Why a positive link between age and income-related health inequality?

NORDIN, MARTIN 1,2,*
GERDTHAM, ULF-G. 1,2,3

1 Department of Economics, Lund University
2 Health Economics & Management, Institute of Economic Research, Lund University
3 Centre for Primary Health Care Research, Lund University

Abstract: This study uses Statistics Sweden’s Survey of Living Conditions (the ULF) to ascertain why the SES-health gradient increases with age. By comparing various SES and health outcome relationships it explores the age increase in health inequality and distinguishes between three types of explanations: i) the effect of SES on individual health is causally increasing with age ii) reversed health effect on SES; and iii) lifecycle variation in the measurement errors in SES. Results indicate that the age increase in health inequality is primarily caused by a reverse causality going from health to income, and the probable mechanism is health affecting the labour supply of the individual. The evidence in this study is not conclusive in that it does not contain a direct proof. However, documented evidence supports this conclusion.

JEL classification: D30, D31, I10, I12

Key words: health inequality, socioeconomic status, income, education, age

1 Introduction

Numerous studies report the existence of a persistent socioeconomic status (SES)-health gradient in every country, regardless of SES and most health outcome measures (see for example; Baum and Ruhm, 2009; Buckley et al., 2004; Deaton and Paxton, 1998; Gerdtham and Johannesson 2000, 2002, 2004; Smith, 2004; van Doorslaer et al., 1997; Wagstaff and van Doorslaer, 2000; van Doorslaer and Koolman, 2004). However, despite the undisputed existence of socioeconomic inequality in health, a long-standing debate on the causal impact of different factors and widespread health policy concerns about the health inequality in a number of governments, there is little conclusive evidence of the underlying mechanisms behind the observed health inequality (e.g. Smith, 1999; Deaton, 2002; Cutler et al., 2008).

Several studies have established that the unequal distribution of health among SES groups increases with age (Baum and Ruhm, 2009; Case and Deaton, 2005; Deaton and Paxton, 1998; Islam et al., 2010), and Cutler et al. (2008) stress that one has to take the lifecycle into account when modelling causality of the SES-health gradient. For example, Case and Deaton (2005) demonstrate that the SES-health gradient for the US increases up to the age of 55 and narrows after that. But, again, less is known about why the SES-health
gradient increases with ageing. This paper aims to expand knowledge in this area by exploring different types of explanations for a positive relationship between age and health inequality.

The literature mentions a number of different explanations for such a positive relationship. The first explanation is the most straightforward: the effect of SES on individual health may causally increase with ageing, e.g. having a low education level and/or a low income makes one more prone to a risky lifestyle that gradually affects one’s health. This explanation is also known as the cumulative-advantage hypothesis (Lynch, 2003).

The second explanation emphasizes that the relationship could well run in the opposite direction, i.e., from health to SES. As Deaton and Paxton (1998) suggest, if labour supply and income are adversely and cumulatively affected by health shocks, ageing may increase the relationship between health and income. Recent studies support the argument that the SES-health gradient primarily runs from health to income (Banks et al., 2007; Case and Deaton, 2005; van Kippersluis et al., 2009a; Smith, 2005), and operates through employment.1

A third explanation could be that the positive age increase in the SES-health gradient is related to measurement errors in the SES, and with lifecycle variation in the measurement errors in SES, the gradient might vary with age.2 Few studies have explored the influence of measurement errors as a candidate for explaining the lifecycle variation in health inequality.

This study uses Swedish data to explore and disentangle the three types of explanations for a positive relationship between age and health inequality: i) age increase in the causal SES effect; ii) reversed health effect on SES; and iii) lifecycle variation in the measurement errors in SES. The main contribution is that we compare the age variation in the income effect on health with the age variation in the education effect; to our knowledge, this has only been assessed by van Kippersluis et al. (2010). Our specific contribution is that we analyse whether measurement errors in income explain the relationship between age and health inequality, which has not been assessed before. Thus, compared to other studies on the age increase in health inequality, our study contributes to improved understanding by being the first to consider several explanations. We should point out that our primary aim is not to document whether the effects of SES on health are causal, but rather to identify what determines the age increase in health inequality. However, in determining the part of health inequality that increases with age, the remaining part (the part that does not vary with age) is largely reduced.

First, since health shocks do not affect a fixed socioeconomic measure such as final education level, the relationship between final education level and health should not increase by ageing due to reverse causality between the variables. Even if it is commonly believed that health inequality exists irrespective of the SES measure, few studies of the lifecycle perspective in health inequality focus on education as the SES measure. An exception is van Kippersluis et al. (2009a) who find, when using educational attainment as their SES measure, an age increase in health inequality for men (but not for women) in the Netherlands. Moreover, the measurement error may be much smaller for final education level than for current income. Therefore, for different age groups we estimate and

---

1 Intergenerational transmissions could also partly explain the reversed causation. Since adult health seems to be determined by chronic conditions in childhood and child health is closely related to parental income (Case et al., 2002), the health effect on income may be related to parental SES.

2 For a further discussion regarding the measurement error in income, see for example Grawe (2006) or Böhlmark and Lindquist (2006).
compare the income and education effects on health. If the age increase in SES-health inequality is caused by reverse causality and/or measurement errors, we expect the age increase to be much smaller for final education level than for current income. Second, by also using permanent income (average of at least three different yearly incomes) we further examine whether the age increase in socioeconomic health inequality might be related to measurement errors in the independent variable.

We explore the SES gradient in health by using three different health indicators: health, overweight and mortality. This way, because the mechanism underlying the different health and SES relationships might vary, we expect to learn more about the age increase in health inequality. For example, financial constraints determining whether one can afford a healthy diet may be a relatively strong predictor of one’s risk of overweight. Besides, since dying is definite, mortality cannot directly be a predictor of one’s socioeconomic position. Still, mortality is also a measure of past health, which is of course a potential predictor of one’s socioeconomic position.

Since health seems to affect income through employment (Banks et al., 2007; Case and Deaton, 2005; Smith, 2005; van Kippersluis et al., 2009a), we further explore the issue of reverse causality by estimating the respective relationships for only the working population (i.e. we exclude the unemployed, the retired and the early-retired). For example, if a health shock affects income through early retirement, this type of reverse association between health and income is ruled out.

Statistics Sweden’s Survey of Living Conditions (the ULF), which contains a random sample of adults interviewed between 1980 and 2005, enables us to use longitudinal data to control for both age and time fixed effects (along with other covariates). Furthermore, by restricting the analysis to ages 20 to 64 we limit the potential channels through which SES could affect the health of the individual. That is, selective mortality and retirement, which might decrease the age variation in the SES-health gradient (Case and Deaton, 2005; van Kippersluis et al., 2009b), have minor effects on health inequality among those of working age.

2 Data

Every year, Statistics Sweden conducts systematic surveys of living conditions, in the form of one-hour personal interviews with randomly selected adults aged 16-84. About 6,000 individuals have been interviewed each year since 1975.

The questions are structured into four main themes: Social relations, work, health and physical environment. A large number of basic questions on each of these themes are asked every year, and particular emphasis is put on each of the themes on a rotating schedule. For example, the years 1980/81, 1988/89, 1996/97 and 2004/05 represent the specific “health–years”. Thus, overweight is only collected for the health-years, whereas self-reported health is measured yearly.

Although the data is primarily cross-sectional, we also observe a panel sample: 11 percent of the individuals are included twice, 10 percent three times and 3 percent four times. As a result, the total sample consists of approximately 120,000 individuals, but the data consists of 167,166 observations. The individuals in the panel are followed up for 7-24 years, depending on when they were first interviewed. We use cross-sectional data from the annual interviews in some of our analyses, but only the health-years or the panel sample in others.

The survey data is merged with register data and register information on survival status at the end of a follow-up. Thus, the dataset includes information on risk factors such as weight, physical activity, and socio-economic characteristics including place of
residence, education, occupation, employment status and actual income. The mortality experience of the total sample was recorded until December 31, 2005.

Our study restricts the sample to the working ages, i.e. 20 to 64, which decreases it to 87,701 individuals or 120,002 observations. For students, the measurement error in final education level and current income (as a measure of lifetime income) is large. Therefore, excluding 11,392 students (and 616 observations doing military service), the sample contains 80,206 individuals or 108,610 observations. Finally, we exclude those with an annual income of zero, which leaves us with 73,196 individuals or 98,469 observations (361 observations are also lost because of missing values for the education variable or the self-reported health variable). Because individuals with an annual income of zero are often financially supported by a spouse or some other relative, an annual income of zero probably does not represent their true financial means. Information regarding overweight is only collected in the panel sample; we lose another 50 to 60 percent of the sample for this outcome.

Buckley et al. (2004) argue that a wealth variable is preferred to an annual income variable, because wealth captures the continuous influence of one’s financial resources better. It is true that a wealth variable partly measures something other than annual income does, but because the financial argument is only one explanation (and not necessarily the most important) for health inequality, we prefer the (in many respects more standard) annual income measure. Moreover, with the aim of estimating the income effect for different periods of life, we use a measure that fluctuates, i.e. we want to keep the lifecycle variation. To compare the income-health relationship with the education-health relationship, we use individual measures in both cases. However, the results for income do not differ much when using household income instead of individual income.

Statistics Sweden provides us with an annual income measure that contains income from employment and business, as well as transfers (pension payments, unemployment benefits, paid sick-leave, housing assistance etc.). However, to find out whether the age increase in health inequality is caused by measurement errors, we also use a measure of permanent income; that is, we compute an average of (at least) three yearly observations of the individual’s annual income. This income measure is therefore only calculated for the sample belonging to the panel.

Our educational attainment variable is constructed according to SUN (Swedish Educational Terminology), the standard system for classifying education in Sweden, which contains the following years of schooling; eight, nine, eleven, twelve, fourteen, sixteen and eighteen. Since we exclude students, we assume (at least for those above the age of 25) that the education measure is final education level.

Our health variables are self-reported health, overweight and mortality. Overweight is measured as having a BMI over 25, according to the standard BMI methodology, i.e. weight(kg)/height(m)². We have also looked at BMI and obesity (BMI over 30). Whereas BMI gives identical results to overweight, the small fraction of obese individuals (especially in the younger age groups) hinders the analysis.

Because the interviews in the panel are (in general) performed every eighth year, the mortality indicator is constructed so that an individual is classified as deceased if he/she dies during the following eight-year period. By constructing such a mortality variable, we exploit more of the variation in data compared to using an indicator variable measuring only whether the individual dies during the entire time period. Nonetheless, the analysis focuses on the probability of surviving the next period of eight years.

Table 1 contains summary statistics for the main variables used in the study. To find out whether there are lifecycle variations, the statistics are reported separately for five-year age groups. Table 1 shows, as expected, that self-reported health decreases over
the lifecycle, and that overweight and the probability of dying increase with age. Furthermore, the income of both men and women increases up to the age of 49, and then it gradually falls. Table 1 also shows that younger age groups are more educated. The youngest cohort, however, has relatively low education, probably because these young individuals have yet to reach their final education level. Finally, the probability of early retirement or retirement does of course increase considerably with age, whereas the level of unemployment is much higher for young cohorts (and somewhat higher for the oldest cohort).

Table 1. Descriptive statistics. Observations, means and standard deviations for the age groups.

<table>
<thead>
<tr>
<th>Age group:</th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>20-24</td>
<td>25-29</td>
</tr>
<tr>
<td>N</td>
<td>4,593</td>
<td>5,702</td>
</tr>
<tr>
<td>%</td>
<td>9.2%</td>
<td>11.4%</td>
</tr>
<tr>
<td>Self-reported health (mean)</td>
<td>2.895</td>
<td>2.894</td>
</tr>
<tr>
<td>Overweight (mean)</td>
<td>0.005</td>
<td>0.005</td>
</tr>
<tr>
<td>Mortality (mean)</td>
<td>0.022</td>
<td>0.331</td>
</tr>
<tr>
<td>Ln Income (mean)</td>
<td>7.212</td>
<td>7.525</td>
</tr>
<tr>
<td>Retired (mean)</td>
<td>0.001</td>
<td>0.003</td>
</tr>
<tr>
<td>Unemployed (mean)</td>
<td>0.143</td>
<td>0.078</td>
</tr>
<tr>
<td></td>
<td>0.005</td>
<td>0.004</td>
</tr>
</tbody>
</table>

Notes: Standard deviations are shown below the means. The sample sizes are smaller for Overweight.
3 Empirical specification

In order to analyse the age increase in health inequality, we estimate a health equation based on pooled (i.e. we treat each individual-year observation as a single observation) cross sectional data where we model separate SES effects for each age group. That is, we estimate the following equation:

\[ Health = \alpha_n \text{Age}_n + \lambda_j \text{Year}_j + \sum_{m=1}^{9} \delta_m \beta_m \text{SES} + \gamma X + \varepsilon \]  \hspace{1cm} (1)

for which we create \( M \) dummy variables \( \delta_m \), where \( m \in [1,...,9] \), one for each age group, and interact them with the SES measure. This means that we use a separate SES variable for each age group, and \( \beta_m \) gives us the SES effect for each of the nine age groups. We allow for age and time differences in health by including age and time fixed effects. With \( X \) we control for additional covariates (civil status, children, immigrant status, region and parental background). Clustered standard deviations (on the individual level) are calculated.

According to van Kippersluis et al. (2009b), cohort fixed effects might seem more important than time fixed effects. For example, health conditions common during an individual’s childhood might generate cohort differences that could confound with the age variation in health inequality (van Kippersluis et al., 2009b). However, using cohort fixed effects instead of time fixed effects does not change any of our results. Using the panel sample in the data, we could estimate a panel-data individual fixed-effects model. However, this is not an appropriate specification with which to analyze the age increase in health inequality, because such a model would eliminate the effects of interest in this study, i.e. the between-individual variation in income at a certain age would be removed.

The specification above is separately estimated for men and women and for each health-SES combination, i.e. where self-reported health, overweight or mortality are regressed on either log annual income (and permanent annual income) or years of schooling. Using standardized (subtracting the mean and dividing by the standard error) health variables, we can easily compare the results for different relationships, i.e. the income or education effects for different health outcomes are then comparable. Also, when using standardized variables, ordinal computations of the health outcomes do not affect our results.

4 Results

4.1 The relationship between health and annual income

Figure 1, for men, and Figure 2, for women, illustrate the age variation in the (log) income effect on different health outcomes. Together with Figure 3 and Figure 4, these figures report the main results of our study.

Figures 1 and 2 generally show that the expected age increase in health inequality is present (the income effect grows with age), and for self-reported health the variation is particularly large. For self-reported health many of the income effects for different age groups are significantly different from each other. The age increase seems to be somewhat larger for women than for men, but there are clearly large differences in the relationships,

3 For the binary outcomes we estimate a linear probability model.
4 Along with the results in other studies (see for example Buckley et al., 2004) the income effect on health is relatively unaffected by the inclusion of covariates (see for example Buckley et al., 2004). Tables of these results are available upon request.
Why a positive link between age and income-related health inequality? 193

Fig. 1-2 Age variation in the relationship between income and health.

Notes: The size of the income effect is on the y-axis and the x-axis describes the age groups. The income effects for the age groups are estimated in a pooled OLS model (see equation 1).

and it is quite obvious that there is a stronger association between income and self-reported health than between income and overweight and mortality, respectively.

Before comparing the health outcomes, we focus further on self-reported health. Besides decreasing for ages 60 to 64 (relative to ages 45 to 59), the age increase in the income effects is substantial and fairly linear for self-reported health. The income effect is about four to five times larger for ages 50 to 55 compared to ages 20 to 29. The decrease in the income effect for the oldest cohort is in accordance with other results (see for
example Case and Deaton, 2005), and the decrease is most likely linked to retirement behaviors affecting current income.\(^5\)

When focusing on overweight and mortality, the income effects show a different pattern compared to self-reported health. Although few of the income effects are significant (or significantly different from each other), the main picture for overweight is that there is, in particular, a difference in the income effects for lower age groups (20 to 34 for men and 20 to 39 for women) compared to higher age groups (35 to 64 for men and 40 to 64 for women). For women, there is a tendency for a linear age increase in the income effect. For the lower age groups, some income effects on overweight are negative, which indicates that there is a positive relationship (since we have changed sign) between income and overweight for younger cohorts. In the case of mortality we find, particularly for men, a significant income effect for older age groups, but overall the age variation in health inequality for mortality is very small.

A plausible mechanism for an age increase in health inequality (and health inequality in general) is that income may explain food expenditures; i.e., to afford a healthy diet one needs a high income. But if this is the driving mechanism for health inequality, we may expect the age increase in the income effect on overweight to be as large as (or larger than) the income effect on self-reported health. Since this is not the case, the age increase in health inequality must primarily have some other explanation. We pursue the task of finding this explanation by studying the education effects on health.

### 4.2 The relationship between health and education

Figure 3 and Figure 4 illustrate the lifecycle variation in the education effect on health. The figures report no clear age increase in the relationship between years of schooling and health, and there are apparent differences in comparison to Figure 1 and Figure 2 (the income effects on health). However, the education effects on self-reported health and overweight are significant (not reported). Hence, while education seems to determine the health of the individual, the effect does not increase with age. This finding indicates that the age increase in the income-related health inequality may be related to reverse causality, which is an important and rather surprising result that needs to be further analyzed.

### 4.3 Measurement errors in current income

To try to explore whether lifecycle variation in the measurement error in current income is causing our (with age) increasing income effect, we use a measure of permanent income (with less measurement error), instead of current income, as our income measure. If the income effect when using permanent income, increases for the young cohorts (for which the measurement error might be substantial) relative to older cohorts, it indicates that the age variation in health inequality is related to measurement errors in current income. Figure 5 (for men) and Figure 6 (for women) report that the lifecycle variation in income effects on self-reported health strengthens rather than diminishes when permanent income is used instead of current income. The lifecycle variation in the income effects for overweight and mortality does weaken though. We therefore conclude that measurement errors in the income variable do not seem to cause the age variation in health inequality.

\(^5\) For a subsample we are able to separate the retired and the early-retired, and by excluding the retired we find that the income effects are somewhat higher for the oldest age group. Moreover, since the results are unchanged when excluding those who die during the next eight years, selective mortality does not explain the low income effect for the oldest age group.
Why a positive link between age and income-related health inequality?

Why a positive link between age and income-related health inequality?  195

Fig. 3-4 Age variation in the relationship between education and health.

Notes: The size of the education effect is on the y-axis and the x-axis describes the age groups. The education effects for the age groups are estimated in a pooled OLS model (see equation 1). For comparability we change the sign of the education effects on overweight and mortality; otherwise these relationships go in the opposite direction compared to self-reported health.

4.4 Excluding the unemployed, retired and early-retired

The next step of the analysis is to estimate the relationships for only the working population. In the case of lasting reduced working capacity, the Swedish disability insurance program replaces foregone earnings of early-retired workers. Before 1997 unemployed individuals could be granted disability benefits for labor-market reasons and health reasons combined. The disability compensation rate today is 64 percent of the income, up to a ceiling, and it has remained fairly constant over time. Thus, when excluding the unemployed, the early-retired and the retired (after the age of 60 the individual can decide to retire), the extent to which current income is affected by health
Fig. 5-6 Age variation in the relationship between permanent income and health.  
Notes: The size of the permanent income effect is on the y-axis and the x-axis describes the age groups. The permanent income effects for the age groups are estimated in a pooled OLS model (see equation 1). The measure of permanent income is an average of (at least) three yearly observations of the individual’s annual income. For comparability we change the sign of the income effects on overweight and mortality; otherwise these relationships go in the opposite direction compared to self-reported health.

shocks is constrained. This analysis will be performed only for the income-self-reported health relationship, because final education is not affected by labor-market status.

In Figure 7 for men and Figure 8 for women, the age variation in the relationship between self-reported health and income is illustrated when excluding the unemployed, the retired and the early-retired. For comparability, the figures also contain the relationships for the entire population (the relationships illustrated in Figure 1 and 2).
Why a positive link between age and income-related health inequality?

**Fig. 7-8** Age variation in the relationship between self-reported health and income when excluding the unemployed, retired and early-retired.

*Notes:* The size of the income effects is on the y-axis and the x-axis describes the age groups. The income effects for the age groups are estimated in a pooled OLS model (see equation 1).

When restricting the sample to the working population, much (70 percent) of the increase in the health inequality between age 20 and 54 vanishes. The figures do not show it, but the decrease in age variation is almost entirely caused by excluding the retired and the early-retired. Together, these findings are additional evidence of reversed causality being the primary explanation for the age increase in health inequality. Because health shocks resulting in sickness absence or part-time work are not excluded, a minor part of the age...

---

6The sick leave replacement rate in Sweden is about 80 percent (and has varied over time), and depending on the duration of the sick leave, the benefit is paid either by the employer (sick-pay) or the National Sickness Insurance (sickness benefit).
increase in the health gradients remains. On the other hand, employers generally cannot fire sick employees or lower their wages.

5 Discussion

That the age increase in health inequality does not exist for final education level indicates that there is something different going on in the relationship between education and health relative to the relationship between income and health. Moreover, since the age variation does not seem to operate one’s weight, the possible SES-related channels through which SES might influence health are further limited. Furthermore, we do not find any major age increase of health inequality in mortality (as already mentioned dying is definite and cannot affect one’s SES). Thus, if we assume that reverse causality does not (or only marginally) affect the education-health relationship and that the measurement error in final education level is minor, we can draw the following conclusion: the age increase in health inequality is caused by reverse causality and/or lifecycle variation in the measurement error in current income. The conclusion might seem to hinge on an assumption that the potential causal mechanisms behind the age increase in the income effect and the education effect are the same. However, this is only a necessary assumption when the correlation between education and income is zero. In the more empirically realistic case, where education and income are highly correlated, a specific causal mechanism for one of the SES indicators will show up for the other SES indicator as well, i.e. if there is a causal mechanism behind the age increase in the income effect, it will to some extent also show up for education. Therefore, when there is no age increase in the educational effect, the causal mechanisms do not even have to be identical.

The results, when using a measure of permanent income instead of current income, show that measurement error in the income variable is not responsible for the age increase in health inequality. The final evidence indicating that it is health affecting income (and not the other way around) is that, when restricting the sample to the working population, the age variation in health inequality decreases substantially, i.e. when excluding the sample where health definitively affects the income of the individual, the age increase in health inequality is very small.

A non-random measurement error in the health measures may exist, for example if high SES groups overestimate their health. But for this to affect the results of this paper, the non-random measurement error has to vary with age as well, otherwise the measurement error will only affect the age-constant part of the relationship between SES and health. We find such an effect unlikely, particularly as it will then only affect the age increase in the income effect on self-reported health, and not the age increase in the education effect. Moreover, van Doorslaer and Gerdtham (2003) show that there is no systematic adjustment of self-reported health reporting by SES in the ULF.

6 Concluding remarks

Our analysis supports the notion that the age increase in health inequality is primarily caused by a reverse causality going from health to annual income, and the probable mechanism is health affecting the labour supply of the individual. The evidence in our study is not conclusive in that we do not directly prove this, but all the documented evidence agrees with and supports this conclusion.

7 Excluding the unemployed, the retired and the early-retired might confuse the influence from reversed causality with the influence from a sample change. But since the income effects on health are linear (these results are available upon request), excluding subgroups should not affect the results.
Still, the financial resources of the individual obviously determine the health care of the individual, but in Sweden, with its egalitarian health policies, this is a rather unlikely mechanism for the age increase in health inequality (and a mechanism that, plausibly, would also run through the education level of the individual), even if there is some evidence of pro-rich distribution in doctors’ visits; i.e., despite the above mentioned egalitarian policy goals (Gerdtham, 1997; Gerdtham and Trivedi, 2001; van Doorslaer et al., 2006).

Our results have key policy interests for several reasons. One reason is that our results put emphasis of the effect of health on income in the explanation of the persistent socioeconomic health inequality, which has been reported in an endless number of studies. This information is vital to policy makers in developing effective strategies to reduce socioeconomic inequalities in health (Deaton, 2002).

Acknowledgements

We would like to thank the participants at seminars in Antwerp and Lund for their valuable comments, which have helped improve this paper. Financial support from the Swedish Council for Working Life and Social Research (dnr 2006-1217 and dnr 2007-0318), Government Grant for Clinical Research (“ALF”) and Region Skåne is gratefully acknowledged.
References


Why a positive link between age and income-related health inequality?


