Aging and inequality in income and health

Angus S. Deaton

and

Christina H. Paxson

Research Program in Development Studies
Princeton University

December 14, 1997
This version: January 9, 1998

This paper was prepared for the AEA Session on Demographic Trends and Economic Consequences on January 3, 1998 in Chicago. The research was supported by the National Institute of Aging through grants P01-AG05842 and R01-AG11957.
In our previous work, Deaton and Paxson (1994, 1997), we showed that, in a large group of countries, inequality in consumption increases with age within cohorts of individuals. This finding was motivated by a well-known feature of standard autarkic intertemporal choice models, that under appropriate assumptions consumption follows a martingale, see Hall (1978). The theory implies that, provided that shocks to consumption are not perfectly correlated across individuals, within-cohort consumption inequality should rise over time as cohorts age. The same should be true of income, at least up to the date of retirement, and of earnings, if employers pay workers their expected marginal product, see Farber and Gibbons (1996).

More recently we have examined whether inequality in health status also increases with age, and how the joint distribution of health and income evolve over the life-cycle. It is plausible that health shocks have both permanent and transitory components. The presence of the former implies that health status will be non-stationary so that, provided health shocks are not perfectly correlated across individuals, the dispersion of health status with grow with age. This view of health status as a non-stationary random variable is consistent with stress models in which health is the result of “the piling up of adverse life experiences,” Ryff and Singer (1997, p. 90.)

Health status, along with income and consumption, is an important determinant of welfare, so that our interest in health inequality stems from a more general interest in the distribution of welfare. Furthermore, health is not independent of economic status. There is a well-documented but poorly understood “gradient” linking socioeconomic status to a wide range of health outcomes, see Adler et al (1994) and McIntyre (1997) for reviews. The gradient has both a life-cycle and a temporal component; differences in mortality across socioeconomic groups are widest in late middle age, Kitigawa and Hauser (1973), Duleep (1995), and Elo and Preston (1996), and are increasing over time, Feldman et al 1989, Pappas et al (1993), and Elo and Preston (1995).
In our earlier work, we used data from the National Health Interview Survey (NHIS) to examine life-cycle patterns in health status and in the joint distribution of health status and income (Deaton and Paxson, 1998). In this paper we summarize and extend those results, and provide new evidence from the Panel Study of Income Dynamics (PSID). Both surveys contain a measure of household income, and collect information on an ordinal measures of self-reported health status (SRHS) that ranges from 1 (excellent) to 5 (poor). Section I concerns problems related to the measurement of inequality in health. Section II presents evidence from the two surveys on health and income inequality.

I. Measurement Issues

The measurement of health inequality raises two important issues. The first is the difficulty of identifying a measure of health status that is useful over the complete adult life-cycle. For example, measures of the inability to complete “activities of daily living,” (ADLs) such as dressing or bathing, have been fruitfully used to assess the health of the elderly. However, these measures do not adequately capture health differences among younger people. Self-reported “days of illness” or “doctor visits,” are themselves conditioned by socio-economic status, and sometimes show perverse correlations with income, with better-off people apparently perceiving and treating their illnesses more seriously. The properties of the measure of self-reported health status used in this paper have been studied extensively. First, it predicts subsequent mortality. A large number of studies that use data from a variety of countries indicate that reports of poor health are significantly related to subsequent mortality, see Idler and Kasl (1995) for summary of this research. The correlation between SRHS and subsequent mortality remains strong even after controlling for objective measures of health status (obtained from doctor’s examinations) and life-style factors
such as smoking. This fact has led some to argue that SRHS is itself an independent determininant of longevity: individuals with healthier self images live longer. An alternative to these psycho-social explanations is that individuals have information about their health that is unobserved by others, including physicians. For our purposes, it is the raw correlations between self-rated health and mortality that are of interest, since we want to identify a variable that can serve as a single summary measure of health status. Other research has found that those with low SRHS are more likely to develop problems with ADLs (Idler and Kasl, 1995) and miss more work due to illness (Marmot et. al., 1995.)

Once a measure of health is identified, the second issue is how to measure inequality in health status. Although it is straightforward to compute measures of dispersion in SRHS, it is not clear how we should judge such measures in terms of social welfare. Consider, for example, the familiar result that if a distribution $F_1$ (second order) stochastically dominates a distribution $F_2$, then $F_1$ will result in higher social welfare, when social welfare is represented as the integral over the population of a monotone increasing and concave function of the variable in question. Although we are used to the assumption that social welfare is increasing and concave in income or consumption, it is much less clear why it should be increasing and concave in an ordinal self-reported measure of health status. Nevertheless, the literature on SRHS provides some support for the idea that changes in SRHS have a larger effect on mortality when SRHS is “poor” than when it is “excellent.” If so, a mean-preserving spread in SRHS will lower average life expectancy, and will lower the expected value of any function that is concave in life expectancy, for example one that prefers a decrease in infant mortality to an increase in longevity at older ages. Of course, to focus solely on life expectancy ignores the quality of life. SRHS may well give
some indication of quality as well as the likely length of life, so that changes in the distribution of SRHS could still have adverse welfare consequences even in the absence of a relationship between SRHS and mortality. We also note that much of the literature on health inequality is not concerned with inequality in years lived, but with the inequalities in health outcomes across socioeconomic groups. That these are quite different has been emphasized by Wilkinson (1986) who points out that, over the 20th century in Britain, socioeconomic differences in mortality have increased while the inequality in years lived has decreased, essentially because of the decline in infant mortality.

II. Evidence on life-cycle patterns in the distribution of health and income

The NHIS is an annual survey of approximately 50,000 adults (plus children) that collects information on health, illnesses, doctors visits, spells of hospitalization, and basic socioeconomic characteristics. We use data on all adults between the ages of 20 and 70, inclusive, interviewed from 1983 through 1994. The survey provides sampling weights, which we use, so that the results should be representative of the whole US population. The PSID is a panel survey of households that has been in existence since 1968, and since 1984 has collected information on the self-reported health status of household heads and their spouses. We use a sample of 3,435 men and 4,561 women who were either heads of households or their spouses in all years between 1984 and 1992. Unlike the NHIS, this is not a nationally representative sample of all adults, both because it is only households heads and spouses, and because the PSID oversampled poor households in 1968. Under these circumstances, we did not use sampling weights with the PSID.

Although the NHIS has much larger sample sizes and more extensive health information than the PSID, it has far less information on income. The measure of family income in the NHIS is
bracketed and is top-coded at $50,000 in nominal dollars. The brackets are narrow and are not a major concern, but such serious top-coding cannot be ignored in computing measures of dispersion. In the NHIS results that follow, we have used the Tobit procedure described in Deaton and Paxson (1998), but one reason for extending our work to the PSID is to attempt to reproduce our results with much higher-quality data on income.

Our approach is to track the moments and co-moments of SRHS and family income over time for individuals from the same birth cohort. The NHIS is large enough for each cohort to be defined by the exact year of birth; for the PSID we define cohorts using non-overlapping five-year birth intervals. Cohorts are identified by their age (or, for the PSID, the midpoint of the age range) in the earliest year of observation, 1983 for the NHIS and 1984 for the PSID. There are 62 cohorts for the NHIS, and 9 for the PSID. It should be kept in mind that the PSID is used to construct “true” cohorts: we actually follow the same individuals over time as they age. With the NHIS, we track randomly-selected representatives from the (still-living) populations of people born in the same year. These populations are not fixed, because some group members die each year. The evolution of inequality in health and income with age will reflect both changes in inequality within a fixed group, and the effects caused by selection of some members, though death, out of the group.

We first compute moments (mean and variances) and co-moments of health status and income, for each cohort in each year, and for men and women separately. These become the “raw data” for our analysis, and much can be learned by looking at graphs of these data. Figure 1 shows the cohort plots for males and females from the PSID; the same information for the NHIS is in Figure 4 of Deaton and Paxson (1998). It shows the age-profiles for the mean of SRHS (top
panel), the variance of health status (middle panel), and the correlation between health status and income (bottom panel). Each line on the graph shows the experience of a single cohort over time.

Note first that, as expected, average health status deteriorates with age for both men and women, although women report worse health than men at younger ages. That SRHS worsens with age is perhaps to be expected, but it implies that when people report their health status, they do not completely “norm” their answers with respect to the experience of those at the same age. The patterns of SRHS with age in the NHIS are similar except health is better on average at all ages for men and women, which is perhaps not surprising given the oversampling of poor households by the PSID. Second, inequality in health increases with age and the results for the PSID in the middle panel of Figure 1 are consistent with the evidence from the NHIS. Although we do not show it, in both the PSID and the NHIS the dispersion in the joint distribution of income and health status rises with age up to the age of retirement and then levels off.

Third, the bottom panel shows a consistently negative correlation between health status (measured negatively) and the logarithm of family income, so that the gradient between mortality and income extends to SRHS. Moreover, and again in line with the literature, the correlation varies with age; it is small among those in their early 20s, but becomes steadily larger (in absolute value), reaching a peak value of around –0.4 between ages 50 and 60. The small sample sizes in the PSID (relative to the NHIS) yield only noisy measures of this correlation; to clarify the results, and to facilitate comparisons between the two data sets, we regressed the correlations between health and income on a set of age and cohort dummy variables. The cohort dummies account for the fact that the correlation between the two variables—the gradient—may differ across groups born in different years, while the age effects capture the life-cycle profile of the gradient, the shape of
which, by assumption, is held fixed across cohorts. Figure 2 shows the estimated age effects in both data sets.

The two data sets yield very similar patterns in the age-profile of the correlation of income and health. For both men and women, the correlation between SRHS and income weakens after age 60, as SRHS deteriorates in general. But this is not simply a matter of the elderly having uniformly poor health status. As the top panel of Figure 1 shows, health status deteriorates with age, but the middle panels does not show a collapse in the variance after age 60; instead, the fact is that, at older ages, differences in SRHS are less well-predicted by income.

There are several possible interpretations of these results, none of which necessarily excludes the other. One is that labor supply and earnings ability are adversely and cumulatively affected by health shocks, so that poor health and low income are increasingly correlated with age. The correlation may weaken in old age, since health shocks received after retirement will not affect pensions and social security (although they could affect asset income if sick people run down assets to pay for care.) There is also undoubtedly some causality running from income to health. Poorer people are more prone to life-styles with enhanced risk factors—obesity or cigarette smoking—and have less access to health care, including preventative health care, and live and work in less healthy environments. There is also a literature documenting the adverse health consequences of unemployment. The provision of Medicare at older ages could reduce the correlation by making one determinant of health—medical care—available to everyone. Sorting out the respective contributions of these various mechanisms remains an important task for future research.

Perhaps even more important than life-cycle patterns, is the question of changes in over time in the relationship between income and health status. It is difficult to discern cohort effects,
represented by upward or downward shifts in the traces for different cohorts, from a visual inspec- 
tion of Figure 1. However, a more systematic approach shows that there are significant 
differences across cohorts. We first examined the cohort effects from regressions of each of the 
variables (mean health, the variance of health, and the correlation of health and income) on sets of 
age and cohort dummy variables. This was done separately for men and women, and for the PSID 
and the NHIS. The estimated cohort effects are jointly significant at the 5 percent level or better 
for each of the variables and samples. To summarize the size and sign of these cohort effects, we 
then regressed each of the variables on a complete set of age dummies and a linear cohort trend. 
The results are as follows: First, for females from both data sets and males in the NHIS, average 
health status has improved over time across cohorts. The effect is small—approximately 0.004 
units per year of birth. However, a visual inspection of the cohort effects indicates that they are 
not linear. There has been no improvement, and possibly some deterioration, in health status 
across cohorts born after 1945, and there were larger improvements across those born before 
1945. The results for the males in the PSID are at odds with the other groups. The estimate of the 
cohort trend indicates that more recently-born groups are significantly less healthy, by about 
0.009 units per year of birth. This is largely due to declines in reported health, controlling for age, 
of the youngest four cohorts. These declines, which can be seen in the raw data graphed in the top 
panel of Figure 1, warrant further analysis. Second, for all of the samples, younger cohorts have a 
lower variance of health status.

Third, the results provide some support for the findings cited above that the gradient between 
income and health is becoming stronger over time. The coefficient on the cohort trend ranges 
from 0.001 (for females in the NHIS) to 0.003 per year (for males in the PSID), a positive sign
indicating that for more recently born cohorts there is a larger correlation between income and health. For example, the actual correlation between the logarithm of income and SRHS is -0.40 for PSID males born between 1940 and 1944, when they reached the age of around 50 in 1992. Our results imply that the correlation for the cohort born ten years later, between 1950 and 1954, will equal -0.43 when this group reaches the age of 50.
References


Macintyre, Sally, 1997, “The Black report and beyond: what are the issues?” Social Science and Medicine, 44, 723–45.


Figure 1: Health status, the variance of health status, and the correlation between health and income for cohorts in the PSID
Figure 2: Estimated age effects in the correlation between SRHS and log income—the gradient.