

# Social Causation Versus Health Selection in the Life Course: Does Their Relative Importance Differ by Dimension of SES?

Rasmus Hoffmann<sup>1,2</sup>  · Hannes Kröger<sup>1,3</sup> · Siegfried Geyer<sup>4</sup>

Accepted: 2 March 2018  
© The Author(s) 2018

**Abstract** A person's socioeconomic status (SES) can affect health (social causation) and health can affect SES (health selection). The findings for each of these pathways may depend on how SES is measured. We study (1) whether social causation or health selection is more important for overall health inequalities, (2) whether this differs between stages of the life course, and (3) between measures of SES. Using retrospective survey data from 10 European countries (SHARELIFE,  $n = 18,734$ ), and structural equation models in a cross-lagged panel design, we determine the relative explanatory power of social causation and health selection through childhood, adulthood, and old age. We use three ways to measure SES: First, as a latent variable capturing different aspects of SES, second as material wealth, and third as occupational skill level. Between childhood and adulthood, social causation and health selection are equally important. In the transition from adulthood to old age, social causation becomes more important than health selection, making it the dominant mechanism in old age. The three measures of SES produce similar results. Only material wealth shows a stronger effect on health (between childhood and adulthood); it is also more affected by health (between adulthood and old age) than the other measures.

**Keywords** Socioeconomic status · Life course · Health inequality · Structural equation model · Social causation · Health selection

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s11205-018-1871-x>) contains supplementary material, which is available to authorized users.

✉ Rasmus Hoffmann  
rasmus.hoffmann@eui.eu

<sup>1</sup> Department of Political and Social Sciences, European University Institute, Via Dei Roccettini 9, 50014 San Domenico di Fiesole, Italy

<sup>2</sup> Max Planck Institute for Demographic Research, Konrad-Zuse-Straße 1, 18057 Rostock, Germany

<sup>3</sup> Socio-Economic Panel Study (SOEP) at the German Institute for Economic Research (DIW), Mohrenstraße 58, 10117 Berlin, Germany

<sup>4</sup> Hannover Medical School, Carl-Neuberg-Straße 1, 30625 Hannover, Germany

# 1 Introduction

That health is poorer and mortality is higher in groups with lower socioeconomic status (SES) has been documented in numerous studies and for all periods and countries for which data are available. Health inequalities usually amount to between 5 and 10 years' difference in life expectancy and between 10 and 20 years' difference in disability-free life expectancy between groups with low and high SES (Mackenbach 2006). Health inequalities can be analyzed with different indicators for SES, the most commonly used being education, occupational class, occupational status, and material factors such as income or wealth. While health differences have been found for all these SES indicators (Hoffmann 2008; Lahelma et al. 2006; Miething et al. 2013; Torssander and Erikson 2010), the size of the gradients differs between them, especially in multivariate analyses, and it is generally agreed that they should not be used interchangeably. This is because they represent different aspects of SES that are related to each other, and linked to health by different mechanisms (Geyer et al. 2006), each of them worth investigating in order to better understand the links between SES and health. This does not call into question SES as a construct. In case of SES indicators correlating highly, using only one would suffice for empirical research on health inequalities, and the others would thus be superfluous. However, due to high degrees of status inconsistency, correlations between SES indicators are usually small or moderate (Geyer 2008; Lahelma et al. 2008). This implies that it is necessary to consider all dimensions separately and to check whether the results differ from each other and from results with latent constructs that exploit joint variation of different SES dimensions. This approach has also been used with outcomes other than health, e.g. students' achievements in a cross-country comparison (Marks 2011).

A second line of research on health inequalities deals with the question of whether the relationships between SES and health are primarily due to the influence of SES on health (social causation) or to the influence of health on SES, or due to common background factors that influence both SES and health. Our study combines these two lines of analysis and investigates the relative explanatory power of social causation and health selection over the life course, and whether the results vary between age groups and dimensions of SES.

## 2 Background

### 2.1 Health Inequalities and the Causal Direction Between SES and Health

Morbidity and mortality rates are systematically higher among people with lower SES. While average health and life expectancy have improved over time in almost all countries, relative health inequalities are also increasing (Mackenbach et al. 2015b). The economic costs of health inequalities in the EU have been estimated at about 1000 billion € per year, which is 9.5% of EU GDP (Mackenbach et al. 2011). Despite the fact that health inequalities are an urgent and recognized public health problem (Elo 2009), there is still some controversy surrounding discussions on the fundamental mechanisms that create health inequalities (Chandola et al. 2003; Galama and van Kippersluis 2010). The debate is centered on the relative importance of social causation and health selection. Although both pathways involve causal effects, throughout the text we will refer to the first as *causation* and to the second as *selection*. A third causal model to explain health inequalities

is that (unknown) background factors influence both SES and health (indirect selection) (Goldman 2001a). These factors may be genetic endowment, family background, or individual characteristics (genetic or acquired), such as height, personality, or preferences with regard to behavior and lifestyle. This third causal model is most difficult to test empirically. Hence, in our study, we concentrate on the relative explanatory power of causation versus selection for the creation of health inequalities over the life course. We also deal with indirect selection, both empirically and by discussing to what extent it can be addressed in empirical studies by different approaches.

In disciplines such as health sociology and health economics, there is disagreement concerning the relative importance of causation versus selection, due not only to different underlying ideas of the relation between social structure and individual agency, but also to different research designs and methods, as well as divergent concepts of causality. Besides empirical findings that make social causation plausible, there are theoretical arguments against health selection which address the timing of cause and effect over the life course. For example, it is claimed that although health changes in middle and old age can no longer influence formal education, there are considerable health differences between educational groups (Haan et al. 1989). A central proposition of the selection hypothesis is that social mobility is partly due to health. While there are indications for a certain level of health-related social mobility in early adulthood when people enter the labor market (Smith 1999), the relationship between health and social mobility is fairly weak (Kröger 2015a). Moreover, the chronology of social mobility at younger ages and increasing health problems at older ages seems to contradict the proposition of the health selection hypothesis, which postulates the reverse order: pre-existing health differences going on to influence SES. While these arguments against health selection and related findings seem persuasive, “it is too easy to write off health selection as of little or no significance” (West 1991, p. 373), if one wrongly assumes that it works through mainly biological or genetic mechanisms: This ignores the social and sociological aspects, for example the social significance of health and processes of social mobility. From a methodological perspective, it is often overlooked that reverse causality from health to SES can bias the coefficients of conventional statistical models if the direction of causality is simply assumed (Hertzman et al. 1994).

Few epidemiological studies have examined the possibility of health selection (e.g. Blane et al. 1993; Chandola et al. 2003; Lundberg 1991) and the majority of authors in the fields of health sociology and social epidemiology believe that health selection is not very important (Goldman 2001b; Manor et al. 2003; Marmot et al. 1997). Although a small direct effect of education on health is widely agreed on Gathmann et al. (2015), many economists think that the influence of material resources on health is low, and that the influence of health on material status is the strongest overall causality in the relationship between SES and health (Galama and van Kippersluis 2010).

A recent literature review on the relative importance of causation versus selection evaluated 34 out of 2952 reviewed studies from the past 20 years, qualitatively and in a meta-analysis (Kröger et al. 2015). The result is that, across disciplinary boundaries, there is no preference for one of the two causal directions. 12 studies supported causation, and 10 supported selection—the other studies supported both directions equally.

Our first aim is to find out whether pathways of causation have more explanatory power than pathways of selection. The causal direction between SES and health is part of an ongoing discussion between different fields of social sciences, with several important normative and political implications regarding the acceptability of health inequality, and appropriate interventions to tackle it. It also raises complex methodological questions

concerning the analysis of causal effects in a longitudinal life course framework. We do not claim to reveal causal effects as they are understood in the potential outcome or counterfactual framework—for example, what happens to the SES of an average individual after a change in health or vice versa (Rubin 2005). Instead, we estimate how much of the variation in SES in our sample can be explained by health earlier in the life course and, complementarily, how much of the health variation can be explained by prior SES. Throughout the text we use the terms *effect* and *influence* in terms of explanatory power.

Our second aim is to establish whether our findings differ between age groups. It has been proposed that, to the extent that selection is important, it should peak in childhood (Marmot et al. 1997; West 1991) and in early old age, when inchoate health problems affect the ability to work (Smith 1998, 2003). Therefore, we study the interplay between SES and health in two broad stages of the life course. Finally, our third aim is to find out if there are differences in the relative importance of social causation and health selection between different dimensions of SES and different ways to measure it. For the mere relation between health and SES, differences between dimensions of SES have been found in earlier research (e.g. Elgar et al. 2016). Warren (2009) has looked at social causation versus health selection, testing different health measures. We add to this literature by comparing material factors and occupation as two dimensions of SES, and a latent variable for SES, to see how this affects the results for causation versus selection, which has not been done before. Beyond this, our study combines a number of innovative strengths; first, we start early in the life course by measuring the very beginning of the development of health and SES, gradually proceeding to old age. This is crucial for disentangling causality between related processes (Heckman 1981). Second, we use measurement models for latent variables, which reduces measurement error. The substantial influence of measurement error on results and conclusions in a cross-lagged panel design has been shown in previous research (Kröger et al. 2016b). Third, we use structural equation models that can simultaneously model two pathways (causation and selection), also taking into account indirect selection to the extent to which it creates a correlation between health and wealth.

## 2.2 Pathways Between SES and Health for Various SES Indicators

Studies on health inequality often choose one or more stratification variables, such as education, material wealth, or occupation, to measure the association between SES and health. The health selection hypothesis suggests that the same variables for social stratification can also be influenced by health; for example, poor health in childhood can limit educational achievement. In the following, we will briefly review existing evidence on the pathways between health and the most common SES indicators—education, occupation, and material wealth—illustrating why the relative importance of causation versus selection may vary between these indicators. We do not consider other determinants of health and health inequalities, such as place of living, gender, and ethnicity, that partly have to do with fixed non-social characteristics; we focus on dimensions of individual SES that are the result of social processes of distribution, and separate all analyses by gender.

*Education* means knowledge regarding health risks and health-promoting behavior. It provides cognitive skills for dealing with complex information, such as the effect of behavior on health and access to healthcare institutions. Better education reduces stress and provides more effective coping and flexible problem-solving behavior (Braveman 2006). Research on human capital has shown that education promotes cognitive and non-cognitive skill formation. These skills in turn facilitate the accumulation of health capital through

self-regulation and choices (Cunha and Heckman 2007). The positive health effect of education as 'learned effectiveness' is cumulative and self-amplifying, because education increases the sense of control, which shapes healthy life styles. The perceived success in controlling one's health determinants (e.g. weight control) creates incentives for further investments in health and other life domains (e.g. sporting activity) (Mirowsky and Ross 2003). The observed overall association between education and mortality can be partly explained by material factors and behavior, because higher education means higher income and more healthy behavior (Schrijvers et al. 1999). Better educated people also have more rewarding and healthier jobs, which is another indirect effect of education on health (Mirowsky and Ross 2003). The use of education as an indicator for SES is widespread, because of its simplicity, availability, and comparability, even internationally. However, its significance as an indicator of SES and its benefit for health is also disputed. Some authors argue that education is not a direct measure of SES, but rather a mechanism by which individuals gain positions, and that it has little direct effect on health (Blane 2006). Others claim that education is an important marker of SES (Lynch and Kaplan 2000). Analyses of school reforms as natural experiments have demonstrated a small positive causal effect of education on health (Gathmann et al. 2015), but there is also ample evidence that education is affected by poor health in childhood (Haas et al. 2011) because poor health may limit children's school participation and opportunities to develop educational and cognitive skills.

Although we consider education to be an important dimension of SES and to be closely related to health, we do not include it in our comparison of alternative SES measurements, because it is fixed in early adulthood and can only be influenced by health up to this point. Likewise, we do not interpret education as a common background factor for health and SES (West 1991). Instead, similar to Warren (2009), we represent education as a mediator that can contribute to causation and selection between childhood and adulthood, and can have a long-lasting effect on old age via adult SES and adult health.

*Occupation* influences health through the general position in society that a job can provide, and through physical and mental health risks at the workplace. Occupation is partly determined by education and in turn determines income, and may not have much effect on health net of education and income (Warren and Kuo 2003). However, there is a long research tradition demonstrating specific causal pathways from occupations with a negative effort-reward balance to stress and heart disease (Siegrist and Wahrendorf 2016). A related argument is that jobs that involve more productive self-expression, rather than self-suppression, favor health, and that better educated people are more likely to find such jobs (Mirowsky and Ross 2003). Occupation may also be affected by health selection, in the sense that health affects career trajectories, occupational status, and occupational class, because unhealthy people are less able to invest energy and time in their occupation and because of labor market discrimination against unhealthy people (Kröger 2015a).

The practical implementation of occupation as an SEP indicator is limited by the fact that not all people work, be they homemakers or retired people no longer exposed to current work conditions. Occupational class is thus considered less important than education and income for retired people (Hoffmann 2008; Huisman et al. 2003), but it is unknown to what extent the association between past occupational class and health decreases after retirement.

*Material wealth* has been shown to be strongly associated with mortality (Tarkiainen et al. 2012). It influences health and mortality through the affordability of health care, environmental hazards, consumption, and the psychological burden of being poor. Consumption includes a healthy diet, good housing in a safe environment, and quality healthcare.

Besides material explanations of the benefits of income, it also enhances effective capabilities, control, freedom, and the general ability to achieve goals (Mirowsky and Ross 2003; Robeyns 2011; Sen 1999). The effect of income on mortality has been found to be large compared to education and occupation (Torssander and Erikson 2010). Nevertheless, there is also a strong association between education and income, partly because higher education provides better opportunities on the labor market (Autor 2014). Consequently, income (or material conditions more generally) has indeed been found to partly mediate the effect of education and occupation on mortality (Schrijvers et al. 1999). Income and other material factors may also be influenced by health, because of two mechanisms: First, unhealthy people often earn less, because they invest less in their occupation and because they are less likely to be promoted (Contoyannis and Rice 2001); second, unhealthy people may spend financial resources on medical care expenditures (Smith 2004). In fact, for the association between material factors and health, there is wide-spread debate as to whether they reflect a causal effect of material factors on health or the effect of health on material wealth (Galama and van Kippersluis 2010; Martikainen et al. 2009).

Our hypothesis for the comparison of three different SES measures is that wages and income, as sub-dimensions of material wealth are especially prone to being influenced by health, because of their volatility. They should be more influenced by health than education (that can, by definition, only be influenced by health early in life) and more than occupational position, which also reacts more slowly to health than income. The opposing hypothesis is that all three measures reveal a similar relative importance between causation and selection, despite different mechanisms linking them to health. This hypothesis can be based on the fundamental causes hypothesis by Link and Phelan (1995), who argued that each dimension provides a set of resources that enables the individuals to develop and enact strategies in different circumstances that are beneficial to health. It is claimed that the socioeconomic status “embodies an array of resources, such as money, knowledge, prestige, power, and beneficial social connections, that protect health no matter what mechanisms are relevant at any given time” (Phelan et al. 2004, p. 265). A similar argument is that while each dimension of SES has a unique relation to health, between them, there is “some common element of social ordering that may be operating to influence health” (Adler and Ostrove 1999, p. 10). Others have claimed that “control” is one such overarching resource that is beneficial for health (Geyer 2016).

### 2.3 Life Course Framework

Health inequalities can be best explained by a life course approach (Mayer 2009), because they are the outcomes of the interaction of two life-long processes: First, the development and changes in a person’s SES; these changes can be understood as a process with critical periods and transitions, with path dependencies and accumulations (Dannefer 2003). Second, the development and changes in a person’s health, which are determined by fixed individual characteristics, social influences, behavior, and institutional settings, and also have critical periods, path dependencies, and accumulations of health problems or recovery (Kuh et al. 2003). The causal direction between SES and health can be studied best in a life course framework, because the interplay between SES and health is a long-term process. The life course perspective offers the additional possibility to check whether the relative importance of causation and selection changes over time, which indirectly also reflects the fact that different aspects of SES have their own unique timing and change their effect on, and their responsiveness to, health over time. For example, material factors

may be especially important for elderly people, because material factors take over the role of occupational class and status after retirement (Avlund et al. 2003). Studying such differences in a life course perspective reflects the processual character of SES, health, and health inequalities and is important for understanding health inequalities, measuring them at different ages, and also for addressing them with policies that might apply to different ages and life course stages.

### 3 Data and Methods

#### 3.1 Data

We use the third wave (SHARELIFE, version 5.0.0) of the Survey of Health Aging and Retirement in Europe (SHARE) (Börsch-Supan 2016). The data represents the population 55+ and their spouses living in households in the selected European countries. SHARELIFE contains information on health and SES at the time of the interview, and detailed retrospective information on events and changes in health and SES (Börsch-Supan et al. 2013). It has been used in numerous life course studies on SES and health (e.g. Kröger et al. 2016a; Pakpahan et al. 2017) The data was collected with personal interviews at home using computerized questionnaires. We limit our analysis to 10 countries (Austria, Belgium, Denmark, France, Germany, Italy, Netherlands, Spain, Sweden, and Switzerland) because for Poland, the Czech Republic, and Greece information on wages was not comparable over the life course or contained too many missing values. We also limit the analyses to persons aged 55–90 at the time of the interview in 2008/2009. The sample size is 18,734 and the average response rate across countries is approximately 60% (ranging from 40 to 80%). Details of participating countries are provided online (<http://www.share-project.org/data-access-documentation/sample.html>). For a description of the sample and the variables see Table 1.

#### 3.2 Measures

We divide the life course into three periods: childhood (age 0–15), adulthood (age 30–50), and old age (age 55–90). For childhood we use available indicators in SHARE that refer to childhood in general or to age 10. For adulthood we decided to start at age 30 when most educational trajectories are finished and individuals' own occupation and income can be reported with sufficient validity. The starting age of the oldest group is set at age 55, because we use the third wave of SHARE that was representative of the population aged 50+ in its first wave 6 years earlier.

For each of these stages, we employ three different strategies to measure SES: The first strategy is to use measurement models that estimate latent variables, using indicators from different dimensions of SES (composite measure). In childhood these indicators are: the number of books in the household, the number of rooms per person, and the occupational skill level of the father within four levels of the International Standard Classification of Occupations (ISCO). The first is an indicator of the educational level and the cultural capital of the household, the second is a measure of material welfare, and the third is a measure of occupational skill level of the father that is assumed to also influence the childhood conditions of the respondents. Education, occupation, and material welfare are related to health by the mechanisms discussed in Sect. 2.2. For adulthood, we use two indicators:

**Table 1** Description of the data (variables, categories, distributions)

Latent construct	Variable	Category	N = 18,734	
Country	West	Austria	945	5.0%
		Belgium	2584	13.8%
		France	2223	11.9%
		Germany	1762	9.4%
		Netherlands	2069	11.0%
		Switzerland	1157	6.2%
	South	Italy	2292	12.2%
		Spain	2035	10.9%
		Denmark	1806	9.6%
	Age in 2009 (Wave 3)	Sweden	1861	9.9%
		Mean	68.3	
		Min–Max	55–90	
Gender	Male	8598	45.9%	
	Female	10,136	54.1%	
Education (number of years)	Mean	10.6		
	Min–Max	0–25		
	Missing	2379		



**Table 1** (continued)

Latent construct	Variable	Category	N = 18,734		
CSES (Childhood SES—age 10)	Number of books	0–10 books	8118	43.9%	
		11–25 books	3939	21.3%	
		26–100 books	3916	21.2%	
		101–200 books	1216	6.6%	
		> 200 books	1299	7.0%	
	Father's occupational skill level Based on ISCO categories	Missing	246		
		ISCO 9	3619		20.5%
		ISCO 4-8	11,514		65.3%
		ISCO 3	834		4.7%
		ISCO 1-2	1668		9.5%
Rooms per capita	Missing	1099			
	Mean	0.8			
	Min–Max	0–10			
Number of facilities	Missing	326			
	Mean	2.0			
	Min–Max	0.0–5.0			
		Missing	129		

Table 1 (continued)

Latent construct	Variable	Category	N = 18,734	
CHEALTH (Childhood health—age 15)	Self-rated health	Poor	457	2.5%
		Fair	1235	6.7%
		Good	4777	25.8%
		Very good	5814	31.3%
		Excellent	6265	33.8%
		Missing	186	
		Yes	2166	11.6%
		No	16,441	88.4%
		Missing	127	
			Hospitalized	Yes
		No	17,489	93.8%
ASES (Adult SES—age 30–50)	Occupational skill level Based on ISCO categories	Missing	95	
		ISCO 9	2632	17.2%
		ISCO 4-8	8379	54.7%
		ISCO 3	1642	10.7%
		ISCO 1-2	2656	17.3%
		Missing	3425	
		Mean	1301	
AHEALTH (Adult health—age 30–50)	Average wages	Mean	1301	
		Min-Max	8–6126	
		Missing	9714	
		Mean	97.3	
	Percentage of years of non-illness	Mean	97.5	
	Percentage of years of non-poor health	Mean	97.5	
	Percentage of years of non-stress	Mean	90.8	

**Table 1** (continued)

Latent construct	Variable	Category	N = 18,734	
OSES (Old age SES—age 55–90)	Household income	Mean	35,289	
		Min–Max	0–755,089	
	Household wealth	Missing	2132	
		Mean	161,356	
		Min–Max	– 784,644 to 7153,102	
		Missing	618	
Supervisory status	None	11,157	67.7%	
	1–5 supervisees	2506	15.2%	
	6+ supervisees	2827	17.1%	
	Missing	2244		
OHEALTH (Old age health—age 55–90)	Self-rated health	Poor	2336	12.5%
		Fair	5116	27.4%
		Good	6895	37.0%
		Very good	2886	15.5%
	Grip strength	Excellent	1424	7.6%
		Missing	77	
		Mean	33.6	
	Min–Max	1–85		
	Missing	1576		

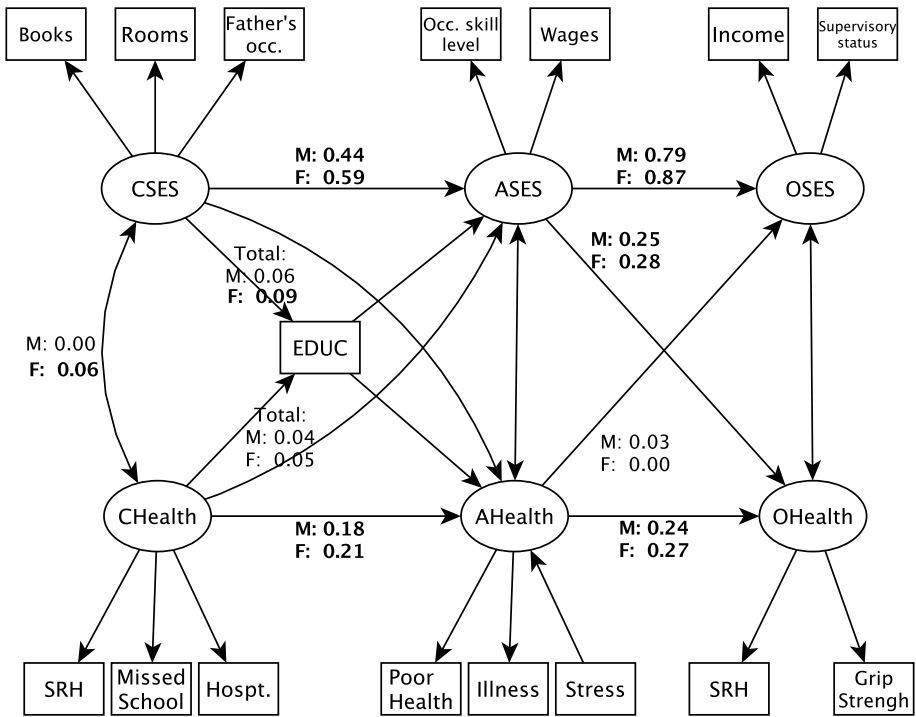
occupational skill level (ISCO) and estimates of average monthly wages over the 20 years between age 30 and 50, corrected for purchasing power and inflation by purchasing power parities (PPP) relative to German Euros in 2006 (Weiss 2012). Again, occupational skill level and income are assumed to be related to health by the mechanism described above. The skill level is based on the ten 1-digit ISCO categories that we transformed into four levels according to their skill level (see Table 1). We used the skill level with the longest duration between ages 30–50. In old age, we measure SES with the net-equivalent household income at the time of the interview (ages 55–90) and with the supervisory status, which is the number of employees (if self-employed) or the number of supervised persons (for employees) in the current or last job. These two variables reflect the material wealth and the occupational status of the respondents that are assumed to influence health. The different measurements of SES over age reflect both data availability and the fact that, at different stages of the life course, different resources are relevant for an individual's SES: first one's parents' and then one's own occupation and salary, and ultimately the job or pension-based income and status that has been acquired over the life course.

The second strategy is to focus on occupational position as an observed SES variable in each life course period. This is the occupational skill level of the father in childhood, one's own skill level in adulthood, and the supervisory status in the last or current job in old age. It is noteworthy that the overall SES associated with a certain skill level has changed between e.g. the 1950s and the 1990s, but we assume that despite distributional changes in the skill levels in the population, the ranking is still a valid indicator for this dimension of SES. The third strategy is to focus on material wealth as a measure of SES. In childhood, the indicators are the number of rooms per person and a summary index of features of the household (cold water, hot water, toilet, bath, heating). In adulthood we use wages, and in old age the net-equivalent household income. The variables used in each approach can be seen in Fig. 1 (SES as composite measure), Fig. 2 (material wealth), and Fig. 3 (occupational skill level).

Health in childhood is measured by three indicators: self-assessed health in five categories, whether school was missed because of health for 1 month or more, and whether 1 month or more was spent in hospital. In adulthood, our health measure is based on three indicators reflecting the number of years individuals reported suffering from bad health, illness, or stress, respectively. In old age, health is measured with the indicators current self-rated health and grip strength, combining a subjective and an objective health indicator in one latent variable. Self-rated health is considered to be a good health measure and predictor for mortality. It measures the absence of disease and is a comprehensive measure of impairment, disease, and distress (Jylhä 2009). Grip strength is an objective measure and has become a popular indicator of physical functioning in surveys, being indicative of overall muscular and physical functioning, and predictive of mortality (Syddall et al. 2017). We do not include further control variables (for example on health behavior) because our aim is to estimate the total effects between SES and health, and we expect other variables to mediate these effects.

### 3.3 Analysis

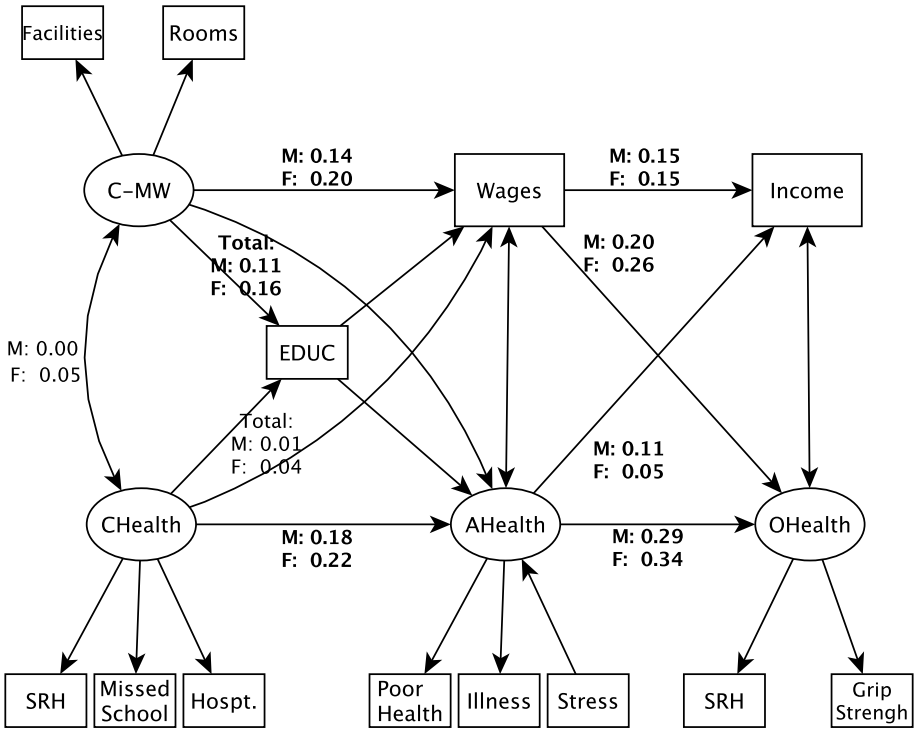
We chose a model-based approach to causal analysis, using life-long retrospective data to study the interplay between SES and health in two different stages of the life course. The advantage of a model-based approach, compared to design-based approaches such as quasi-experiments, is the potential for simultaneously modeling two related processes



**Fig. 1** Structural equation model for bidirectional relationships between SES and health over the life course, with standardized coefficients; SES measured as composite measure. *Note:* countries=Austria, Germany, Netherlands, France, Switzerland, Belgium, Sweden, Denmark, Spain, Italy; observed variables are shown as boxes and latent variables as ellipses; unidirectional arrows are path coefficients, bidirectional arrows are correlations; C=childhood (0–15); A=adulthood (30–50); O=old age (55–90); M/F=male/female; bold numbers are statistically significant ( $p < 0.05$ , two-tailed)

(causation and selection) in which the outcome of one process is the predictor of the other. The assumption of this approach is that all relevant confounders are accounted for, resulting in a quasi-random distribution of the stimulus (conditional independence). *Conditional independence* in such a model means that the two processes are independent, conditional on the history of the joint processes (Blossfeld et al. 2007). The main confounders in our study are the early histories of SES and health. The causal concept behind our approach is “causation as a generative process”, introduced by Cox (1992), and further developed in sociology (Blossfeld 2009; Goldthorpe 2001). It emphasizes the processual character of social phenomena and causal effects.

Observing almost complete life courses from parental and childhood characteristics to old age enables us to get as close as possible to a design that enables causal claims about SES and health. Many existing studies identify short-term causal effects between specific aspects of SES and health, for example with short time windows of five years and controlling for SES and health at baseline (Foverskov and Holm 2016). We claim that such studies suffer from endogeneity bias if they do not control for baseline status, because baseline health is always influenced by prior SES, and SES influenced by prior health. On the other hand, controlling for baseline status disregards the important history of SES and health, without which causal estimates will be biased, too. Instead, our long-term life

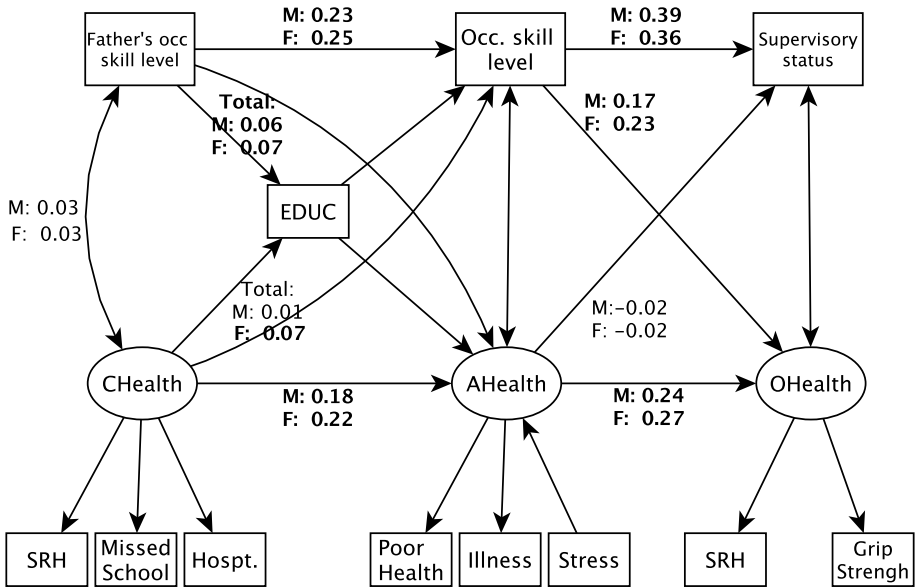


**Fig. 2** Structural equation model for bi-directed relationships between SES and health over the life course, with standardized coefficients; SES measured as material wealth. *Note:* see Fig. 1

course approach starts as early as possible and takes the joint history of SES and health into account.

We estimate the parameters of a structural equation model in a cross-lagged panel design (Bollen 1989; Pakpahan et al. 2015) that represents the relative importance of causation and selection in different stages of the life course (Fig. 1). The detailed description of the model in equation form can be found in the supplementary material. Structural equation models in a cross-lagged panel design have been used before to address the longitudinal interplay between SES and health (Warren 2009), dimensions of SES and deprivation (Laurijssen and Spruyt 2015), or between health and life satisfaction (Headey and Muffels 2016). A structural equation model consists of two kinds of model: measurement models to estimate latent variables based on observed indicators, and a structural model that estimates the paths between latent or observed variables. We model SES and health at three different stages as latent variables, except for the models where only one variable is used to measure SES (see Figs. 1, 2, 3). Our structural model estimates the paths' parameters between these latent or observed variables and observed education.

The parameters are estimated using mean and variance-adjusted weighted least squares (WLSMV) that have no distribution assumption (Finney and DiStefano 2006). We present standardized coefficients in a uniform value range of  $-1$  to  $1$ , making them comparable across paths and models. Our model takes the correlation between SES and health at all ages into account. Especially in childhood, this correlation can be due to common



**Fig. 3** Structural equation model for bi-directed relationships between SES and health over the life course, with standardized coefficients; SES measured as occupational position. *Note:* see Fig. 1

unobserved background factors, for example genetic factors or unobserved characteristics of the family. Thus, we address the common background factors mentioned above to the extent that such factors create a correlation between health and SES. The path parameters can be divided into two groups: first, the autoregressive parameters showing the effect of SES at  $t_1$  on SES at  $t_2$  (and the same for health); second, the cross-lagged parameters, showing how SES at  $t_1$  influences health at  $t_2$  (causation) or health at  $t_1$  influences SES at  $t_2$  (selection). The cross-lagged coefficients in the first life course transition (from childhood to adulthood) can be subdivided into direct and indirect effects, the latter being mediated by education. Direct and indirect effects can be combined to give the total effect. In previous analyses, we had added direct paths from childhood to old age, but these were not statistically significant and did not change the overall results (not shown below). This suggests that most of these effects are moderated by SES and health in adulthood (Pakpahan et al. 2017). For the same reason, we do not model direct paths from education to old age outcomes.

We calculated the models on two levels: first, one analysis for all 10 countries combined; second, with countries grouped into three European regions (West, South, North) (only shown in the supplementary material). The regions roughly reflect different welfare state models: Austria, Germany, Netherlands, France, Switzerland, and Belgium represent West European welfare states, Denmark and Sweden the Scandinavian welfare model, and Spain and Italy Southern Europe. We consider this to be an explorative approach that illustrates the acceptable degree of heterogeneity within SHARE, and we do not claim to test specific welfare-state hypotheses. We use country dummies to control for unobserved national differences. We apply cross-sectional weights, provided by SHARE, that account for the unequal probability of becoming a sample member, and weight the observations to reflect differences in population size between countries. All models are calculated separately for

men and women in order to check systematic gender differences in the relation between SES and health. Age at interview, also reflecting the birth cohort, is a control variable, but shows no significant effect on the results. Data preparation is performed in Stata 14.1, including user-written programs (Kröger 2015b), and analyses in Mplus 7.4 (Muthen and Muthen 2015).

## 4 Results

Results from structural equation models are shown in Figs. 1, 2 and 3 (as a graphical illustration of the model and the different types of measurement), in Table 2 (all coefficients, standard errors, and goodness-of-fit measures, including direct and indirect effects), and in Fig. 4 (only coefficients for comparing causation and selection). Factor loadings for all measurement models are given in Supplementary Table 1. The following presentation of the results refers to Table 2, while Figs. 1, 2 and 3 serve as a visual simplification of the main outcomes and different measurements.

When the cross-lagged effects are taken into account, the correlations between SES and health in each age group is generally negligible, except among men when the composite first measurement type ('SES') is used, where this correlation clearly increases with age, and reaches 0.47 in old age. The autoregressive coefficients for both socioeconomic status and health in Phase 1 (the transition from childhood to adulthood) range between 0.14 and 0.25, except for the path from childhood SES (CSES) to adult SES (ASES); this is 0.44 for men and 0.59 for women when the composite measurement is used. The coefficients can be interpreted as follows: 0.59 means that one standard deviation increase in CSES predicts a 0.59 standard deviation increase in ASES, or that 59% of the variation of ASES can be explained by CSES. These results show that SES and health depend on their prior status, and this effect is greater when using the composite measure of socioeconomic status.

CSES has a lot of explanatory power for education, which in turn explains to some extent adult health (AHEALTH). These two paths multiplied show the indirect causation path between childhood and adulthood. The direct causation path is weak, as is the total causation path from CSES to AHEALTH. Among the mechanisms that contribute to selection in Phase 1, the only strong path is that from education on ASES, but since health in childhood does not predict education, the total effect of selection is even smaller than the total effect of causation: It only reaches statistical significance among women when looking at occupational measures. To summarize the role of education: It is a strong mediator between CSES and ASES and displays some contribution to causation, but none to selection, because it does not depend on health in childhood. That means that, in the 10 SHARE countries, we cannot confirm previous findings that health in childhood affects adult SES through educational attainment (Haas et al. 2011).

In phase 2 (the transition from adulthood to old age) the model is simpler because it does not involve indirect effects. As in Phase 1, the autoregressive parameters are high, and again highest for the composite type of measurement ('SES'). Causation is clearly higher than selection and always statistically significant.

The goodness-of-fit indicators show variable results: The Chi square is always significant, while the Root Mean Square Error of Approximation (RMSEA) shows very good fit for all models, i.e. the upper bound of the confidence interval is below 0.05. The Comparative Fit Index (CFI) is borderline, with a range between 0.839 and 0.902. We think that the lower values of CFI are acceptable, because the purpose of the study



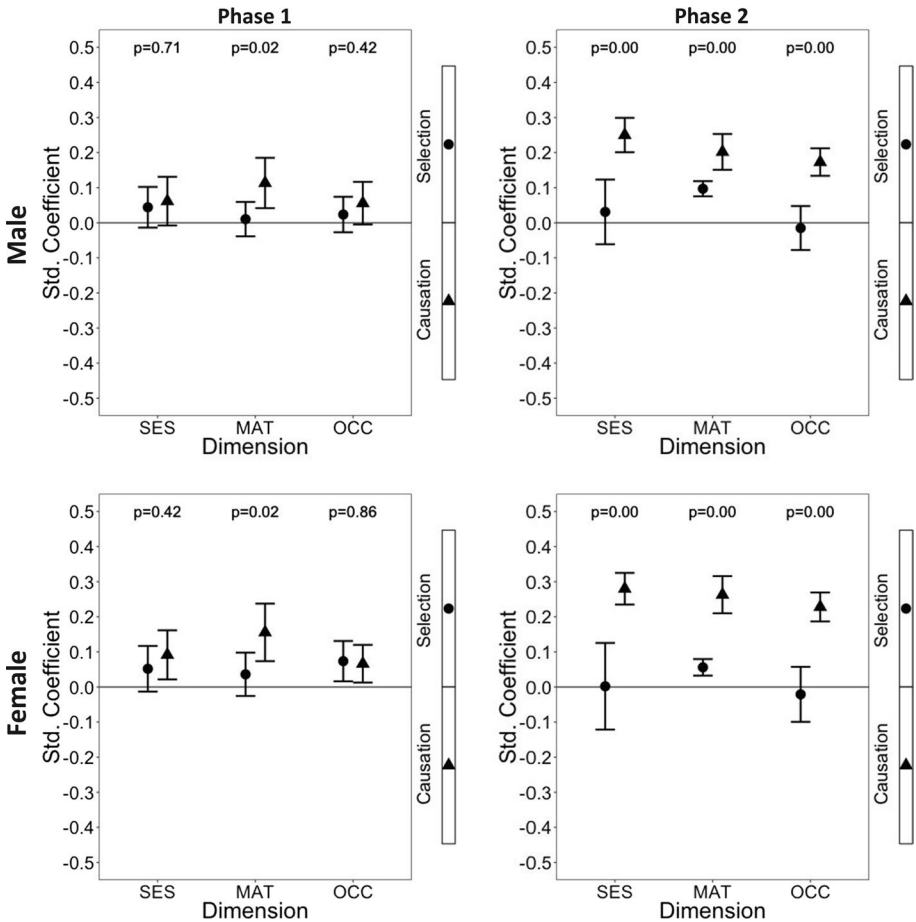
**Table 2** Results from structural equation models on the relation between the socioeconomic status and health over the life course

Parameter	Male			Female		
	SES	MAT	OCC	SES	MAT	OCC
	Coef. S.E.	Coef. S.E.	Coef. S.E.	Coef. S.E.	Coef. S.E.	Coef. S.E.
Correlation	CSES ↔ CHEALTH	0.00	0.00	0.03	0.05	0.03
	ASES ↔ AHEALTH	0.03	0.03	0.03	0.03	0.03
Phase 1	ASES ↔ AHEALTH	<b>0.21</b>	<b>-0.11</b>	<b>0.10</b>	<b>-0.19</b>	<b>-0.08</b>
	OSES ↔ OHEALTH	0.05	0.04	0.03	0.04	0.04
Autoregression	OSES ↔ OHEALTH	<b>0.47</b>	<b>0.10</b>	<b>0.19</b>	<b>0.08</b>	0.03
		0.12	0.03	0.04	0.04	0.05
Causation	CSES → ASES	<b>0.44</b>	<b>0.14</b>	<b>0.23</b>	<b>0.20</b>	<b>0.25</b>
	CHEALTH → AHEALTH	0.03	0.03	0.02	0.04	0.02
Indirect 1	CHEALTH → AHEALTH	<b>0.18</b>	<b>0.18</b>	<b>0.18</b>	<b>0.22</b>	<b>0.22</b>
		0.03	0.03	0.03	0.03	0.03
Indirect 2	CSES → EDUC	<b>0.57</b>	<b>0.54</b>	<b>0.32</b>	<b>0.53</b>	<b>0.26</b>
	EDUC → AHEALTH	0.01	0.02	0.01	0.01	0.01
Indirect 1 × Indirect 2	EDUC → AHEALTH	<b>0.14</b>	<b>0.19</b>	<b>0.15</b>	<b>0.11</b>	<b>0.11</b>
	(CSES → EDUC) × (EDUC → AHEALTH)	0.04	0.04	0.03	0.04	0.03
Direct	CSES → AHEALTH	<b>0.08</b>	<b>0.10</b>	<b>0.05</b>	<b>0.06</b>	<b>0.03</b>
		0.02	0.02	0.01	0.02	0.01
Total	CSES → AHEALTH	-0.02	0.01	0.01	<b>0.10</b>	0.04
		0.05	0.05	0.03	0.05	0.03
Selection	CSES → AHEALTH	0.06	<b>0.11</b>	<b>0.06</b>	<b>0.16</b>	<b>0.07</b>
		0.04	0.04	0.03	0.04	0.03
Indirect 1	CHEALTH → EDUC	0.01	0.00	-0.01	-0.03	-0.01
		0.02	0.02	0.02	0.02	0.02
Indirect 2	EDUC → ASES	<b>0.46</b>	<b>0.23</b>	<b>0.54</b>	<b>0.24</b>	<b>0.53</b>
		0.02	0.02	0.01	0.03	0.02

**Table 2** (continued)

Parameter	Male			Female		
	SES	MAT	OCC	SES	MAT	OCC
	Indirect 1 × Indirect 2 (CHEALTH → EDUC) × (EDUC → ASES)	0.00	0.00	0.00	- 0.01	- <b>0.01</b>
Direct	Coef.	0.01	0.01	0.01	0.00	0.01
	S.E.	0.04	0.01	0.03	<b>0.06</b>	<b>0.08</b>
	Coef.	0.03	<b>0.02</b>	0.02	0.03	0.03
Total	S.E.	0.04	0.01	0.02	0.05	<b>0.07</b>
	Coef.	0.03	0.02	0.03	0.03	0.03
	S.E.	<b>0.79</b>	<b>0.15</b>	<b>0.39</b>	<b>0.87</b>	<b>0.15</b>
Phase 2 Autoregression	Coef.	0.06	0.02	0.02	0.09	0.02
	S.E.	<b>0.24</b>	<b>0.29</b>	<b>0.24</b>	<b>0.27</b>	<b>0.27</b>
	Coef.	0.02	0.02	0.02	0.02	0.02
Causation	S.E.	<b>0.25</b>	<b>0.20</b>	<b>0.17</b>	<b>0.28</b>	<b>0.23</b>
	Coef.	0.02	0.03	0.02	0.02	0.02
	S.E.	0.03	<b>0.10</b>	- 0.02	0.00	<b>0.06</b>
Selection	Coef.	0.05	0.01	0.03	0.06	0.04
	S.E.	2109	1339	1067	1878	1119
	$\chi^2$	0	0	0	0	0
RMSEA 90% confidence interval	<i>p</i> value	0.851	0.866	0.865	0.902	0.847
	CFI	0.036	0.037	0.037	0.031	0.035
	RMSEA	0.035-0.038	0.036-0.039	0.035-0.039	0.03-0.038	0.035-0.038

Standardized regression coefficients; countries = Austria, Germany, Netherlands, France, Switzerland, Belgium, Sweden, Denmark, Spain, Italy; S.E. = standard errors; C = childhood; A = adulthood (30-50); O = old age (55-90); SES = socioeconomic status; the three types of measurement are SES (composite measure), MAT (material factors), and OCC (occupational position); Phase 1 = transition from childhood to adulthood; Phase 2 = transition from adulthood to old age; for interpretation of the coefficients: e.g. 0.5 means that one standard deviation change in the independent variable results in 0.5 standard deviation change in the dependent variable; statistically significant coefficients ( $p < 0.05$ , two-tailed) are printed in bold



**Fig. 4** Relative explanatory power of social causation and health selection, by life course phase, gender, and type of measure for socioeconomic status (SES, MAT, OCC). *Note:* countries=Austria, Germany, Netherlands, France, Switzerland, Belgium, Sweden, Denmark, Spain, Italy; Phase 1=transition from childhood to adult age; Phase 2=transition from adulthood to old age; the confidence intervals show whether an estimate is different from zero ( $p < 0.05$ , two-tailed), while the  $p$  values are from a direct Wald-test for difference between the standardized coefficient for causation and selection

is the comparison of the causation and selection pathways, and not to develop a best possible model of all interrelationships between material factors and health throughout the life course.

For a more direct comparison of the two reciprocal effects, Fig. 2 shows only the related coefficients, by life course stages, gender, and type of SES measure. The confidence intervals in the graphs indicate whether a coefficient is statistically significantly different from zero, while the  $p$  values are from a Wald test of the difference between the two coefficients for causation and selection. For men in Phase 1, both pathways have the same low explanatory power. Only for the material measurement is causation higher than selection. The same is true for women in Phase 1. In the transition from adulthood to old age (phase 2), causation has much more explanatory power than in phase 1,

and much more than selection, which makes causation the dominant pathway in phase 2 for both men and women. This overall pattern is the same across all types of measurement, with two exceptions regarding the material measurement: First, in phase 1, it affects health more than the other SES measures, which means that childhood health affects adult wages more than adult occupational skill level. Second, in phase 2, material wealth is slightly more affected by health than the supervisory status. It is noteworthy that there are no substantial gender differences in our results.

We also calculated results for two European regions and present them in Supplementary Figures 1 and 2. In the North, the results deviate from the SHARE average for men in phase 2 for two types of measures (SES and OCC), and in the Western region no deviations appear. We conclude that these regional differences do not suggest a systematic pattern of differences for certain regions and that the result for the aggregated SHARE sample does not consist of too heterogeneous or inconsistent sub-samples. Unfortunately, the model with the two Southern countries did not converge, which shows that an analysis with only two countries can easily bring the sample size below the minimum needed for the estimation of our complex model.

## 5 Discussion

This study showed from a comprehensive life course perspective that, firstly, health and all three measures of SES substantially depend on their prior status and, secondly, in the transition from childhood to adulthood, that the social causation path was as important as the health selection path, while in the transition from adulthood to old age, causation was much more important than selection.

We found neither systematic gender differences nor clear differences between European regions that would allow the conclusion that welfare systems substantially influence the relationship between SES and health on this very general level. We acknowledge the possible effects of national historical settings on our findings. To the extent that these are fixed country effects, they are addressed by controlling for country.

In the transition from childhood to adulthood (phase 1), selection has very low explanatory power for all SES measures, which raises doubts about the influence of childhood health on education, occupation, or income in adulthood. Causation is only slightly stronger, and is mainly due to indirect effects via education. In the transition from adulthood to old age (phase 2) the explanatory power of selection remains weak, while causation is much stronger compared to phase 1 and compared to selection. This contradicts previous evidence that selection is especially important in older working ages, where many health problems start to become more prevalent, which influences the ability to work up into early old age and secure a high SES (Smith 1998, 2003). Unfortunately, due to lack of case numbers, our data do not permit further stratifications by age. Future research should investigate the role of selection in higher ages by separating those who still work from retired persons.

Our main differentiation into three types of SES measures did not reveal large differences. All three measures show the same results with respect to the broad observed changes over age. However, material wealth shows two slight deviations from the overall pattern, which is a confirmation of previous findings and our assumption that income reacts more to health than occupational position. However, the size of these differences between measures is small. We confirm that different dimensions of SES exhibit similar bi-directed

relations to health over the life course, which gives some credibility to the fundamental cause hypothesis that has been tested more directly elsewhere (Mackenbach et al. 2015a). Our findings do not confirm the special role of material wealth as an SES dimension that is heavily affected by health (Galama and van Kippersluis 2010; Martikainen et al. 2009). Different time frames matter for the role of material wealth and may be responsible for inconsistent findings in the literature.

In order to compare our findings with earlier research, three studies were selected that use similar age groups, indicators, and methods. The first study uses prospective data ( $n=705$ ) from the USA covering an age range of 41–88 years, corresponding only to our second life phase, and comparing the cross-lagged effects between two interviews in 1974 and 1994. It shows that both causal directions are present—causation slightly more than selection—and with only small gender differences (Mulatu and Schooler 2002). They also use a latent variable for SES with the indicators education, family income, and occupation. In a sensitivity analysis they show that adding paths from income and education to health, as two separate resources that have effects beyond the effect of SES, does not change the main results. In our study, we prefer to directly compare three ways of measuring SES (latent variable, occupation, material wealth), because it is simpler and more informative for the question if different measures produce different results. Mulatu and Schooler provide evidence on mediators, e.g. differences in weight and sleeping behavior that seem to matter moderately in the process of health selection. Instead, we look at education as a mediator of effects between childhood and adulthood and show that it mediates causation more than selection.

The second study using Finnish registry data ( $n=211,639$ ) covering an age range of 17–66 years shows causation to be slightly more important than selection (Aittomäki et al. 2012). Despite a large sample with high-quality register data, the indicator for health is limited to sickness allowances. This reflects absence from work due to sickness, but many other forms of poor health are not covered and, more importantly for a life course perspective, this definition of health cannot cover childhood and retirement ages.

The third study uses data from the US Wisconsin Longitudinal Study ( $n=5290$ ) which is a prospective study in the age range 18–65 years, that also contains retrospective information, and finds no evidence for health selection (Warren 2009). Like our results, Warren also finds no effect of childhood health on educational achievement. As mentioned above, Warren examines the robustness of the results for different health measures and finds no substantial differences between them. In a complementary fashion, our study examines different ways to measure SES, and does not reveal substantial differences between them.

We identified eight previous studies that not only compared social causation and health selection, but that traced different dimensions of socioeconomic position from childhood into adulthood in a life course perspective (Case et al. 2005; Eaton et al. 2001; Elovainio et al. 2011, 2012; Huurre et al. 2005; Palloni et al. 2009; Power et al. 2002; Stansfeld et al. 2011). Most of these studies ended in midlife. We summarize their main characteristics in Supplementary Table 2. Of the eight studies, three gave preference to social causation, two to health selection, one gave equal weight to both mechanisms, one said none was important and one did not draw a direct conclusion on the question. None of these studies directly compared different indicators for SES and none compared the results on the causal question between young and old age groups.

Some limitations to our approach are noteworthy: First, the fact that our data cover a long time span comes at the cost of using retrospective data that can be affected by recall bias (Smith and Thomas 2003). To maximize accuracy in remembering the occurrence and temporal order of events in the life course, SHARE uses a life grid approach

that visualizes the life course to the respondent in a calendar (Blane et al. 2007; Börsch-Supan and Schröder 2011). Another advantage of retrospective data over many prospective cohort studies is that we have continuous yearly information on occupation from age 16 onwards. This allows us to average occupational skill level during midlife, which should reduce measurement error and recall bias. A problem remains if current health status influences the answers to the retrospective health status or SES. In this case, measurement error is no longer random, but systematic, and can bias the results in an unknown direction. However, in the measurement of health in old age, our modeling approach using a latent variable takes only the common variation of self-rated health and grip strength into account. This means that the objective health status in old age would need to influence the reporting behavior on childhood health status. We cannot rule out this possibility, but we think that it is much less likely than in studies only using a subjective indicator. If we do have differential measurement error due to current health status and this relates only to health, we would be overestimating the autoregressive parameters in the health equations. If current health status also influenced reporting of SES dimensions in childhood, we might also get indirect bias into the estimates of the social causation and health selection pathways. However, the indicators of childhood SES are mostly reports of simple objective items and therefore should be less prone to this systematic type of measurement error. Besides these principle concerns, there are studies demonstrating that retrospective reports on income or health are influenced and biased by current conditions, mostly making the past similar to the present (Schwarz 2007). However, several studies have also shown that the retrospective measurement of health and SES, including the SHARE data, is relatively valid (Garrouste and Paccagnella 2011; Haas 2007; Mazzonna and Havari 2011). The remaining disadvantages of retrospective data need to be balanced with the fact that it enables the study of longer periods than in previous research based on prospective data (Adams et al. 2003; Stowasser et al. 2011), and that retrospective data does not suffer from panel attrition (Corna 2013).

Second, our multiple indicators for SES and health show differences in the percentage of missing values. In general, survey data has high percentages of missing values, and this is especially prevalent in sensitive questions on financial resources. The highest percentage of missing values in our data is for the variable ‘wages’ (51.9%), one of the indicators for adult SES (see Table 1). All other variables have much lower numbers of missing values, for example occupation (18.3%) and income (11.4%). We performed a sensitivity analysis, excluding cases with missing wages, and found that this does not change the main results (Supplementary Figure 3): In phase 2, two significant differences between causation and selection disappeared, for material measures among men and for SES measures among women, which is mainly due to the increase of the health selection effect. This points to the fact that those who are more tightly bound to the labor market (and therefore can report sufficient information on their wages in adult life) are more subject to health-related selection, as was argued in Kröger et al. (2015). Some of these missing wages are from individuals who are primarily taking care of home and family. We thus conducted a second sensitivity analysis excluding 1241 women and 51 men who reported, for at least 75% of the years between ages 30 and 50, to have mainly worked in the household. This sensitivity analysis produced the same results as the main analysis (Supplementary Figure 4).

Third, the high percentages of good health in childhood and adulthood may produce imprecise estimates or it may even bias downwards the coefficients for social causation, because there is only little variance in health to be explained by SES. We concede that a different measure for adult health with more variance could have resulted in higher

coefficients for social causation in Phase 1. On the other hand, the low prevalence of bad health in adulthood is not unrealistic and suggests that differences in childhood SES did not influence health in adulthood to a large extent.

Fourth, our analysis does not take mortality into account, which is a good health indicator, but logically cannot be used to predict changes in SES. We can only claim to study the surviving population. This might be selected, but we assume that, while selective mortality decreases health inequality in the surviving population, e.g. by impoverished, unhealthy people dying first, it does not systematically and substantially bias the comparison between social causation and health selection. Besides this, a general health-related participation bias might also apply to the SHARE sample in the sense that less healthy people are less likely to participate.

In conclusion, our data and analyses deliver valuable insight into the complex relationship between SES and health over the life course. Future improvements in the availability of comprehensive life course data would enable a more detailed analysis, a more accurate chronology of causes and effects, and better identification of policy-relevant mechanisms to tackle health inequality. In the absence of good observable variables for common background factors, it is crucial to start the observation as early as possible to avoid endogeneity bias. Causation and selection both contribute to health inequalities, but the main pathway is from SES to health, regardless of the exact SES variable used.

**Acknowledgements** Open access funding provided by Max Planck Society. This paper uses data from SHARE wave 4 release 1.1.1, as of March 28th 2013 (<https://doi.org/10.6103/share.w4.111>) or SHARE waves 1 and 2 release 2.6.0, as of November 29th 2013 (DOIs: <https://doi.org/10.6103/share.w1.260> and <https://doi.org/10.6103/share.w2.260>) or SHARELIFE release 1.0.0, as of November 24th 2010 (<https://doi.org/10.6103/share.w3.100>). The SHARE data collection has been primarily funded by the European Commission through the 5th Framework Programme (Project QLK6-CT-2001-00360 in the thematic programme Quality of Life), through the 6th Framework Programme (Projects SHARE-I3, RII-CT-2006-062193, COMPARE, CIT5-CT-2005-028857, and SHARELIFE, CIT4-CT-2006-028812) and through the 7th Framework Programme (SHARE-PREP, N° 211909, SHARE-LEAP, N° 227822 and SHARE M4, N° 261982). Additional funding from the U.S. National Institute on Aging (U01 AG09740-13S2, P01 AG005842, P01 AG08291, P30 AG12815, R21 AG025169, Y1-AG-4553-01, IAG BSR06-11 and OGHA 04-064) and the German Ministry of Education and Research, as well as from various national sources, is gratefully acknowledged (see [www.share-project.org](http://www.share-project.org) for a full list of funding institutions). We thank Dr. Eduwin Pakpahan for his contributions to the early stages of data preparation and analysis.

**Funding** This work was supported by a Starting Grant from the European Research Council [Grant Number 313532].

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Adams, P., Hurd, M. D., McFadden, D., Merrill, A., & Ribeiro, T. (2003). Healthy, wealthy, and wise? Tests for direct causal paths between health and socioeconomic status. *Journal of Econometrics*, *112*(1), 3–56.
- Adler, N. E., & Ostrove, J. M. (1999). Socioeconomic status and health: What we know and what we don't. *Annals of the New York Academy of Sciences*, *896*(1), 3–15.

- Aittomäki, A., Martikainen, P., Laaksonen, M., Lahelma, E., & Rahkonen, O. (2012). Household economic resources, labour-market advantage and health problems—A study on causal relationships using prospective register data. *Social Science and Medicine*, *75*(7), 1303–1310.
- Autor, D. H. (2014). Skills, education, and the rise of earnings inequality among the “other 99 percent”. *Science*, *344*(6186), 843–851.
- Avlund, K., Holstein, B. E., Osler, M., Damsgaard, M. T., Holm-Pedersen, P., & Rasmussen, N. K. (2003). Social position and health in old age: The relevance of different indicators of social position. *Scandinavian Journal of Public Health*, *31*(2), 126–136.
- Blane, D. (2006). Commentary: The place in life course research of validated measures of socioeconomic position. *International Journal of Epidemiology*, *35*, 139–140.
- Blane, D., Davey Smith, G., & Bartley, M. (1993). Social selection: What does it contribute to social class differences in health? *Sociology of Health & Illness*, *15*(1), 1–15.
- Blane, D., Netuveli, G., & Stone, J. (2007). The development of life course epidemiology. *Revue d'Épidémiologie et de Santé Publique*, *55*, 31–38.
- Blossfeld, H.-P. (2009). Causation as a generative process. The elaboration of an idea for the social sciences and an application to an analysis of an interdependent dynamic social system. In H. Engelhardt, H.-P. Kohler, & A. Fürnkranz-Prskawetz (Eds.), *Causal analysis in population studies, the Springer series on demographic methods and population analysis* (Vol. 23, pp. 83–109). Dordrecht: Springer Netherlands.
- Blossfeld, H.-P., Golsch, K., & Rohwer, G. (2007). *Event history analysis with stata* (1st ed.). New York: Psychology Press.
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York, NY: Wiley-Interscience.
- Börsch-Supan, A. (2016). Survey of Health, Ageing and Retirement in Europe (SHARELIFE) Wave 3. Release version: 5.0.0. SHARE-ERIC. Data set.
- Börsch-Supan, A., Brandt, M., Hunkler, C., Kneip, T., Korbmacher, J., Malter, F., et al. (2013). Data resource profile: The Survey of Health, Ageing and Retirement in Europe (SHARE). *International Journal of Epidemiology*, *42*(4), 992–1001.
- Braveman, P. (2006). Health disparities and health equity: Concepts and measurement. *Annual Review of Public Health*, *27*, 167–194.
- Case, A., Fertig, A., & Paxson, C. (2005). The lasting impact of childhood health and circumstance. *Journal of Health Economics*, *24*, 365–389.
- Chandola, T., Bartley, M., Sacker, A., Jenkinson, C., & Marmot, M. (2003). Health selection in the Whitehall II study, UK. *Social Science and Medicine*, *56*, 2059–2072.
- Contoyannis, P., & Rice, N. (2001). The impact of health on wages: Evidence from the British Household Panel Survey. *Empirical Economics*, *26*(4), 599–622.
- Corna, L. M. (2013). A life course perspective on socioeconomic inequalities in health: A critical review of conceptual frameworks. *Advances in Life Course Research*, *18*(2), 150–159.
- Cox, D. R. (1992). Causality: Some statistical aspects. *Journal of the Royal Statistical Society. Series A*, *155*, 291–301.
- Cunha, F., & Heckman, J. (2007). The technology of skill formation. *American Economic Review*, *97*(2), 31–47.
- Dannefer, D. (2003). Cumulative advantage/disadvantage and the life course: Cross-fertilizing age and social science theory. *The Journals of Gerontology, Series B: Psychological and Social Sciences*, *58*(6), 327–337.
- Eaton, W. W., Muntaner, C., Bovasso, G., & Smith, C. (2001). Socioeconomic status and depressive syndrome: The role of inter- and intra-generational mobility, government assistance, and work environment. *Journal of Health and Social Behavior*, *42*(3), 277.
- Elgar, F. J., McKinnon, B., Torsheim, T., Schnohr, C. W., Mazur, J., Cavallo, F., et al. (2016). Patterns of socioeconomic inequality in adolescent health differ according to the measure of socioeconomic position. *Social Indicators Research*, *127*(3), 1169–1180.
- Elo, I. T. (2009). Social class differentials in health and mortality: Patterns and explanations in comparative perspective. *Annual Review of Sociology*, *35*, 553–572.
- Elovainio, M., Ferrie, J. E., Singh-Manoux, A., Shipley, M., Batty, G. D., Head, J., et al. (2011). Socioeconomic differences in cardiometabolic factors: Social causation or health-related selection? Evidence from the Whitehall II Cohort Study, 1991–2004. *American Journal of Epidemiology*, *174*(7), 779–789.
- Elovainio, M., Pulkki-Råback, L., Jokela, M., Kivimäki, M., Hintsanen, M., Hintsu, T., et al. (2012). Socioeconomic status and the development of depressive symptoms from childhood to adulthood: A longitudinal analysis across 27 years of follow-up in the Young Finns study. *Social Science and Medicine*, *74*(6), 923–929.



- Finney, S. J., & DiStefano, C. (2006). Non-normal and categorical data in structural equation modeling. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 269–314). Greenwich, CT: Information Age Publishing Inc.
- Foverskov, E., & Holm, A. (2016). Socioeconomic inequality in health in the British household panel: Tests of the social causation, health selection and the indirect selection hypothesis using dynamic fixed effects panel models. *Social Science and Medicine*, *150*, 172–183.
- Galama, T., & van Kippersluis, H. (2010). *A theory of socioeconomic disparities in health over the life cycle*. RAND Labor and Population working paper series WR-773.
- Garroute, C., & Paccagnella, O. (2011). Three examples of consistency across SHARE and SHARE-LIFE data. In M. Schröder (Ed.), *Retrospective data collection in the Survey of Health, Ageing and Retirement in Europe—SHARELIFE methodology* (pp. 62–72). Mannheim: Mannheim Research Institute for the Economics of Ageing.
- Gathmann, C., Jürges, H., & Reinhold, S. (2015). Compulsory schooling reforms, education and mortality in twentieth century Europe. *Social Science and Medicine*, *127*, 74–82.
- Geyer, S. (2008). Einzelindikator oder Index? Maße sozialer Differenzierung im Vergleich [Single indicator or index? Comparison of measures of social differentiation]. *Das Gesundheitswesen*, *70*(5), 281–288.
- Geyer, S. (2016). Soziale Ungleichverteilungen von Gesundheit und Krankheit und ihre Erklärungen - Schulbildung, Einkommen und Beruf. In M. Jungbauer-Gans & P. Kriwy (Eds.), *Handbuch Gesundheitssoziologie*. Wiesbaden: Springer.
- Geyer, S., Hemström, Ö., Peter, R., & Vågerö, D. (2006). Education, income, and occupational class cannot be used interchangeably in social epidemiology. Empirical evidence against a common practice. *Journal of Epidemiology and Community Health*, *60*(9), 804–810.
- Goldman, N. (2001a). Social inequalities in health: Disentangling the underlying mechanisms. *Annals of the New York Academy of Sciences*, *954*, 118–139.
- Goldman, N. (2001b). Mortality differentials: Selection and causation. In N. J. Smelser & P. B. Baltes (Eds.) *International encyclopedia of social and behavioral sciences* (pp. 10068–10070). Amsterdam: Elsevier.
- Goldthorpe, J. H. (2001). Causation, statistics, and sociology. *European Sociological Review*, *17*, 1–20.
- Haan, M. N., Kaplan, G. A., & Syme, S. L. (1989). Socioeconomic status and health: Old observations and new thoughts. In D. S. Bunker, B. H. Gomby & B. H. Kehler (Eds.), *Pathways to health* (pp. 76–135). Menlo Park, CA: The Henry J. Kaiser Family Foundation.
- Haas, S. A. (2007). The long-term effects of poor childhood health: An assessment and application of retrospective reports. *Demography*, *44*, 113–135.
- Haas, S. A., Glymour, M. M., & Berkman, L. F. (2011). Childhood health and labor market inequality over the life course. *Journal of Health and Social Behavior*, *52*(3), 298–313.
- Headey, B., & Muffels, R. (2016). Towards a theory of medium term life satisfaction: Two-way causation partly explains persistent satisfaction or dissatisfaction. *Social Indicators Research*, *129*(2), 937–960.
- Heckman, J. J. (1981). The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process. In C. F. Manski & D. L. McFadden (Eds.), *Structural analysis of discrete data and econometric applications*. Cambridge: MIT Press.
- Hertzman, C., Frank, J., & Evans, R. G. (1994). Heterogeneities in health status and the determinants of population health. In R. G. Evans, M. L. Barer, & T. R. Marmor (Eds.), *Why are some people healthy and others not? determinants of health and populations, social institutions and social change series* (pp. 67–92). New York: Aldine de Gruyter.
- Hoffmann, R. (2008). *Socioeconomic differences in old age mortality, the Springer series on demographic methods and population analysis* (Vol. 25). Dordrecht: Springer.
- Huisman, M., Kunst, A. E., & Mackenbach, J. P. (2003). Socioeconomic inequalities in morbidity among the elderly: a European overview. *Social Science and Medicine*, *57*, 861–873.
- Huurte, T., Rahkonen, O., Komulainen, E., & Aro, H. (2005). Socioeconomic status as a cause and consequence of psychosomatic symptoms from adolescence to adulthood. *Social Psychiatry and Psychiatric Epidemiology*, *40*(7), 580–587.
- Jylhä, M. (2009). What is self-rated health and why does it predict mortality? Towards a unified conceptual model. *Social Science and Medicine*, *69*(3), 307–316.
- Kröger, H. (2015a). The contribution of health selection to occupational status inequality in Germany—Differences by gender and between the public and private sectors. *Public Health*, *133*, 67–74.
- Kröger, H. (2015b). Newspell—Easy management of complex spell data. *Stata Journal*, *15*, 155–172.
- Kröger, H., Fritzell, J., & Hoffmann, R. (2016a). The association of levels of and decline in grip strength in old age with trajectories of life course occupational position. *PLoS ONE*, *11*(5), e0155954.

- Kröger, H., Hoffmann, R., & Pakpahan, E. (2016b). Consequences of measurement error for inference in cross-lagged panel design—The example of the reciprocal causal relationship between subjective health and SES. *Journal of the Royal Statistical Society. Series A*, *179*, 607–628.
- Kröger, H., Pakpahan, E., & Hoffmann, R. (2015). What causes health inequality? A systematic review on the relative importance of social causation and health selection. *European Journal of Public Health*, *25*(6), 951–960.
- Kuh, D., Ben-Shlomo, Y., Lynch, J., Hallqvist, J., & Power, C. (2003). Life course epidemiology. *Journal of Epidemiology and Community Health*, *57*, 778–783.
- Lahelma, E., Laaksonen, M., Martikainen, P., & Rahkonen, O. (2008). Die Mehrdimensionalität der sozioökonomischen Lage-Konsequenzen für die Analysen gesundheitlicher Ungleichheit. In U. Bauer, U. H. Bittlingmayer, & M. Richter (Eds.), *Health Inequalities—Erklärungsansätze gesundheitlicher Ungleichheit* (pp. 143–166). Wiesbaden: VS-Verlag.
- Lahelma, E., Laaksonen, M., Martikainen, P., Rahkonen, O., & Sarlio-Lähteenkorva, S. (2006). Multiple measures of socioeconomic circumstances and common mental disorders. *Social Science and Medicine*, *63*(5), 1383–1399.
- Laurijssen, I., & Spruyt, B. (2015). Not for people like us? A 6-year panel study of the mutual relationship between feelings of relative deprivation and occupational status among young adults in Flanders. *Social Indicators Research*, *124*(2), 617–635.
- Link, B. G., & Phelan, J. (1995). Social conditions as fundamental causes of disease. *Journal of Health and Social Behavior*, *35*(extra issue), 80–94.
- Lundberg, O. (1991). Childhood living conditions, health status, and social mobility: A contribution to the health selection debate. *European Sociological Review*, *7*, 149–162.
- Lynch, J., & Kaplan, G. (2000). Socioeconomic position. In L. F. Berkman & I. Kawachi (Eds.), *Social epidemiology* (pp. 13–35). Oxford: Oxford University Press.
- Mackenbach, J. P. (2006). *Health inequalities: Europe in profile*. London: UK presidency of the EU.
- Mackenbach, J. P., Kulhánová, I., Bopp, M., Deboosere, P., Eikemo, T. A., Hoffmann, R., et al. (2015a). Variations in the relation between education and cause-specific mortality in 19 European populations: A test of the “fundamental causes” theory of social inequalities in health. *Social Science and Medicine*, *127*, 51–62.
- Mackenbach, J. P., Kulhánová, I., Menvielle, G., Bopp, M., Borrell, C., Costa, G., et al. (2015b). Trends in inequalities in premature mortality: A study of 3.2 million deaths in 13 European countries. *Journal of Epidemiology and Community Health*, *69*(3), 207–217.
- Mackenbach, J. P., Meerding, W. J., & Kunst, A. E. (2011). Economic costs of health inequalities in the European Union. *Journal of Epidemiology and Community Health*, *65*(5), 412–419.
- Manor, O., Matthews, S., & Power, C. (2003). Health selection: the role of inter- and intra-generational mobility on social inequalities in health. *Social Science and Medicine*, *57*(11), 2217–2227.
- Marks, G. N. (2011). Issues in the conceptualisation and measurement of socioeconomic background: Do different measures generate different conclusions? *Social Indicators Research*, *104*(2), 225–251.
- Marmot, M., Ryff, C. D., Bumpass, L. L., Shipley, M., & Marks, N. F. (1997). Social inequalities in health: Next questions and converging evidence. *Social Science and Medicine*, *44*(6), 901–910.
- Martikainen, P., Valkonen, T., & Moustgaard, H. (2009). The effects of individual taxable income, household taxable income, and household disposable income on mortality in Finland, 1998–2004. *Population Studies*, *63*(2), 147–162.
- Mayer, K. U. (2009). New directions in life course research. *Annual Review of Sociology*, *35*, 413–433.
- Mazzonna, F., & Havari, E. (2011). *Can we trust older people's statements on their childhood circumstances? Evidence from SHARELIFE*. Rochester, NY: Social Science Research Network.
- Miething, A., Lundberg, O., & Geyer, S. (2013). Income and health in different welfare contexts: A comparison of Sweden, East and West Germany. *Scandinavian Journal of Public Health*, *41*(3), 260–268.
- Mirowsky, J., & Ross, C. E. (2003). *Education, social status, and health*. New York: De Gruyter.
- Mulatu, M. S., & Schooler, C. (2002). Causal connections between socio-economic status and health: Reciprocal effects and mediating mechanisms. *Journal of Health and Social Behavior*, *43*, 22–41.
- Muthen, L. K., & Muthen, B. O. (2015). *Mplus User's Guide* (7th ed.). Los Angeles: Muthén & Muthén.
- Pakpahan, E., Hoffmann, R., & Kröger, H. (2017). The long arm of childhood circumstances on health in old age: Evidence from SHARELIFE. *Advances in Life Course Research*, *31*, 1–10.
- Pakpahan, E., Kröger, H., & Hoffmann, R. (2015). Statistical methods for causal analysis in life course research: An illustration of a cross lagged structural equation model and a latent growth model. *International Journal of Social Research Methodology*, *20*(1), 1–19.
- Palloni, A., Milesi, C., White, R. G., & Turner, A. (2009). Early childhood health, reproduction of economic inequalities and the persistence of health and mortality differentials. *Social Science and Medicine*, *68*, 1574–1582.

- Phelan, J. C., Link, B. G., Diez-Roux, A., Kawachi, I., & Levin, B. (2004). "Fundamental Causes" of Social inequalities in mortality: A test of the theory. *Journal of Health and Social Behavior*, 45(3), 265–285.
- Power, C., Manor, O., & Li, L. (2002). Are inequalities in height underestimated by adult social position? Effects of changing social structure and height selection in a cohort study. *BMJ*, 325(7356), 131–134.
- Robeyns, I. (2011). Capability approach. In J. Pell & I. V. Staveren (Eds.), *Handbook of economics and ethics* (pp. 39–46). Cheltenham: Edward Elgar.
- Rubin, D. B. (2005). Causal inference using potential outcomes. *Journal of the American Statistical Association*, 100, 322–331.
- Schrijvers, C. T. M., Stronks, K., van de Mheen, D., & Mackenbach, J. P. (1999). Explaining educational differences in mortality: the role of behavioral and material factors. *American Journal of Public Health*, 89(4), 535–540.
- Schröder, M. (2011). Concepts and Topcis. In M. Schröder (Ed.), *Retrospective data collection in the Survey of Health, Ageing and Retirement in Europe. SHARELIFE Methodology* (pp. 11–19). Mannheim: MEA
- Schwarz, N. (2007). Retrospective and concurrent self-reports: The rationale for real-time data capture. In A. Stone, S. S. Shiffman, & A. Atienza (Eds.), *The science of real-time data capture: Self-reports in health research* (pp. 11–26). New York: Oxford University Press.
- Sen, A. (1999). Health in development. *Bulletin of the World Health Organization*, 77(8), 619.
- Siegrist, J., & Wahrendorf, M. (2016). *Work stress and health in a globalized economy. The model of effort-reward imbalance*. Berlin: Springer.
- Smith, J. P. (1998). Socioeconomic status and health. *The American Economic Review*, 88(2), 192–196.
- Smith, J. P. (1999). Healthy bodies and thick wallets: The dual relation between health and economic status. *Journal of Economic Perspectives*, 13(2), 145–166.
- Smith, J. P. (2003). *Consequences and predictors of new health events*. NBER Working Paper 10063, National Bureau of Economic Research.
- Smith, J. P. (2004). Unraveling the SES-health connection. *Population and Development Review*, 30(supplement(Aging, health, and public policy: demographic and economic perspectives)), 108–132.
- Smith, J. P., & Thomas, D. (2003). Remembrances of things past: test-retest reliability of retrospective migration histories. *Journal of the Royal Statistical Society Series A*, 166, 23–49.
- Stansfeld, S. A., Clark, C., Rodgers, B., Caldwell, T., & Power, C. (2011). Repeated exposure to socioeconomic disadvantage and health selection as life course pathways to mid-life depressive and anxiety disorders. *Social Psychiatry and Psychiatric Epidemiology*, 46(7), 549–558.
- Stowasser, T., Heiss, F., McFadden, D., & Winter, J. (2011). "Healthy, wealthy and wise?" Revisited: An analysis of the causal pathways from socioeconomic status to health. In A. W. David (Ed.), *Investigations in the economics of aging* (pp. 267–317). Chicago: University of Chicago Press.
- Syddall, H. E., Westbury, L. D., Dodds, R., Dennison, E., Cooper, C., & Sayer, A. A. (2017). Mortality in the Hertfordshire Ageing Study: Association with level and loss of hand grip strength in later life. *Age and Ageing*, 46(3), 407–412.
- Tarkiainen, L., Martikainen, P., Laaksonen, M., & Valkonen, T. (2012). Trends in life expectancy by income from 1988 to 2007: Decomposition by age and cause of death. *Journal of Epidemiology and Community Health*, 66(7), 573–578.
- Torssander, J., & Erikson, R. (2010). Stratification and mortality—A comparison of education, class, status, and income. *European Sociological Review*, 26(4), 465–474.
- Warren, J. R. (2009). Socioeconomic status and health across the life course: A test of the social causation and health selection hypotheses. *Social Forces*, 87, 2125–2153.
- Warren, J. R., & Kuo, H.-H. (2003). How to measure "What People do for a Living" in research on the socioeconomic correlates of health. *Annals of Epidemiology*, 13, 325–334.
- Weiss, C. T. (2012). *Two measures of lifetime resources for Europe using SHARELIFE*. SHARE Working Paper (Vol. 06-2012).
- West, P. (1991). Rethinking the health selection explanation for health inequalities. *Social Science and Medicine*, 32, 373–384.