PARENTAL INCOME, LIFETIME INCOME, AND MORTALITY

Mårten Palme
Stockholm University

Sofia Sandgren
Uppsala University, Norwegian University of Science and Technology

Abstract
This article studies the relation between parental economic resources and mortality later in life. We use a data set on a cohort of individuals born in 1928 in the county of Malmö in southern Sweden, which contains exceptionally detailed measures of parental household income from five years during the individuals' childhood between 1929 and 1942. The data also contain very rich information on individual earnings throughout these individuals' entire life cycle that allows us to construct a measure of lifetime earnings. Date and cause of death are obtained from national registers. Using Cox proportional hazard models, we find an inverse relationship between parental income and mortality, also when controlling for individual lifetime income and when studying those with high education separately. A competing risk analysis shows the relation between parental income and mortality to apply to cancer as the cause of death. (JEL: D31, I10, I12, J10)

1. Introduction

The nonlinear inverse relation between the individual absolute income level and mortality found in several empirical studies has attracted interest in both academic and policy circles in recent years. An important background to this interest is the policy implication that income redistributions from rich to poor would increase the average life expectancy in a society (see, e.g., Deaton (2003) for an

The editor in charge of this article was Jordi Galí.

Acknowledgments: We thank three anonymous referees, John Bishop, Anders Björklund, Robert Erikson, Per Johansson, Erik Mellander, Torsten Persson, Denny Vågerö, as well as seminar participants at the Center for Health Equity Studies (CHESS) in Stockholm and the Swedish Institute for Social Research (SOFI), EALE/SOLE conference in San Francisco, EEA conference in Amsterdam, ENTER Jamboree meeting in Brussels, and the Nordic Summer Institute Meeting in Labor Economics in Uppsala for helpful comments and suggestions at different stages of this work. We thank the Institute of International Education at Stockholm University for letting us use their data. Mårten Palme gratefully acknowledges financial support from the Swedish Council of Working Life and Social Research (FAS) and Sofia Sandgren from the Swedish Agency for Innovation Systems (VINNOVA). Palme is a member of IZA.

E-mail addresses: Palme: Marten.Palme@ne.su.se; Sandgren: Sofia.Sandgren@nek.uu.se

Journal of the European Economic Association June 2008 6(4):890–911

overview). The mechanisms behind the relation are, however, not fully understood and there are alternative explanations for it, with very different policy implications. One "third factor," causing a spurious correlation between individual income and mortality (see, e.g., Smith (1999) for an overview), that has been suggested is exposure to poor living conditions very early in life, or even *in utero*, which in the medicine literature has been shown to affect mortality later in life (e.g., Barker 1997). The spurious correlation between income and mortality is then caused through the linkage in earnings across generations through intergenerational transmission of human capital.

A recent study, Case, Lubotsky, and Paxson (2002), shows that parental income has a protective effect on child health, and a subsequent study—Case, Fertig, and Paxson (2005)—shows differences in economic resources during childhood and adolescence to have lasting effects on both adult health status and earnings. Furthermore, van den Berg, Lindeboom, and Portrait (2006) show that economic recessions experienced early in life affect mortality later on. There is also a growing literature in epidemiology and medical sociology on parental social class and health status in adult life (e.g., Forsdahl 2002; Marmot et al. 2001; Smith et al. 1997).

Although a relation between parental socioeconomic status and health in the adult life has been documented in several empirical studies, the background to this relation is still controversial. Case, Fertig, and Paxson (2005) distinguishes between *life course* and *pathway* models for describing this relation. *Life course* models allow for both direct and indirect influences of childhood health and living conditions on adult health status and economic conditions. Direct, because children from economically deprived homes, with on average inferior health, are likely to enter their adult life with lower health status and ultimately higher mortality rate. Indirect, because children with inferior health may face more difficulties in acquiring education and the skills necessary for reaching a higher ranked position in the society as an adult.

Pathway models (see, e.g., Marmot et al. (2001) for empirical evidence), on the other hand, argue a causal relation between childhood and adult living conditions through intergenerational transmission of socioeconomic status. Subsequently, there is a separate causal relation between adult socioeconomic status and health. However, according to these models, there is no direct or indirect relation between childhood circumstances and health status later in life.

In this study we first analyze if there is a relationship between economic resources during childhood and mortality later in life. Following the previous discussion, such a relation may include both direct and indirect effects of childhood living conditions on mortality as well as, following the pathway model, be an effect of an essentially spurious correlation. Because our data include very detailed information about economic conditions during their entire life course, we are also able to evaluate to what extent there is a conditional effect of parental economic resources conditional on lifetime income and educational attainments—that is, a separate and lasting effect of childhood circumstances on mortality later in life.

We use a data set containing information on a cohort of men and women born in the county of Malmö in southern Sweden in 1928. The original interviews, focusing on cognitive ability and social background, were conducted in 1938. In this first interview, detailed data on parental family income and economic resources were obtained. Subsequently, there have been several re-interviews, the last one in 1993, primarily focusing on economic resources and education. In addition, tax register data on earnings have been obtained for twelve separate years since 1948, which enables us to construct a measure of lifetime earnings. Finally, data on date and cause of death until 2003 have been obtained from the National Cause of Death Register and matched onto the data. To our knowledge, this is the only longitudinal data set including individual information on income and socioeconomic conditions covering the individuals' entire life cycle in existence.

We use Cox proportional hazard models in the empirical analysis. We find that an inverse relation between parental economic resources and mortality persists when controlling for the measure of individual lifetime earnings as well as household income. Furthermore, we analyze high and low educational groups separately and find parental income to have a significant effect in the group with high education. Because this group of individuals are likely to have lived most their adult life in relatively good economic conditions, this result supports the interpretation that there is a persistent effect of parental income on mortality later in life. Finally, we use a competing risk model and analyze the two main causes of death, cancer and circulatory diseases, separately. A significant relation is found between parental income and death from cancer.

The rest of the article is organized as follows. Section 2 gives a detailed description of the data and Section 3 discusses the empirical specifications. Section 4 presents the results and Section 5 concludes.

2. Data and Measurement

The original sample of the Malmö Longitudinal Study consists of all boys and girls who attended third grade in any school in the county of Malmö in 1938.[1] The first round of interviews was done in 1938, when the majority of the individuals included were aged ten. The last round of data collection was made 55 years later, in 1993 when most of them had reached the normal retirement age at age 65. The original purpose of collecting the data was to investigate the relationship between

1. Malmö is the third largest city in Sweden, located in the South. The survey includes pupils in the city of Malmö and surrounding municipalities. In 1938, the population size was 151,247.

social background and cognitive ability.[2] Therefore, extensive social background information was collected, and an ability test distributed, for all children in the sample.

Table 1 shows some descriptive statistics of the data used in this study. The first column shows that 1,542 pupils were tested and interviewed in the original survey. It also shows that most of them, 1,342 (87%), were born in 1928, that is, were "normal aged," but that both over-aged and under-aged pupils were included. From the original survey, we use information on parental income collected on five

TABLE 1. Sample means.

	Original Sample $n = 1,542$	Sample 1 $n = 1,441$	Sample 2 $n = 1,064$	Sample 3 $n = 1,341$
Parental background variables:				
Parental income 1929–1942, means	-	5,181	5,063	5,012
		(8,125)	(6,608)	(6,378)
Educational level of father,%				
Primary school	45.7	45.1	44.1	45.1
Primary school and apprenticeship	37.2	37.4	38.0	37.7
Vocational	8.6	8.8	9.3	8.9
Lower secondary school	4.3	4.4	5.0	4.5
Upper secondary, or higher	4.1	4.3	3.6	3.8
Social class, %				
1	33.6	29.8	28.3	30.1
2	34.1	36.0	36.5	36.7
3	18.1	19.3	20.6	19.2
4	14.2	14.9	14.6	14.0
Individual variables:				
Year of birth, number of individuals				
<1928	171	143	84	128
1928	1,342	1,269	962	1,194
>1928	29	29	18	19
Educational level, %				
Primary school	47.4	45.8	34.1	44.4
Vocational school	25.3	26.0	31.6	26.8
Lower secondary school	14.6	15.3	18.6	15.6
Upper secondary school	8.2	8.3	10.1	8.5
Academic studies	4.4	4.6	5.6	4.7
Household income 1971, means	51,466	51,466	55,054	51,466
	(32,855)	(32,855)	(31,432)	(32,855)
Average early lifetime earnings		-	434,237	-
			(293,744)	
Average lifetime earnings	-	-	6,026,744	-
			(4,217,907)	
Deceased in December 2003, %	33.4	33.1	31.2	35.2

Standard deviation within parentheses.

2. The study was initiated by a doctoral student named Siver Hallgren, who wrote his thesis on this subject (see Hallgren 1939).

occasions between the years 1929 and 1942, father's education, and social class. Table 1 reports the mean and standard deviation for the parental income measure. It also reports information on the frequency distribution of father's education in five levels. Very few fathers, less than 10%, have an educational level beyond vocational schooling. Socioeconomic status was assigned in four categories on basis of four items of information: father's occupation, family income in 1937, number of children at home, and appearance in the social welfare registers of the Malmö schools. Table 1 also shows the sample percentage distribution of these levels.

There are two income measures of key importance in this study: the individuals' own income and parental income. Measuring parental economic resources during the individual's childhood involves several considerations. One of these is what time period should be included. Although there are likely to be some credit constraints, it is reasonable to assume that the household may smooth out temporary income changes. This suggests that income from a longer period than just one year is preferable.[3] There are two main sources of variation: variation over an individual's career and transitory variation between different years. Because our data set does not contain any information on parental age or job tenure, we are unable to deal with the former source of variation. However, because we have access to data on parental income from more than one year, we are able to account for income variation between years.

We use earnings information, collected by taxation rate-books for the years 1929, 1933, 1937, 1938, and 1942. Thus, we have information on parental income during the years between ages 1 and 14 for normal-aged individuals in this sample. For each year, there is missing information on incomes for between 100 and 200 families in the sample. There is information on income for each year for 1,082 families. There is missing information on one and two years for 258 and 101 families, respectively. For 101 families, there are missing values for three years or more of the five years included in the data.

We have chosen to exclude the 101 individuals with three years or more of missing parental income information. Descriptive statistics for the original sample of 1,542 individuals compared to the 1,441 remaining after this selection can be found in Table 1 (sample 1). For those with parental income missing for one or two years, we have used fixed effects models, including year effects estimated on the entire sample, to predict the missing years of parental income. We have then used CPI to obtain measures of real income for each year and, finally, averaged over the five years included.

We obtain three alternative measures for the individual's own income: lifetime earnings, early lifetime earnings, and household income in 1971. To construct the

3. Mazumder (2005) surveys different methods of considering income smoothing when measuring income across generations.

measure of lifetime earnings, we use information from the numerous follow-ups of the Malmö study, which have been done since 1938. Four questionnaires have been distributed, in 1964, 1971, 1984, and 1994. Among other things, the questionnaires have included occupational information for each year between 1942 and 1993. Register information on own income, family income, capital, and property has been collected with three- to five-year intervals between 1948 and 1993.[4]

In 1948, the normal-aged men and women in the study turned 20. However, from the questionnaires, we know that many of them started work as early as in 1942—when they were fourteen years old and had just finished primary school. Through the questionnaires, we also know what kind of job they had. We have carefully studied wage statistics for the relevant years (1942–1947) and assigned each individual the wage of a minor in that profession during that year.

A problem with the earnings data is that for the first years (1948, 1953, 1958, and 1963) there is missing information on earnings for some of the individuals, probably because only the county registers of Malmö were searched these first years. However, because we have occupational information, we have estimated wage equations on those for whom we do have earnings information. Then, we have used this information to predict the earnings for those without registered earnings, but with reported occupation (predictions are made for between 4% and 25% of the sample). For this estimation, we have used an extended version of the Mincer earnings function, with dummies for larger occupational categories, also including ability measures. For the years between the observed ones, we use the two surrounding years to interpolate assuming a linear increase of earnings. If the sample individual's first year on the labor market was 1950, for example, and there are data on earnings for 1953 and 1958, we have extrapolated from these figures. For a more thorough explanation of the lifetime earnings measure, see Sandgren (2007).

The advantage of lifetime earnings as a measure of economic well-being is that it considers both the problem of variation over the life cycle as well as transitory income variations over a shorter period of time (see, e.g., Mazumder 2005). The measure has three major disadvantages, however. The first is an obvious endogeneity problem when relating lifetime earnings to mortality due to the fact that those who die at a relatively young age contribute with fewer years of earnings and may, to a larger extent, be restricted from working full time for health reasons. We use two different strategies to avoid this problem. First, we predict earnings for the remaining years for those who died before reaching the normal retirement age. We use a simple model with fixed effects, age and age squared.

4. This means that we have this kind of data for the years 1948, 1953, 1958, 1963, 1968, 1971, 1974, 1978, 1982, 1986, 1999, and 1993.

Then, we inflate each year's earnings to the price level of 1993,[5] once more using the CPI, and add these to obtain a measure of each individual's earnings from the moment he or she entered the labor market and until reaching the age of 65. Second, we construct a measure of early lifetime earnings. For this purpose, we use exactly the same procedure as for the measure of lifetime earnings, but we only accumulate earnings up to the age of 43. At that age, only 36 of the individuals in the original sample had passed away.

A second problem with the lifetime earnings measure is that we have excluded those who did not answer the questionnaire in 1964 from both lifetime earnings measures, because in those cases, we do not know when the individual entered the labor market. In column 3 of Table 1, we show the statistics for this group (sample 2). This procedure creates an obvious attrition problem. By comparing columns 1 and 3, it can be seen that those who refused to answer the questionnaire in 1964, and were thus excluded from the sample, on average have lower earnings and obtained a lower educational level. They also have a higher mortality rate. However, comparing the social background variables, the differences are smaller and go both ways. For example, the original sample actually has a higher average family income in 1937, and more fathers with a higher education, than sample 2.

A third potential disadvantage of the lifetime earnings measure is that it does not take pooling of economic resources within the household into account. This aspect is likely to be most important for females since, in general, they have a lower labor force participation rate. As an alternative measure of the individual's own economic resources, we use household income in 1971, that is, when most individuals in our sample were aged 43. This income measure is obtained from tax registers and refers to the sum of all taxable income (*sammanräknad hushållsinkomst*) from both spouses in the household. An advantage of this income measure is that it does not rely on survey responses, but can be obtained for all men and women still alive by that year and living in Sweden. This enables us to extend the sample as compared to that used for lifetime income, because we only need to exclude the 36 individuals in the sample who died before 1 January 1972.

The obvious disadvantage of our measure of household income is that it neglects both life-cycle variation and transitory variation. However, using data from the Health and Retirement Study complemented with earnings histories from the Social Security registers, Haider and Solon (2006) suggest there to be a very high rate of correlation between income around the age of 40 and lifetime income. Böhlmark and Lindquist (2006) confirm this general conclusion on Swedish data. Column 4 in Table 1 shows descriptive statistics for this extended sample (sample 3).

We have supplemented the data from the Malmö study with information on mortality from a national register provided by the National Board of Health

5. The year the normal-aged individual reached the normal retirement at the age of 65.

and Welfare. There is information on all deceased up until 31 December 2003, when those born in 1928 were 75 years old. By 31 December 2003, 515 of the original 1,542 persons had passed away. For 66 of these (those who died in 2002 or 2003), information on the cause of death is missing. Of the remaining 449, 176 persons died from cancer and 168 of circulatory diseases. No other cause of death is sufficiently common to deserve attention.[6] Figure 1 shows Kaplan–Meier survival function estimates for different groups in the sample. The top left panel shows estimates for the first (poorest) and fifth (richest) quintile groups in the distribution of parental income. The top right panel shows the corresponding estimates for the first and fifth quintile groups of household income in 1971. In the bottom left panel, the division is made on basis of lifetime earnings. Finally, the bottom right panel shows the survival function estimates for males and females, respectively.

Although the sample size for some of the groups is somewhat limited (around 250 individuals for the quintile groups), and the survival function estimates are consequently "jumpy,"[7] there are clear differences in the survival rates for the groups considered. The difference between those originating from the first and fifth quintiles of the distribution of parental income is about 7 percentage points at the age of 70. This difference is substantially lower than the corresponding difference between the quintile groups of household income in 1971 and the measure of lifetime earnings, which is about 13 percentage points. This difference is, in turn, similar to that between men and women around the same age. We can see that about 55% of the men and 75% of the women are alive by the end of the observation period.

When making these mortality comparisons, it should be noted that income inequality has decreased between the 1930s, when parental income was measured, and the early 1970s, when the individual's own family income is measured. Estimates of the Gini coefficient in the sample show a decrease from 0.368 in 1938 to 0.306 in 1971. This means that the relative income difference between the first and the fifth quintiles is larger for the distribution of parental income than for the individual's own household income in 1971. This, in turn, implies that, provided

6. The classification of cause of death follows the WHO standard of underlying cause of death. The data distinguish between 20 different causes of death corresponding to the international standard (ICD).

7. This is particularly true for the survival function estimates for the groups with high and low life-time income, shown in the bottom left panel. As can be seen in the figure, there are very few early deaths in the high income group and mortality is, on average, lower in this sample. This is primarily related to the fact that the mortality rate is higher among those who did not answer the questionnaire in 1963 and, as explained previously, were excluded from the sample used for the life-time income measure. The first death in the fifth quintile of the life-time earnings distribution was recorded in 1986. By that time, 59 in the entire sample used for the life-time earnings measure have deceased, compared to 119 in the sample used for the measure of household income in 1971. The background to the lower starting point for the survival function of the first quintile is that 19 of the individuals included in this group died before age 50.

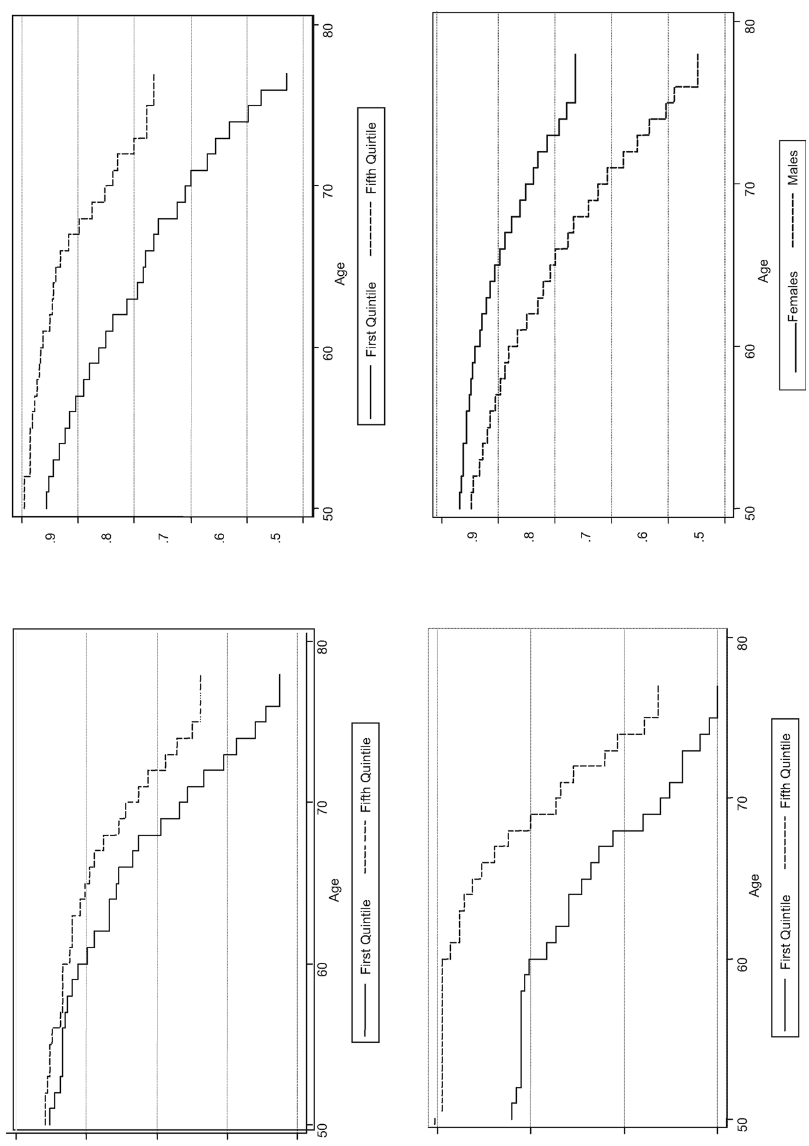

FIGURE 1. Kaplan–Meier non-parametric estimates of survival functions for different groups in the sample.

Note: Top left panel: first versus fifth quintile in the distribution of parental income. Top right panel: first versus fifth quintile in the distribution of household income in 1971. Bottom left panel: first versus fifth quintile in the distribution of lifetime earnings. Bottom right panel: men versus women.

that absolute income differences has any importance for mortality differences in both distributions, the effect of income differences in parental income is overestimated compared to the one from the differences in the distribution of household income in 1971.

3. Empirical Specification

In the empirical analysis, we use (discrete time) Cox regression models, namely,

$$\lambda_{1i}(t) = \lambda_0(t) \exp\left\{g\left(I_i^p\right) + \beta' \mathbf{x}_i\right\}, \tag{1}$$

where the dependent variable is the mortality hazard (the risk of the event), $\lambda_0(t)$ is the baseline hazard, I_i^p is our measure of parental income, $g(\cdot)$ is a general function and x_i is a vector of observable characteristics that may affect mortality. This Cox regression model is semi-parametric in the sense that it does not impose any functional form restriction for the baseline hazard. Because the hypothesis on a persistent effect of parental income does not impose any functional form on the relationship, we try specifications, both without and with logarithmic transformation.

As additional control variables for parental background, we use dummy variables for father's educational level and social class for the parental household. This means that the variation in parental income identifying differences in mortality later in life is within cells with the same father's educational level and social class. The motivation for including these variables is that mortality may be related to social rather than economic differences during childhood, and we want to isolate the effect of economic resources. However, we are still not able to control for variation in parental income that can be attributed to differences in the father's job tenure, age, and hours of work, which are not included in the data set.

For the two alternative measures of the individual's own income, we use the same strategy to determine the functional form relation, as described for parental income. We also include dummy variables for the individual's own educational level. The motivation for including the latter variable is to also control for differences in social resources which may be affected by the individual's education and which, in turn, may affect mortality. The indicators for educational level may also reflect unmeasured economic resources.

All specifications include a dummy variable for gender. The obvious reason is that the gender groups have very different mortality patterns. The analysis is also done separately for each gender group, which relaxes the assumption that parental income has a similar effect on gender groups. It also enables us to see whether the observed relation can primarily be attributed to one of the gender groups.

In addition to the separate analysis for gender groups, we split up the sample along three additional dimensions. First, we do separate analyses for individuals with low and high educational levels where low education is defined as only vocational or no post-compulsory schooling. High education is secondary education or beyond that. The idea behind this analysis is, once more, to see whether any observed relation can primarily be attributed to one of these groups, but also to examine whether there is a persistent effect of economic conditions during childhood for those with an upward social mobility, that is, those obtaining a higher educational level from low-income families. Second, we do a corresponding analysis, but divide the sample on the basis of father's education to account for potential endogeneity of educational attainment.

Finally, we estimate competing risk models for the two main causes of death in our sample: cancer and circulatory diseases. In these models, we maintain the discrete time Cox regression framework, treating all who passed away from other causes than the one under study as right censored at the date of death (e.g., Katz and Meyer 1990). This procedure gives consistent estimates for the separate causes of death under the assumption of independence between the latent causes of death, namely, if death from the other causes of death can be treated as random right censoring.

There are several medical reasons why the latent probability of death from cancer and circulatory diseases may be correlated (e.g., Honoré and Lleras-Muney 2006). Most of these come from lifestyle-related health conditions associated with alcohol consumption, smoking, and obesity. If there is perfect correlation between different latent causes of death, the competing risk model is not identified and the model for any cause of death will coincide with the original Cox regression model. In this sense, the results from the Cox regression model and the competing risk model described above form bounds for the estimates, where the former assume perfect correlation between the latent causes of death and the latter complete independence.

4. Results

Table 2 shows the first set of results. The first specification includes only a gender dummy and a linear variable in parental income to get the bivariate relation between parental income and mortality. In order to facilitate interpretations, we have, throughout, chosen to present estimated hazard ratios rather than coefficients. The hazard ratios measure the percentage change in the hazard corresponding to a one unit change in the underlying variable. A hazard ratio on 1 corresponds to "no effect" of the underlying variable. An estimate below one corresponds to an inverse effect. For example, the 0.51 estimate of the hazard ratio for the women indicator variable should be interpreted as women facing a mortality

TABLE 2. Estimates from Cox proportional hazard models. Men and women.

	(1)	(2)	(3)	(4)
Log of parental income	0.734 (−3.53)	862.9 (2.36)	1031.3 (2.37)	1680.5 (2.65)
Log of parental income^2	—	0.655 (−2.46)	0.646 (−2.48)	0.633 (−2.69)
Log of real lifetime earnings	—	164.0 (1.84)	—	—
Log of real lifetime earnings^2	—	0.836 (−1.92)	—	—
Log of early real lifetime earnings	—	—	2.730 (0.54)	—
Log of early real lifetime earnings^2	—	—	0.943 (−0.63)	—
Log of household income	—	—	—	2.132 (1.12)
Log of household income^2	—	—	—	0.946 (−1.49)
Female	0.510 (−6.90)	0.437 (−5.34)	0.420 (−4.05)	0.491 (−6.56)
Parental income, p-value for joint significance	0.000	0.023	0.019	0.024
Life-time income, p-value for joint significance	—	0.043	—	—
Early life-time income, p-value for joint significance	—	—	0.341	—
Household income, p-value for joint significance	—	—	—	0.000
Father's educational level, p-value for joint significance	—	0.089	0.101	0.287
Social class of parents, p-value for joint significance	—	0.645	0.238	0.399
Own educational level, p-value for joint significance	—	0.645	0.286	0.197
Log likelihood	−3,343.1	−2,160.4	−2,155.0	−2,698.2
Sample size	1,441	1,033	1,026	1,213

Notes: Father's as well as own education levels are measured by dummy variables corresponding to six different education levels, respectively. Parental social class is measured by dummy variables corresponding to four different levels. t-values in parentheses.

hazard (risk) that is, on average, 49% lower than that of the men in the sample. The corresponding coefficient estimate is -0.673, that is, $\exp\{-0.673\} = 0.51$. The t-ratios, shown in parentheses below each hazard ratio, indicate whether or not the underlying coefficient estimate is significantly different from zero.

The estimated hazard ratio and the t-ratio for the parental income variable shows a significant inverse relation between mortality and parental income. The hazard ratio for the log-linear variable measuring parental income shows that a log unit increase in parental income (an about 72% change) corresponds to a 26.6 percentage point decrease in mortality risk. This means that a 1% change in parental income corresponds to an about 0.37% reduction in mortality hazard.

For the other three specifications also shown in Table 2 we have added controls for the three different concepts of the individual's own income discussed in Section 2. Column 2 shows the results when we used the measure of lifetime income, column 3 when we used the measure of early lifetime income and, finally, column 4 when we used the measure of household income in 1971. Each

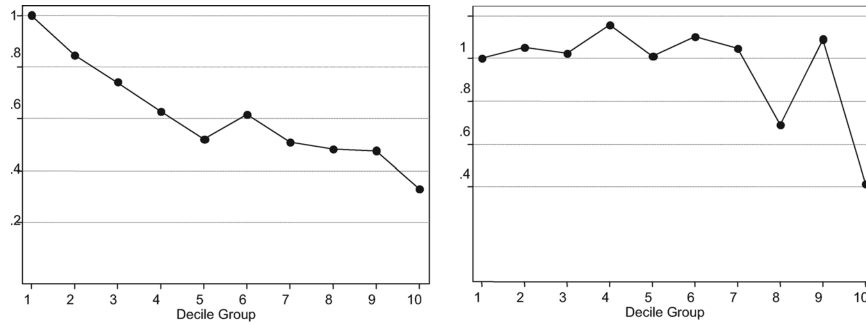

FIGURE 2. Separate hazard ratio estimates for different decile groups.

Note: Left panel: distribution of the individual life time income. Right panel: distribution of parental income. Specification as in column 2 in Table 2 except for the functional forms in household and parental income.

specification also include dummy variable's for the individual's own education in six different levels, dummy variables for father's educational level and social class. As we explained in Section 3, the interpretation of the coefficients for the parental income variables is now the effect of parental income conditional on the other included variables on mortality.

The functional relation between both parental income as well as the different measures of the individual's own income is of particular importance in this study. We have performed a parallel model selection process for both these variables. First, we added successive polynomial components in these income variables. The cubic component is in all cases insignificant and we maintained the quadratic polynomial specification in both the different measures of the individual's own income and parental income. Another strategy is to use dummy variables for decile groups in the two income distributions. The hazard ratios from the estimates of this model are shown in Figure 2 when we use household income in 1971 as a measure of the individual's own income. The left panel in Figure 2 shows the results for the individual's own income and the right panel the corresponding results for parental income. For both income measures, the first decile group is omitted in the model and the hazard ratios are consequently set to 1.

The results in Table 2 show that the measure of parental income is statistically significant in all specifications. The specification with dummy variables for each decile group shown in the right panel of Figure 2 shows that the only dummy variable being separately significant is the one for the 10th percentile in the income distribution. This gives us the extra information that the main effect on mortality comes from differences in the upper end of the income distribution.[8]

8. We have also tried a specification where we included the dummy variable for the 10th decile in the distribution of parental income in the quadratic specification. Because it turned out to be

Table 2 also shows that the alternative measures of the individual's own income are highly significant, except for the measure of early lifetime earnings. The lack of significance of the early lifetime income measure is likely to be due to the fact that the variation in the variable measuring early lifetime earnings is, as can be seen in Table 1, much smaller than for the alternative measures, because the higher earnings rates for people with higher education to a larger extent are counteracted by the earnings loss during the years of studies.

The estimates shown in Table 2 suggest a concave relation between mortality and individual's own economic status, although, as shown by the t-values, the underlying coefficient estimates are not separately significant. This is very different from the convex relation obtained in most previous studies (see, e.g., Deaton (2003) or Smith (1999) for overviews, or Gertham and Johannesson (2004) for empirical evidence). This issue could be further examined in the dummy variable specification shown in the left panel of Figure 2. These results provide no strong support for a concave relation. In fact, the only estimate in conflict with the usual convex relation is the one for the 10th decile group and this is not significantly different from the level of the hazard ratio for the 9th decile group. Taken together, we do not find the results sufficient for any definite conclusions on the functional relation.

In addition to the models presented in Table 2, we have performed two different types of sensitivity analyses.[9] First, as additional measures of the individual's own income, we use register data on annual earnings for the years 1963, 1971, 1982, and 1990, respectively. These models all correspond to slightly different samples, but the results are very robust, which means that they are not driven by mortality-attrition. Second, we have used a number of different estimation techniques for estimating the models in Table 2. These techniques include a probit model and hazard regression models with different parametric specifications for the base hazard. Overall, the results are remarkably similar. This indicates that it is the occurrence of death during the follow-up period, rather than the timing of death during that period, that is driving the results.

Although the results in Table 2 provide evidence of a robust statistical significance of parental income, they give very little guidance on the magnitude of the effect on mortality. To analyze the magnitudes of the effects, we use an estimated model including a quadratic polynomial of log parental income and log household income in 1971 along with a dummy variable for gender, to predict survival functions for different groups in the sample.[10] Figure 3, which contains

statistically insignificant, we concluded that the quadratic polynomial in parental income provided a parsimonious specification.

9. The results are available from the authors upon request.

10. We have also made predictions from the model presented in column 4 in Table 2, namely, when also including the insignificant dummy variables for father's education level and social class along with the dummy variables for own educational level. The results turned out to be similar to those presented in Figure 2.

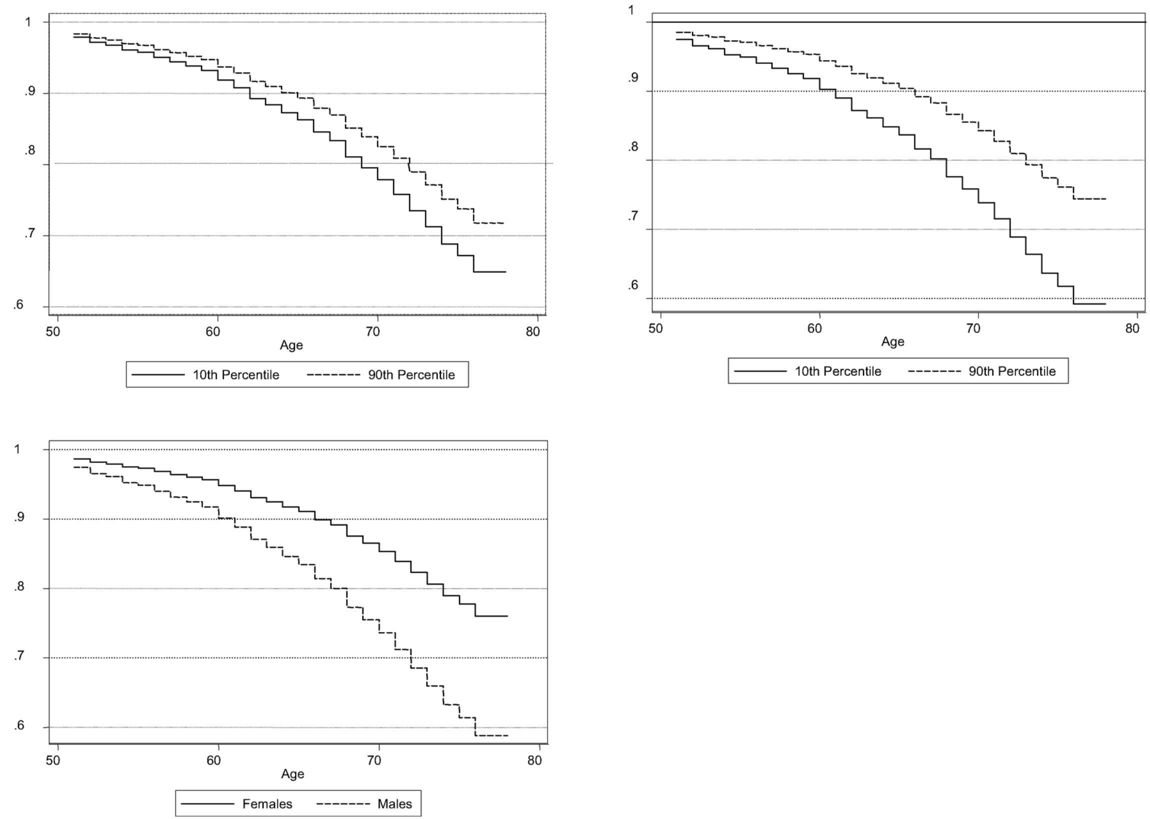

FIGURE 3. Predicted survival functions.

Note: Top left panel: 10th versus 90th percentile in the distribution of parental income. Top right panel: 10th versus 90th percentile in the distribution of household income in 1971. Bottom left panel: men versus women.

three panels, shows the results from this exercise. The top left panel shows the difference between the survival functions for the 10th and 90th percentiles in the distribution of parental income, respectively. The top right panel shows the same thing for the individual's own income measured as household income in 1971. Finally, the bottom left panel shows the survival functions for men and women, respectively, as a comparison.

The results are very similar to the nonparametric Kaplan–Meier estimates of the unconditional differences between the different groups reported in Figure 1: The differences in survival rates are much larger for the individual's own income than for parental income. At the age of 70, there is a 5 percentage-point difference between the 10th and 90th percentiles of the parental income distribution, compared to about a 10 percentage-point difference between the same percentiles of the distribution of household income in 1971. The latter difference is, once more, very similar to that between men and women.

Table 3 shows the results from estimates where the specifications from columns 2 and 4 in Table 2 have been separately estimated for individuals with

TABLE 3. Cox proportional hazard models for low- and high-education-level groups, separately. Men and women.

	Low Education		High Education	
	(1)	(2)	(3)	(4)
Log of parental income	42.929	60.353	94.969	1,746.8
	(0.75)	(0.36)	(0.81)	(1.37)
Log of parental income^2	0.794	0.781	0.734	0.622
	(−0.73)	(−0.88)	(−0.96)	(−1.52)
Log of real lifetime earnings	1205.76	-	95.974	-
	(1.86)		(0.90)	
Log of real lifetime earnings^2	0.779	-	0.852	-
	(−1.93)		(−0.96)	
Log household income	-	1.024	-	3.934
		(0.03)		(1..10)
Log household income^2	-	0.984	-	0.916
		(−0.36)		(−1.38)
Female	0.423	0.479	0.465	0.660
	(−4.70)	(−5.96)	(−2.62)	(−2.10)
Parental income, p-value for joint significance	0.717	0.514	0.034	0.025
Lifetime income, p-value for joint significance	0.045	-	0.238	-
Household income, p-value for joint significance	-	0.000	-	0.008
Father's educational level, p-value for joint significance	0.263	0.510	0.113	0.010
Social class of parents, p-value for joint significance	0.700	0.707	0.057	0.075
Log likelihood	−1,458.6	−2,015.3	−507.5	−674.2
Sample size	678	863	359	444

Notes: The sample for low education is restricted to those with own highest education corresponding to level 3, that is, no education or vocational schooling only. The sample for high education is restricted to those with an education higher than level 3, that is, junior secondary school, higher secondary school, or university education. Father's educational level is measured by dummy variables corresponding to six different education levels. Parental social class is measured by dummy variables corresponding to four different levels. t-values in parentheses.

TABLE 4. Cox proportional hazard models for low- and high-educated fathers, separately. Men and women.

	Low Education		High Education	
	(1)	(2)	(3)	(4)
Log of parental income	1.482	3.431	1.17e+07	3.94e+09
	(0.10)	(0.32)	(2.11)	(2.86)
Log of parental income^2	0.980	0.935	0.378	0.272
	(−.08)	(−0.28)	(−2.20)	(−2.91)
Log of real lifetime earnings	1556.48	-	1.612	-
	(2.20)		(0.09)	
Log of real lifetime earnings^2	0.772	-	0.981	-
	(−2.29)		(−0.11)	
Log household income	-	2.780	-	2.570
		(1.33)		(0.56)
Log household income^2	-	0.931	-	0.948
		(−1.71)		(−0.61)
Female	0.400	0.476	0.435	0.511
	(−5.28)	(−6.28)	(−2.37)	(−2.53)
Parental income, p-value for joint significance	0.957	0.727	0.027	0.011
Lifetime income, p-value for joint significance	0.009	-	0.941	-
Household income, p-value for joint significance	-	0.000	-	0.755
Own educational level, p-value for joint significance	0.821	0.191	0.755	0.892
Social class of parents, p-value for joint significance	0.677	0.627	0.503	0.309
Log likelihood	−1745.97	−2238.51	−307.83	−336.33
Sample size	848	1017	211	226

Notes: The sample for low education is restricted to those with own highest education corresponding to primary schooling. The sample for high education is restricted to those with vocational education or higher, that is, junior secondary school, higher secondary school, or university education. Parental social class is measured by dummy variables corresponding to four different levels. t-values in parentheses.

only vocational or no post-compulsory schooling and those with secondary education or more. The most interesting results revealed in Table 3 are that although own income is still significant, parental income completely loses its significance for low education individuals. However, for the highly educated, parental income is still significant. This result supports the hypothesis that parental income has a lasting effect on health status also for individuals obtaining a higher educational level than their parents. A possible explanation for the insignificant results for those with low education is the low variation in parental income. The standard deviation in parental income is substantially higher in the high education group: 0.73 as compared to 0.45.

We have also made the same analysis, but instead of the individuals' own education, we have split the sample according to the educational level of the fathers. As high-educated fathers we classified those that have more than compulsory education. As can be seen in Table 4, the results are similar to when the split was made according to own educational level. For those with low-educated fathers, parental income is insignificant.

Table 5 shows the estimates separately for men and women. In general, the results seem to be very similar for men and women. In the male subsample,

TABLE 5. Estimates from Cox proportional hazard models. Men and women separately.

	(1)	(2)	(3)	(4)
Males				
Log of parental income	0.739	68.554	36.577	55.260
	(−2.77)	(1.51)	(1.14)	(1.25)
Log of parental income^2	-	0.757	0.790	0.776
		(−1.64)	(−1.24)	(−1.30)
Log of real lifetime earnings	-	-	0.299	-
			(−0.20)	
Log of real lifetime earnings^2	-	-	1.023	-
			(0.12)	
Log of household income	-	-	-	1.600
				(0.56)
Log of household income^2	-	-	-	0.960
				(−0.88)
Parental income, p-value for joint significance	0.006	0.055	0.220	0.349
Lifetime/household income, p-value for joint significance	-	-	0.026	0.000
Father's educational level, p-value for joint significance	-	0.245	0.238	0.481
Social class of parents, p-value for joint significance	-	0.534	0.681	0.716
Own educational level, p-value for joint significance	-	-	0.887	0.329
Log likelihood	−2,042.8	−1,975.4	−1,288.8	1,666.5
Sample size	774	751	561	663
Females				
Log of parental income	0.726	4.79e+07	5,182,648	6,350,315
	(−2.19)	(3.11)	(2.53)	(2.71)
Log of parental income^2	-	0.340	0.389	0.387
		(−3.12)	(−2.56)	(−2.70)
Log of real lifetime earnings	-	-	1,356.2	-
			(1.27)	
Log of real lifetime earnings^2	-	-	0.779	-
			(−1.27)	
Log of household income	-	-	-	4.927
				(1.19)
Log of household income^2	-	-	-	0.906
				(−1.41)
Parental income, p-value for joint significance	0.029	0.008	0.035	0.025
Lifetime/household income, p-value for joint significance	-	-	0.445	0.021
Father's educational level, p-value for joint significance	-	0.598	0.203	0.322
Social class of parents, p-value for joint significance	-	0.101	0.195	0.224
Own educational level, p-value for joint significance	-	-	0.822	0.837
Log likelihood	−999.0	−964.8	−657.0	−891.8
Sample size	667	646	472	600

Note: As in Table 2.

TABLE 6. Esitimates from competing risk models separating cancer and circulatory diseases as causes of death. Men and women.

	Cancer			Circulatory Diseases		
	(1)	(2)	(3)	(4)	(5)	(6)
Log of parental income	0.702	8.87e+08	1.01e+08	0.702	3.412	17.076
	(−2.38)	(2.85)	(2.90)	(−2.29)	(0.34)	(0.70)
Log of parental income^2	-	0.291	0.330	-	0.904	0.831
		(−2.82)	(−2.86)		(−0.47)	(−0.75)
Log of real lifetime earnings	-	12.623	-	-	243.90	-
		(0.61)			(1.19)	
Log of real lifetime earnings^2	-	0.916	-	-	0.819	-
		(−0.63)			(−1.29)	
Log of household income	-	-	0.776	-	-	1.273
			(−0.22)			(0.21)
Log of household income^2	-	-	1.001	-	-	0.973
			(0.02)			(−0.44)
Female	0.672	0.657	0.716	0.350	0.277	0.326
	(−2.50)	(−1.64)	(−1.90)	(−5.67)	(−4.69)	(−5.58)
Parental income, p-value for joint significance	0.017	0.017	0.012	0.022	0.319	0.613
Lifetime/household income, p-value for joint significance	-	0.787	0.049	-	0.065	0.015
Father's educational, p-value for joint significance	-	0.018	0.127	-	0.738	0.984
Parental social class, p-value for joint significance	-	0.057	0.041	-	0.511	0.716
Own educational level, p-value for joint significance	-	0.632	0.633	-	0.885	0.471
Log likelihood	−1,171.9	−757.7	−940.2	−1,087.3	−742.1	−950.2
Sample size	1,441	1,033	1,213	1,441	1,033	1,213

Note: As in Table 2.

parental income is highly significant when including father's education and social class, but excluding own income and educational level in the specification. However, the significance is lost when also including own income and educational level in the specification. Parental income has a more robust impact on the female subsample and it is also highly significant when including controls for own income and final educational level.

Table 6 shows the results from the competing risk analysis, where we distinguish between the two main causes of death in the sample: cancer and circulatory diseases. For cancer diseases, parental income is highly significant in all specifications. Parental social class is at least marginally significant and the individual's own income is significant when measured as household income in 1971. For circulatory diseases, however, only the individual's own income is significant and, once more, when measured as household income in 1971 a more precise estimate is obtained.

The result that parental income is important for cancer as a cause of death is reasonable, as it is well known that cancer, at least some forms, takes a very long time to develop. There are several candidates for social, environmental, and nutritional factors during childhood that can be related to cancer diseases later in life. There are, for example, several pieces of evidence on how exposure to helicobacter pylori bacteria is related to overcrowded living conditions during childhood and gastric cancer later in life (see, e.g., Ueda et al. (2003) or Whitaker, Dubiel, and Galpin (1993)). However, our data are not detailed enough to separate out different potential environmental factors.

The insignificant result for circulatory diseases is harder to interpret. It should be stressed that it could be an outcome from a correlation between this cause of death and the latent probability of death from cancer diseases, namely, a violation of the assumption underlying the competing risk model. Such correlation would imply that those with an increased probability of dying from circulatory diseases as a result of poor living conditions during the childhood, also have an increased probability of dying early from cancer. The only conclusion we can make is that our result provides no support for the claim that parental income has an effect on death from circulatory diseases later in life.

5. Conclusions

We find parental income during childhood to be associated with mortality risk later in life. This relation is found in the non-parametric Kaplan–Meier analysis as well as in the Cox regressions when we include controls for parental socioeconomic conditions measured by father's education and social class, own lifetime income, or household income in 1971 (at the age of 43 for most individuals in the sample) as well as the own educational level. The strongest effect is located in the upper end of the distribution of parental income. When dividing the sample into different sub-groups, we find the relation between parental income and mortality to be significant among those with a high educational level and among women. When separately analyzing cancer and circulatory diseases as causes of death in a competing risk model, we find the relation under study to apply to cancer as the cause of death.

What do our results tell us about the two research questions posed in the Introduction? For the first question, whether there is an unconditional relation between parental income and mortality, we find unambiguous evidence for such a relation, also when controlling for father's education and parental social class. This result is very unlikely to be affected by any "reversed causality" problem, namely, that the child's health status would affect the parents' earnings possibilities, because all children in our sample are enrolled in normal schools.

The second question, whether there is a direct (conditional) effect of parental economic resources on the mortality of their children, irrespective of the children's economic status later in life, is naturally a much more demanding empirical issue.

Because our data include several outcomes over the included individuals' entire life cycle, we are able to control for a comparatively rich set of information on educational attainments and economic conditions later in life. However, the data are not sufficiently rich to allow us to measure the relative importance of the effect of inferior chances of acquiring education and subsequent better economic resources for children from poorer backgrounds on mortality, vis-à-vis the persistent effect of parental income on mortality.

The fact that the relation between parental income and mortality applies also when we restrict the sample to the group with high education strengthens the interpretation of a separate effect of parental income on mortality later in life. This shows that the relation between parental income and mortality can be referred to the group of individuals whose health status did not restrict them to obtain at least a secondary education. It is also likely that most, if not all, in this group have lived most of their life under relatively good economic conditions. Still, there is a persistent effect of parental income on mortality within this group.

The result that parental income is related to mortality in cancer diseases gives a potential medical explanation for how health is formed by economic resources on a long-term basis. Cancer is well known to have a long incubation period and previous medical research has shown that there may be a link between the living conditions during childhood and the risk of getting cancer as well as genetic origins. This link is, however, different from the mechanism suggested by the influential studies by David Barker on the relation between living conditions very early on and the prevalence of circulatory diseases later in life.

A third potential research question is whether results obtained in this study could be given a causal interpretation. From our data, it is not possible to separate out any effect of genetic differences, or socially inherited life habits, that may be correlated with parental income. Results obtained by Almond, Chay, and Lee (2005) suggest a comparatively large bias from genetic factors on cross-section estimates of the returns to low birth weight prevention. The linkage between mortality from cancer diseases and parental income does not rule out the possibility of bias from genetic factors, because some cancer diseases, in addition to environmental influences, are well known to be related to genetic backgrounds. This provides an alternative to a causal interpretation of our results.

References

Almond, Douglas, Kenneth Chay, and David Lee (2005). "The Cost of Low Birth Weight." *Quarterly Journal of Economics*, 120, 1031–1084.

Barker, David J. P. (1997). "Maternal Nutrition, Fetal Nutrition, and Disease in Later Life." *Nutrition*, 13, 807–813.

Böhlmark, Anders, and Matthew Lindquist (2006). "Life-Cycle Variations in the Association between Current and Lifetime Income: Country, Cohort and Gender Comparisons." *Journal of Labor Economics*, 24, 879–896.

Case, Anne, Angela Fertig, and Christina Paxson (2005). "The Lasting Impact of Childhood Health and Circumstance." *Journal of Health Economics*, 24, 365–389.

Case, Anne, Darren Lubotsky, and Christina Paxson (2002). "Economic Status and Health in Childhood: The Origins of the Gradient." *American Economic Review*, 92, 1308–1334.

Deaton, Angus (2003). "Health, Inequality and Economic Development." *Journal of Economic Literature*, 41, 113–158.

Forsdahl, Anders (2002). "Observations Throwing Light on the High Mortality in the County of Finnmark. Is the High Mortality Today a Late Effect of Very Poor Living Conditions in Childhood and Adolescence?" *International Journal of Epidemiology*, 31, 302–308.

Gertham, Ulf, and Magnus Johannesson (2004). "Absolute Income, Relative Income, Income Inequality and Mortality." *Journal of Human Resources*, 39, 228–248.

Haider, Steven, and Gary Solon (2006). "Life-Cycle Variation in the Association between Current and Lifetime Earnings." *American Economic Review*, 96, 1308–1320.

Hallgren, Siver (1939). "Intelligens och Miljö: En experimentell undersökning av barn i tredje skolåret vid Malmö folkskolor och privata skolor." Unpublished Licentiate thesis. University of Lund.

Honoré, Bo, and Adriana Lleras-Muney (2006). "Bounds in Competing Risk Models and the War on Cancer." *Econometrica*, 74, 1675–1698.

Katz, Lawrence, and Bruce Meyer (1990). "Unemployment Insurance, Recall Expectations, and Unemployment Outcomes." *Quarterly Journal of Economics*, 105, 973–1002.

Marmot, Michael, Martin Shipley, Edgar Brunner, and Harry Hemingway (2001). "Relative Contributions of Early Life and Adult Socioeconomic Factors to Adult Morbidity in the Whitehall II Study." *Journal of Epidemiology and Community Health*, 55, 301–307.

Mazumder, Bhashkar (2005). "Fortunate Sons: New Estimates of Intergenerational Mobility in the United States using Social Securituy Earnings Data." *Review of Economics and Statistics*, 87, 235–255.

Sandgren, Sofia (2007). "Comparing Annual and Lifetime Earnings: Distributions and Wage Premiums for a cohort of Swedish Men." *Journal of Income Distribution*, 16, 26–49.

Smith, George Davey, Carole Hart, David Blane, Charles Gillis, and Victor Hawthorne (1997). "Lifetime socioeconomic position and mortality: prospective observational study." *British Medical Journal*, 314, 547–552.

Smith, James (1999). "Healthy Bodies and Thick Wallets: The Dual Relationship between Health and Economic Status." *Journal of Economic Perspectives*, 13, 145–166.

Ueda, Mitsue, Shogo Kikuchi, Tatsuzo Kasugai, Tuuki Shunichi, and Chieko Miyakeu (2003). "Helocobacter Pylori Risk Associated with Childhood Home Environment." *Cancer Science*, 94, 914–918.

van den Berg, Gerard, Maarten Lindeboom, and France Portrait (2006). "Economic Conditions Early in Life and Individual Mortality." *American Economic Review*, 96, 290–302.

Whitaker, C. J., A. J. Dubiel and O. P. Galpin (1993). "Social and Geographical Risk-factors in Helocobacter-pylori Infections." *Epidemiology and Infections*, 111, 63–70.