years up to 15 years prior and pos

	All diagnoses*				Major Conditions**			
	men		women		men		women	
	beta	CI	beta	CI	beta	CI	beta	CI
income difference 5 years prior	-542.9	-(867.1 -,218.8)	-449.3	-(666.8 -,231.9)	-1462.0	-(1985.4 ,938.6)	-529.2	-(863. -,195.4)
income difference 5 years after	-1346.2	-(1813.6 -,878.7)	-643.2	-(934.6 -,351.7)	-4516.7	-(5192.2 -,3841.2)	-1297.8	-(1683.2 -,912.3)

Confidence intervals (CI) are 95%

*All diagnoses is cancer, heart disease, stroke, broken bones, diabetes

**Major conditions is cancer, stroke or heart disease

Table 7. Odds ratios of disease in 1992 by earnings at specific ages and averaged over age ranges

	ADL/IADL		Diabetes		Heart		Stroke		Cancer		Depression	
	OR	**CI**	**OR**	**CI**	**OR**	**CI**	**OR**	**CI**	**OR**	**CI**	**OR**	**CI**
all												
age 35	0.95	(0.87 1.04)	**0.85**	(0.77 0.94)	0.99	(0.89 1.10)	0.90	(0.74 1.10)	1.00	(0.87 1.15)	0.88	(0.81 0.97)
age 45	0.92	(0.85 1.00)	0.97	(0.88 1.07)	0.93	(0.84 1.02)	0.92	(0.75 1.12)	1.01	(0.87 1.18)	**0.81**	(0.75 0.88)
age 55	**0.84**	(0.75 0.94)	1.08	(0.92 1.26)	0.94	(0.81 1.08)	1.12	(0.83 1.52)	1.03	(0.83 1.29)	**0.77**	(0.68 0.87)
age 31-35 (avg)	0.94	(0.87 1.02)	0.92	(0.83 1.02)	1.01	(0.91 1.11)	0.83	(0.69 1.00)	0.99	(0.87 1.13)	0.88	(0.81 0.97)
age 41-45 (avg)	**0.85**	(0.79 0.92)	0.93	(0.85 1.02)	0.95	(0.87 1.04)	0.83	(0.70 0.99)	0.95	(0.83 1.07)	**0.81**	(0.75 0.88)
age 51-55 (avg)	**0.88**	(0.81 0.95)	0.92	(0.84 1.02)	**0.85**	(0.78 0.94)	**0.76**	(0.64 0.89)	0.90	(0.79 1.03)	**0.84**	(0.77 0.91)
men												
age 35	0.86	(0.74 1.00)	0.91	(0.77 1.07)	1.00	(0.86 1.17)	0.81	(0.62 1.07)	0.96	(0.71 1.31)	**0.77**	(0.66 0.89)
age 45	0.86	(0.76 0.97)	1.03	(0.89 1.18)	0.89	(0.79 1.01)	0.87	(0.67 1.12)	0.77	(0.62 0.96)	**0.73**	(0.65 0.82)
age 55	**0.74**	(0.62 0.87)	0.95	(0.78 1.17)	0.86	(0.72 1.03)	0.97	(0.69 1.34)	0.90	(0.61 1.34)	**0.74**	(0.62 0.89)
age 31-35 (avg)	**0.81**	(0.69 0.95)	0.97	(0.81 1.17)	1.07	(0.91 1.27)	0.74	(0.57 0.96)	0.88	(0.63 1.23)	**0.79**	(0.67 0.93)
age 41-45 (avg)	**0.75**	(0.67 0.86)	0.91	(0.79 1.04)	0.95	(0.84 1.07)	0.79	(0.63 1.00)	0.81	(0.65 1.01)	**0.72**	(0.63 0.81)
age 51-55 (avg)	**0.77**	(0.68 0.87)	0.87	(0.77 1.00)	**0.84**	(0.75 0.95)	**0.70**	(0.58 0.86)	0.82	(0.66 1.01)	**0.77**	(0.68 0.86)
women												
age 35	0.99	(0.89 1.10)	**0.81**	(0.71 0.93)	0.97	(0.84 1.12)	1.00	(0.78 1.28)	1.01	(0.87 1.18)	0.96	(0.85 1.08)
age 45	0.97	(0.87 1.08)	0.92	(0.80 1.05)	1.02	(0.87 1.19)	1.00	(0.73 1.38)	1.21	(1.02 1.44)	0.93	(0.82 1.05)
age 55	0.94	(0.81 1.10)	**1.36**	(1.10 1.69)	1.13	(0.92 1.40)	1.52	(0.84 2.75)	1.09	(0.83 1.42)	**0.79**	(0.67 0.92)
age 31-35 (avg)	0.98	(0.89 1.08)	0.88	(0.78 1.01)	0.96	(0.85 1.08)	0.91	(0.72 1.15)	1.01	(0.88 1.16)	0.92	(0.83 1.03)
age 41-45 (avg)	0.91	(0.82 1.00)	0.95	(0.83 1.09)	0.95	(0.83 1.08)	0.86	(0.66 1.12)	1.00	(0.87 1.16)	0.90	(0.80 1.00)
age 51-55 (avg)	0.96	(0.87 1.07)	1.01	(0.87 1.18)	0.87	(0.75 1.00)	0.86	(0.63 1.18)	0.97	(0.82 1.14)	0.92	(0.81 1.03)

odds ratios signifcant at alpha = 0.05 are shown in bold type

both SRE and SSE are natural log tranformed earnings

Confidence intervals (CI) are 95%

models control for years of education, gender (in non-gender stratified model), birthyear, race (black vs. white), Hispanic ethnicity, mother's education, father's education, father's occupational category, region of birth (Northeast, West, Midwest, South, non-US)

Table 8. Odds ratios of disease in 1992 for self-report earnings (SRE) with lagged Social security earnings (SSE) for lagged years and averaged over multiple years

	ADL/IADL		Diabetes		Heart Disease		Stroke		Cancer		Depression	
	OR	**CI**	**OR**	**CI**	**OR**	**CI**	**OR**	**CI**	**OR**	**CI**	**OR**	**CI**
all												
1991 (sre)	**0.68**	(0.58 0.80)	**0.74**	(0.63 0.87)	0.88	(0.75 1.03)	**0.63**	(0.47 0.84)	0.98	(0.80 1.21)	**0.79**	(0.67 0.93)
1991 (sse)	**0.81**	(0.72 0.92)	0.91	(0.78 1.06)	0.91	(0.78 1.05)	**0.67**	(0.53 0.86)	0.95	(0.77 1.18)	**0.81**	(0.71 0.93)
1981 (sse)	**0.77**	(0.67 0.88)	0.86	(0.73 1.02)	0.98	(0.82 1.17)	0.90	(0.63 1.26)	0.85	(0.68 1.05)	**0.78**	(0.68 0.91)
1971 (sse)	0.94	(0.80 1.10)	0.86	(0.72 1.03)	0.96	(0.78 1.18)	0.70	(0.47 1.04)	0.95	(0.76 1.18)	0.83	(0.70 0.97)
1961 (sse)	0.89	(0.79 1.01)	0.89	(0.76 1.03)	0.96	(0.83 1.10)	0.94	(0.65 1.35)	1.02	(0.86 1.22)	0.86	(0.74 1.00)
1987-1991 (avg sse)	0.83	(0.72 0.95)	0.91	(0.75 1.09)	0.94	(0.78 1.13)	0.76	(0.56 1.03)	0.87	(0.68 1.12)	**0.80**	(0.68 0.94)
1977-1981 (avg sse)	0.84	(0.70 1.01)	0.88	(0.69 1.11)	0.90	(0.73 1.13)	0.70	(0.45 1.07)	0.86	(0.65 1.14)	**0.74**	(0.61 0.89)
1967-1971 (avg sse)	0.87	(0.74 1.03)	0.81	(0.65 1.01)	0.98	(0.77 1.24)	0.78	(0.47 1.29)	1.09	(0.83 1.44)	0.87	(0.71 1.06)
1961, 1971, 1981, 1991 (avg sse)	**0.67**	(0.54 0.83)	0.74	(0.57 0.95)	0.87	(0.68 1.13)	**0.55**	(0.35 0.88)	0.86	(0.61 1.21)	**0.62**	(0.49 0.78)
men												
1991 (sre)	**0.64**	(0.52 0.79)	**0.78**	(0.64 0.94)	0.88	(0.73 1.05)	**0.65**	(0.46 0.92)	0.89	(0.69 1.15)	0.83	(0.68 1.02)
1991 (sse)	**0.77**	(0.66 0.89)	0.97	(0.60 1.17)	0.93	(0.79 1.10)	**0.68**	(0.51 0.90)	0.86	(0.65 1.13)	0.82	(0.69 0.97)
1981 (sse)	**0.76**	(0.63 0.92)	0.99	(0.77 1.26)	1.05	(0.85 1.30)	0.80	(0.56 1.15)	0.71	(0.53 0.97)	0.83	(0.67 1.03)
1971 (sse)	1.07	(0.81 1.40)	1.00	(0.70 1.42)	0.94	(0.73 1.21)	0.62	(0.38 1.01)	0.87	(0.57 1.35)	0.91	(0.69 1.22)
1961 (sse)	0.88	(0.73 1.06)	0.91	(0.75 1.10)	0.94	(0.80 1.10)	0.87	(0.59 1.29)	1.20	(0.83 1.73)	0.82	(0.66 1.01)
1987-1991 (avg sse)	**0.74**	(0.63 0.88)	0.93	(0.74 1.17)	0.94	(0.76 1.15)	0.71	(0.51 0.99)	0.77	(0.54 1.10)	0.80	(0.65 0.99)
1977-1981 (avg sse)	0.84	(0.66 1.08)	1.08	(0.76 1.54)	0.97	(0.75 1.26)	**0.58**	(0.36 0.92)	0.66	(0.44 0.98)	**0.71**	(0.54 0.91)
1967-1971 (avg sse)	0.90	(0.66 1.23)	0.85	(0.59 1.25)	0.87	(0.64 1.18)	0.61	(0.31 1.19)	1.09	(0.66 1.80)	0.83	(0.59 1.17)
1961, 1971, 1981, 1991 (avg sse)	**0.63**	(0.48 0.83)	0.89	(0.64 1.26)	0.90	(0.77 1.20)	**0.48**	(0.29 0.82)	0.73	(0.43 1.25)	**0.64**	(0.47 0.88)
women												
1991 (sre)	0.77	(0.59 0.99)	**0.64**	(0.47 0.89)	0.89	(0.59 1.35)	0.53	(0.24 1.17)	1.13	(0.82 1.54)	0.75	(0.57 0.98)
1991 (sse)	0.92	(0.73 1.16)	0.75	(0.56 1.01)	0.82	(0.56 1.19)	0.74	(0.40 1.36)	1.09	(0.82 1.43)	0.79	(0.63 0.99)
1981 (sse)	0.76	(0.61 0.96)	**0.73**	(0.58 0.93)	0.83	(0.61 1.14)	1.51	(0.61 3.69)	0.93	(0.70 1.24)	**0.73**	(0.58 0.90)
1971 (sse)	0.89	(0.73 1.10)	0.77	(0.60 0.98)	1.05	(0.72 1.52)	0.90	(0.48 1.69)	0.95	(0.73 1.23)	0.77	(0.62 0.96)
1961 (sse)	0.89	(0.75 1.07)	0.84	(0.56 1.09)	1.08	(0.80 1.46)	1.07	(0.45 2.58)	0.94	(0.74 1.18)	0.96	(0.76 1.20)
1987-1991 (avg sse)	1.03	(0.82 1.30)	0.83	(0.60 1.15)	1.01	(0.64 1.58)	1.36	(0.64 2.89)	1.00	(0.74 1.36)	0.78	(0.61 1.01)
1977-1981 (avg sse)	0.84	(0.63 1.11)	**0.65**	(0.46 0.90)	0.79	(0.54 1.15)	1.63	(0.73 3.67)	1.04	(0.74 1.46)	0.80	(0.61 1.04)
1967-1971 (avg sse)	0.86	(0.69 1.06)	0.76	(0.57 1.02)	1.12	(0.74 1.70)	1.01	(0.52 1.97)	1.09	(0.79 1.50)	0.86	(0.66 1.12)
1961, 1971, 1981, 1991 (avg sse)	0.70	(0.50 0.97)	**0.52**	(0.33 0.80)	0.89	(0.52 1.54)	0.95	(0.38 2.40)	0.91	(0.60 1.40)	**0.59**	(0.41 0.86)

Both SRE and SSE are natural log tranformed earnings

Confidence intervals (CI) are 95%. Odds ratios significant at alpha = 0.05 are shown in bold type

Models control for years of education, gender (in non gender stratified model), birthyear, race (black vs. white), Hispanic ethnicity, mother's education, father's education, father's occupational category, region of birth (Northeast, West, Midwest, South, non-US)

Table 9. Odds ratios of disease in 1992 with nested models controlling for additional measures of earnings and wealth

	OR	CI	OR	CI	OR	CI	OR	CI	OR	CI	OR	CI	OR	CI
ADL/IADL														
race (black vs white)	**1.56**	(1.27 ,1.92)	**1.39**	(1.08 ,1.79)	**1.40**	(1.07 ,1.82)	**1.35**	(1.03 ,1.76)	**1.36**	(1.03 ,1.8)	**1.40**	(1.09 ,1.8)	**1.31**	(1.01 ,1.7)
education (years)	**0.83**	(.8 ,.86)	**0.86**	(.82 ,.89)	**0.86**	(.82 ,.89)	**0.86**	(.82 ,.9)	**0.83**	(.79 ,.88)	**0.86**	(.82 ,.89)	**0.86**	(.82 ,.9)
gender (male vs female)	**0.48**	(.41 ,.57)	**0.55**	(.45 ,.67)	**0.54**	(.43 ,.67)	**0.55**	(.44 ,.69)	**0.49**	(.38 ,.62)	**0.53**	(.43 ,.66)	**0.53**	(.43 ,.66)
physical activity	**1.23**	(1.15 ,1.32)	**1.16**	(1.08 ,1.26)	**1.17**	(1.08 ,1.26)	**1.17**	(1.08 ,1.27)	**1.16**	(1.07 ,1.27)	**1.17**	(1.08 ,1.26)	**1.16**	(1.08 ,1.26)
self-report wage 1991			**0.74**	(.67 ,.81)	**0.67**	(.58 ,.78)	**0.76**	(.68 ,.84)	**0.73**	(.66 ,.82)	**0.72**	(.65 ,.8)	**0.73**	(.66 ,.81)
social security wage 1991					**1.15**	(1. ,1.32)								
social security wage 1981							0.90	(.8 ,1.)						
social security wage 1971									1.07	(.96 ,1.21)				
SSE average (1971,1981,1991)											1.05	(.91 ,1.21)	1.07	(.92 ,1.23)
wealth 1991													0.97	(.94 ,1.)
Diabetes														
race (black vs white)	**1.71**	(1.33 ,2.2)	**1.82**	(1.34 ,2.46)	**1.68**	(1.21 ,2.32)	**1.58**	(1.15 ,2.18)	**1.82**	(1.3 ,2.55)	**1.82**	(1.35 ,2.47)	**1.68**	(1.22 ,2.31)
education (years)	0.98	(.94 ,1.01)	1.00	(.95 ,1.05)	1.00	(.95 ,1.05)	0.98	(.94 ,1.04)	1.01	(.95 ,1.06)	1.00	(.95 ,1.05)	1.01	(.96 ,1.06)
gender (male vs female)	**1.33**	(1.09 ,1.61)	**1.48**	(1.15 ,1.89)	**1.44**	(1.12 ,1.86)	**1.48**	(1.13 ,1.95)	**1.45**	(1.08 ,1.95)	**1.49**	(1.15 ,1.92)	**1.47**	(1.14 ,1.91)
physical activity	**1.18**	(1.09 ,1.27)	**1.13**	(1.04 ,1.23)	**1.14**	(1.04 ,1.25)	**1.13**	(1.03 ,1.24)	**1.15**	(1.05 ,1.26)	**1.14**	(1.04 ,1.24)	**1.13**	(1.04 ,1.23)
self-report wage 1991			**0.83**	(.75 ,.93)	**0.76**	(.66 ,.89)	**0.85**	(.75 ,.95)	**0.81**	(.72 ,.91)	**0.84**	(.75 ,.94)	**0.85**	(.76 ,.96)
social security wage 1991					1.15	(.99 ,1.34)								
social security wage 1981							0.95	(.84 ,1.08)						
social security wage 1971									0.99	(.86 ,1.13)				
SSE average (1971,1981,1991)											0.98	(.84 ,1.15)	1.00	(.86 ,1.18)
wealth 1991													**0.95**	(.92 ,.99)
Heart Disease														
race (black vs white)	0.88	(.68 ,1.14)	0.80	(.57 ,1.11)	0.76	(.53 ,1.08)	0.69	(.48 ,1.)	0.76	(.53 ,1.1)	0.81	(.58 ,1.12)	0.77	(.55 ,1.09)
education (years)	0.97	(.94 ,1.01)	1.00	(.96 ,1.04)	1.00	(.95 ,1.05)	0.99	(.95 ,1.04)	1.00	(.96 ,1.05)	1.00	(.96 ,1.04)	1.00	(.96 ,1.05)
gender (male vs female)	**1.82**	(1.53 ,2.16)	**2.02**	(1.61 ,2.53)	**2.02**	(1.59 ,2.57)	2.10	(1.64 ,2.7)	**1.92**	(1.44 ,2.54)	**2.01**	(1.59 ,2.54)	**2.00**	(1.58 ,2.53)
physical activity	**1.11**	(1.04 ,1.19)	1.04	(.96 ,1.12)	1.03	(.95 ,1.11)	1.04	(.96 ,1.13)	1.04	(.96 ,1.13)	1.04	(.96 ,1.12)	1.03	(.96 ,1.11)
self-report wage 1991			0.92	(.82 ,1.02)	0.94	(.8 ,1.11)	0.93	(.83 ,1.06)	**0.87**	(.77 ,.98)	0.91	(.81 ,1.03)	0.92	(.81 ,1.04)
social security wage 1991					1.01	(.86 ,1.17)								
social security wage 1981							0.96	(.84 ,1.09)						
social security wage 1971									1.10	(.95 ,1.27)				
SSE average (1971,1981,1991)											1.02	(.86 ,1.2)	1.03	(.87 ,1.21)
wealth 1991													0.98	(.94 ,1.01)
Stroke														
race (black vs white)	1.38	(.86 ,2.23)	1.36	(.74 ,2.51)	1.41	(.74 ,2.66)	1.35	(.68 ,2.68)	1.28	(.67 ,2.47)	1.36	(.74 ,2.52)	1.11	(.57 ,2.19)
education (years)	0.99	(.92 ,1.06)	1.05	(.93 ,1.17)	1.03	(.92 ,1.16)	1.05	(.92 ,1.19)	1.01	(.89 ,1.14)	1.05	(.93 ,1.18)	1.07	(.95 ,1.19)
gender (male vs female)	1.43	(.98 ,2.07)	**1.68**	(1. ,2.81)	1.64	(.96 ,2.8)	1.88	(.99 ,3.59)	1.68	(.91 ,3.08)	1.70	(.99 ,2.92)	1.65	(.97 ,2.82)
physical activity	**1.17**	(1. ,1.37)	1.03	(.85 ,1.24)	0.99	(.82 ,1.21)	1.06	(.86 ,1.31)	1.12	(.9 ,1.39)	1.03	(.85 ,1.24)	1.01	(.84 ,1.22)
self-report wage 1991			**0.70**	(.59 ,.83)	0.82	(.63 ,1.07)	**0.69**	(.57 ,.83)	**0.66**	(.55 ,.8)	**0.70**	(.58 ,.85)	**0.72**	(.59 ,.88)
social security wage 1991					**0.78**	(.63 ,.98)								
social security wage 1981							1.08	(.77 ,1.51)						
social security wage 1971									0.89	(.69 ,1.14)				
SSE average (1971,1981,1991)											0.98	(.72 ,1.32)	1.02	(.74 ,1.41)
wealth 1991													**0.91**	(.86 ,.97)
Cancer														
race (black vs white)	0.71	(.48 ,1.06)	0.72	(.45 ,1.15)	**0.56**	(.34 ,.94)	0.66	(.4 ,1.11)	**0.54**	(.31 ,.94)	0.72	(.45 ,1.15)	0.70	(.43 ,1.15)
education (years)	1.00	(.95 ,1.05)	1.00	(.94 ,1.06)	0.99	(.93 ,1.06)	1.00	(.94 ,1.07)	1.01	(.94 ,1.08)	1.00	(.94 ,1.06)	1.00	(.94 ,1.06)
gender (male vs female)	**0.37**	(.28 ,.48)	**0.36**	(.26 ,.51)	**0.33**	(.23 ,.47)	**0.38**	(.26 ,.56)	**0.32**	(.22 ,.47)	**0.36**	(.26 ,.52)	**0.37**	(.26 ,.52)
physical activity	0.97	(.88 ,1.06)	0.98	(.89 ,1.09)	1.00	(.9 ,1.12)	1.02	(.92 ,1.15)	0.96	(.86 ,1.08)	0.98	(.89 ,1.09)	0.98	(.89 ,1.09)
self-report wage 1991			1.06	(.92 ,1.23)	1.13	(.94 ,1.35)	1.06	(.89 ,1.27)	1.05	(.88 ,1.25)	1.07	(.9 ,1.28)	1.08	(.9 ,1.28)
social security wage 1991					0.97	(.81 ,1.16)								
social security wage 1981							0.98	(.83 ,1.17)						
social security wage 1971									0.97	(.83 ,1.13)				
SSE average (1971,1981,1991)											0.97	(.78 ,1.22)	0.98	(.78 ,1.22)
wealth 1991													0.99	(.94 ,1.04)
Depression														
race (black vs white)	**1.27**	(1. ,1.6)	1.26	(.93 ,1.69)	1.29	(.94 ,1.77)	1.23	(.89 ,1.68)	1.12	(.8 ,1.56)	1.25	(.93 ,1.69)	1.03	(.75 ,1.41)
education (years)	**0.89**	(.86 ,.92)	**0.92**	(.87 ,.97)	**0.92**	(.87 ,.96)	**0.91**	(.86 ,.96)	**0.93**	(.88 ,.98)	**0.92**	(.88 ,.97)	**0.94**	(.89 ,.99)
gender (male vs female)	**0.75**	(.63 ,.89)	**0.75**	(.59 ,.94)	**0.70**	(.55 ,.89)	0.78	(.6 ,1.01)	**0.65**	(.49 ,.86)	**0.77**	(.61 ,.99)	**0.75**	(.59 ,.97)
physical activity	**1.31**	(1.21 ,1.42)	**1.18**	(1.07 ,1.29)	1.17	(1.07 ,1.29)	1.15	(1.05 ,1.27)	**1.19**	(1.07 ,1.32)	**1.18**	(1.07 ,1.29)	1.16	(1.06 ,1.28)
self-report wage 1991			**0.81**	(.73 ,.9)	0.88	(.75 ,1.04)	**0.85**	(.75 ,.96)	**0.80**	(.71 ,.9)	**0.83**	(.73 ,.93)	**0.85**	(.75 ,.96)
social security wage 1991					0.95	(.82 ,1.09)								
social security wage 1981							0.89	(.79 ,1.)						
social security wage 1971									0.96	(.86 ,1.08)				
SSE average (1971,1981,1991)											0.93	(.8 ,1.08)	0.97	(.83 ,1.14)
wealth 1991													**0.91**	(.88 ,.93)

Both SRE and SSE are natural log tranformed earnings. Confidence intervals (CI) are 95%. Odds ratios signifcant at alpha = 0.05 are shown in bold type.
In addition to variables as indicated in the above table, models control for birthyear, Hispanic ethnicity, mother's education, father's education, father's occupational category, and region of birth (Northeast, West, Midwest, South, non-US)