Advances in Life Course Research 36 (2018) 23–36

Contents lists available at ScienceDirect

Advances in Life Course Research

journal homepage: www.elsevier.com/locate/alcr

Pathways between socioeconomic status and health: Does health selection or social causation dominate in Europe?

Rasmus Hoffmann[a,b,*], Hannes Kröger[c], Eduwin Pakpahan[d]

[a] European University Institute, Via Dei Roccettini 9, 50014 San Domencio di Fiesole, Italy
[b] Max Planck Institute for Demographic Research, Konrad-Zuse-Straße 1, 18057 Rostock, Germany
[c] Deutsches Institut für Wirtschaftsforschung e.V. (DIW Berlin), SOEP Mohrenstraße 58, 10117 Berlin, Germany
[d] Institute of Neuroscience, Newcastle University, Campus for Ageing and Vitality (CAV), Biomedical Research Building, Newcastle upon Tyne, NE4 5PL, UK

ARTICLE INFO

Keywords:
Socioeconomic status
Health inequality
Causality
Structural equation model
Social causation
Health selection

ABSTRACT

Health differences which correspond to socioeconomic status (SES) can be attributed to three causal mechanisms: SES affects health (social causation), health affects SES (health selection), and common background factors influence both SES and health (indirect selection). Using retrospective survey data from 10 European countries (SHARELIFE, n = 20,227) and structural equation models in a cross-lagged panel design, we determine the relative importance in terms of explanatory power of social causation and health selection in the life course from childhood to old age. Both SES and health heavily depend on their prior status, albeit more for SES than health. During the transition from childhood to working ages, social causation and health selection are equally weak. Turning to the second phase (transition from working ages to old age) causation increases while selection decreases which makes causation the dominant mechanism in older age. While the contribution of common background factors remains difficult to assess, this study shows that both social causation and health selection are responsible for health inequalities; however, their relative importance changes with age. Life course modelling can complement causal analysis by revealing interactions between the processes of SES and health and their contribution to health inequality.

1. Introduction

Social inequalities, i.e. the unequal distribution of education, material wealth or occupational status, can be measured at specific points in time, but a person's socioeconomic status (SES) develops and changes over the life course. These changes can be understood as a process with critical periods and transitions, with path dependencies and accumulations. Health inequalities – systematic differences in health or mortality between groups with different SES – can also be described at specific points in time. However, health is also a life course process that is determined by fixed individual characteristics, social influences, behaviour, and institutional settings; it also has critical periods, path dependencies, and accumulations of problems or recovery. We study how these two processes influence each other, whether pathways from SES to health have more explanatory power than pathways from health to SES, and whether this changes over the life course.

The question of the relative importance of social causation and health selection is part of an enduring discussion between different fields of social sciences, with several important normative and political implications with regard to the acceptability and the reduction of health inequalities. It also includes complex methodological questions concerning the empirical analysis of causal effects in a longitudinal life course framework, a framework which is ideal for explaining health inequalities (Kuh, Ben-Shlomo, Lynch, Hallqvist, & Power, 2003; Mayer, 2009). The explanation of health inequalities by common background factors will also be addressed in this study, but since these factors are largely unobservable, we can take them into account only partially. In the remainder of the introduction, we will first elaborate on the background of health inequalities, and then on the open question of the relative importance of social causation and health selection.

1.1. The mutual relationship between socioeconomic status and health

Morbidity and mortality rates are systematically higher among people with lower SES. 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, and they rate high on the political agenda (Elo, 2009). While average health and life expectancy have improved over time in almost all countries, relative health inequalities are also on the increase (Mackenbach et al., 2015).

* Corresponding author.
E-mail address: rasmus.hoffmann@eui.eu (R. Hoffmann).

https://doi.org/10.1016/j.alcr.2018.02.002
Received 13 April 2017; Received in revised form 28 November 2017; Accepted 20 February 2018
Available online 24 February 2018

The economic costs of health inequalities in the EU have been estimated at 1000 billion € per year, which is 9.5% of EU GDP (Mackenbach, Meerding, & Kunst, 2011).

The mechanisms that create health inequalities are manifold and have been discussed extensively in the literature (Case & Deaton, 2005; Galama & van Kippersluis, 2010; Hoffmann, 2008). A thorough investigation reveals mechanisms where SES influences health, and those where health affects SES. Throughout this text, we will refer to the first as *social causation* and to the second as *health selection*. The first of these models implies, for example, that education influences health through health awareness and risk behaviours, that occupational status influences health through prestige and occupational hazards, and that income and wealth influence health through the affordability of health care, environmental hazards, consumption, and the psychological burden of being poor (Gathmann, Jürges, & Reinhold, 2015; Graham, 2009; Hummer, Rogers, & Eberstein, 1998). Education may also indirectly influence health through its effect on social mobility (Altzinger, Crespo Cuaresma, Rumplmaier, Sauer, & Schneebaum, 2015; Erikson & Goldthorpe, 1992). The second model (health selection) entails health influencing education (during childhood) through the general ability to invest in education, or through specific factors such as mental health conditions (Thorley, 2016), health influencing occupational status through the ability to invest in a career, and health influencing income and wealth, again through the ability to invest in a career and medical care expenditures (Galama & van Kippersluis, 2010; Palloni, White, & Milesi, 2008; West, 1991).

A third 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 behaviour and lifestyle. This third model is most difficult to test empirically. Hence, in our study, we concentrate on the relative explanatory power of social causation versus health selection in generating health inequalities over the life course; however, we also discuss the extent to which we can draw conclusions about indirect selection.

In related disciplines such as health sociology and health economics, there is disagreement concerning the relative importance of social causation versus health 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 to divergent concepts of causality. One argument against health selection addresses the timing of cause and effect over the life course: 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 labour 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 higher 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 such arguments against health selection and related findings seem persuasive, there are also studies that reveal health selection, especially at older working ages, which has been explained by a higher prevalence of poor health, thus affecting the ability to work (Smith, 1998; Smith, 2003). More generally, it is often neglected that reverse causality from health to SES can bias the coefficients of conventional statistical models if the direction of causality is simply assumed (Hertzman, Frank, & Evans, 1994).

Few studies have examined the possibility of health selection. While some authors do not consider it very important (e.g. Chandola, Bartley, Sacker, Jenkinson, & Marmot, 2003; Lundberg, 1991; Manor, Matthews, & Power, 2003), others think that the influence of health on material status is the strongest overall causality in the relationship between SES and health (Galama & van Kippersluis, 2010). A recent systematic literature review on the relative importance of social causation versus health selection evaluated 34 out of 2952 reviewed studies from the past 20 years (Kröger, Pakpahan, & Hoffmann, 2015). The result is that, across disciplinary boundaries, there is no preference for one causal direction. 12 studies supported causation, and 10 supported selection – the other studies supported both directions equally.

This study uses several SES indicators and covers near-complete life courses from childhood to old age, in order to estimate the relative explanatory power of social causation and health selection. Unlike many existing studies, we do not aim to identify local causal effects between specific aspects of SES and a specific measure of health. Instead, we follow a long-term life course approach with broad indicators of SES and health, studying the extent to which variation in SES can be explained by variation in health at different stages in the life course, and vice versa. Separating our analysis into the transition from childhood to adulthood and the transition from adulthood to old age, we contribute to a better understanding how the reciprocal relationship between SES and health changes over the life course to produce inequality, because 'relying on an *a priori* definition of the direction of causality contradicts the recognition that social causation and health selection processes co-

evolve across life stages' (Lee & Jackson, 2017). It also allows us to test the hypotheses from the literature mentioned above, according to which health selection should either be relatively influential in early adulthood because of health-related social mobility at labour market entry, or relatively influential in older working ages when health problems start to become more prevalent. For social causation, based on existing knowledge, we do not expect a particular trend across the life course.

2. Data and methods

2.1. Data

We use the third wave (SHARELIFE, version 5.0.0) of the Survey of Health Aging and Retirement in Europe (SHARE), in which persons aged 50 and older were asked about their current circumstances and, retrospectively, about the development of their SES and their health since childhood. The data is representative for the population 50+ and their spouses living in households in the respective European countries. SHARELIFE contains detailed information on events and changes in health and SES (Börsch-Supan et al., 2013). The data was collected with personal interviews at home using computerised questionnaires. We limit our analysis to 10 countries (Austria, Belgium, Denmark, France, Germany, Italy, Netherlands, Spain, Sweden and Switzerland), because for three SHARE countries information on wages was not comparable over the life course (Poland, Czech Republic) or contained too many missing values (Greece). We also limit the analyses to persons aged 50 to 90 at the time of the interview in 2008/2009 (n = 20,227). The average response rate across countries is about 60% (ranging from about 40% to 80%) and details of participating countries are provided online (http://www.share-project.org/data-access-documentation/sample.html). For a description of the data see Table 1.

In the operationalization of our concepts we divide the life course into three periods: childhood (up to age 15), principal working age (between 30 and 49) and old age (50 to 90). We use three indicators for SES in childhood at age 10: the number of books in the household, the number of rooms per person, and the occupational status of the father within four categories of the International Standard Classification of Occupations (ISCO). An individual's education was measured as number of years spent in education. For working age, we use two social indicators: occupational status (ISCO) and estimates of average monthly wages over 20 years. SHARE only provides a 1-digit ISCO code, which we converted into four skill levels as suggested by the International Labour Organization (ILO 2012): level 1: elementary occupations (ISCO code 9), level 2: ISCO codes 4–8, level 3: ISCO code 3, and level 4: ISCO code 1 and 2. For individuals in the armed forces (< 1%) the

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

Latent Construct	Variable		Category	N = 20227	
	Country	West	Austria	976	4.83%
			Belgium	2785	13.77%
			France	2433	12.03%
			Germany	1885	9.32%
			Netherlands	2222	10.99%
			Switzerland	1272	6.29%
		South	Italy	2483	12.28%
			Spain	2187	10.81%
		North	Denmark	2076	10.26%
			Sweden	1908	9.43%
	Age in 2009 (Wave 3)		Mean	67.17	
			Min	50	
			Max	90	
	Gender		Male	9096	44.97%
			Female	11131	55.03%
	Education (number of years)		Mean	10.75	
			Min	0	
			Max	25	
			Missing	2533	
CSES (Childhood SES – Age 10)	Number of books		0–10 books	8521	42.68%
			11–25 books	4291	21.49%
			26–100 books	4316	21.62%
			101–200 books	1395	6.99%
			> 200 books	1441	7.22%
			Missing	263	
	Father's occupation		ISCO Level 1 (elementary occupation)	3919	20.58%
			ISCO Level 2	12366	64.92%
			ISCO Level 3	922	4.84%
			ISCO Level 4 (manager)	1840	9.66%
			Missing	1180	
	Rooms per capita		Mean	0.78	
			Min	0	
			Max	10	
			Missing	348	
CHEALTH (Childhood health – Age 15)	Self-rated health		Poor	496	2.48%
			Fair	1329	6.64%
			Good	5095	25.44%
			Very good	6291	31.41%
			Excellent	6817	34.04%
			Missing	199	
	Missed school		Yes	2342	11.66%
			No	17750	88.34%
			Missing	135	
	Hospitalised		Yes	1251	6.22%
			No	18871	93.78%
			Missing	105	
ASES (Adult SES – Age 30–50)	Occupation ISCO–Classification		ISCO Level 1 (elementary occupation)	2823	16.97%
			ISCO Level 2	9134	54.91%
			ISCO Level 3	1811	10.89%
			ISCO Level 4 (manager)	2867	17.23%
			Missing	3592	
	Average wages		Mean	1284.03	
			Standard deviation	856.90	
			Min	7.49	
			Max	6125.63	
			Missing	10297	
AHEALTH (Adult health – Age 30–50)	Percentage of years of non-illness		Mean	0.97	
	Percentage of years of non-poor health		Mean	0.97	
	Percentage of years of non-stress		Mean	0.90	

(*continued on next page*)

Table 1 (*continued*)

Latent Construct	Variable	Category	N = 20227	
OSES (Old SES – Age 50+)	Household income	Mean	30064.13	
		Standard deviation	47801.83	
		Min	0	
		Max	755089.4	
		Missing	2634	
	Household wealth	Mean	159793.30	
		Standard deviation	220935.80	
		Min	−784644.10	
		Max	7153102.00	
		Missing	718	
OHEALTH (Old health – Age 50+)	Self-rated health	Poor	2420	12.01%
		Fair	5408	26.85%
		Good	7465	37.06%
		Very good	3222	15.99%
		Excellent	1629	8.09%
		Missing	83	
	Grip strengh	Mean	33.87	
		Standard deviation	12.24	
		Min	1	
		Max	85	
		Missing	1652	

information for occupational status is taken from other job spells. Wages are calculated as the average of all job spells between age 30 and age 50. For these spells the employment income after tax is recorded in the data, and we correct it for purchasing power and inflation by purchasing power parities (PPP) relative to German Euros in 2006 (Weiss, 2012). Because many persons are retired by the time of the interview, we measure SES at higher ages with the net-equivalent income and household net wealth per capita, both at the time of the interview (ages 50–90). Wealth includes property, cars, company shares, and liquid funds, minus debts.

The different measurements of SES reflect that, at different stages of the life course, different resources are relevant for an individual's socioeconomic status in society: first one's parents' and then one's own occupation, and ultimately income and wealth that have been acquired over the life course. The use of different indicators for SES at different ages also implies that they are not directly comparable. This is why we use several variables at each age as indicators for a latent variable for SES, which makes the concept of SES one-dimensional in principle – albeit measured by several indicators – but improves comparability over the life course.

Evidence on the relative importance of causation and selection will always be contingent on the social context, the method, and the indicators used (Huurre, Rahkonen, Komulainen, & Aro, 2005). These indicators cannot be assessed on a simple gradient of validity, but particular dimensions of SES are related via specific mechanisms to certain aspects of health. This complexity constitutes a theoretical and empirical problem that cannot be solved here. Our study is a step towards a critical reflection, providing some preliminary answers to a complex empirical question. To this end, we used multiple and general indicators of SES: The simplification involved in our measurement of SES and health incurs the risk of missing specific dimension-specific causal mechanisms, although we are able with this approach to cover large parts of the life course and many aspects of SES and health. Our results should be interpreted as summary measures and net effects of all factors involved in the relationship between SES and health. Since the SES measures for childhood refer to the parents and the household in childhood, the correlations between childhood and adulthood are mainly *inter*generational, whereas the SES correlations between adulthood and old age and all correlations of health statuses are *intra*generational.

Health in childhood is measured by three indicators covering the age up to 15: self-assessed health in five categories, the question whether school was ever missed because of health for one month or more, and whether one month or more was spent in hospital. At ages 30–49, our health measure is based on three indicators reflecting the years in which individuals reported illness, bad health, or stress, respectively. Respondents could define the start and the end of several periods of illness, the start and the end of one period of poor health, and one period of stress. From these spells we calculated the percentages of years reported to be free of illness, poor health, and stress respectively.

In old age, health is measured with the indicators current self-rated health and grip strength. We thereby combine the advantages of a subjective and an objective health indicator. 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 behaviour) because our aim is to estimate the total effects between SES and health, and we expect other variables to mediate these effects.

2.2. Methods

We chose a model-based approach to causal analysis using life-long retrospective data in order to study the interplay between SES and health. This differs from a design-based approach, such as quasi-experiments, in that it focusses on causality within a model based on theoretical assumptions on how SES and health are related, while the design-based approach uses external variation, e.g. in a natural experiment, needing fewer assumptions. The advantage of a model-based approach is the potential for simultaneously modelling two related processes (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 taken into account and a quasi-random distribution of the stimulus (conditional independence) is achieved. The causal concept behind our approach is "causation as a generative process", introduced by Cox (1992), and since further developed in sociology (Blossfeld, 2009; Goldthorpe, 2010). It stresses the processual character of social phenomena and causal effects.

In social situations, there is often more than one mechanism

involved in the production of an outcome. Therefore, our aim is to use a dynamic model that includes as many relevant factors as possible, representing multiple mechanisms simultaneously in order to integrate all relevant processes between cause and effect. *Conditional independence* in such a model means that the two processes are independent, conditional on the history of the joint processes (Blossfeld, Golsch, & Rohwer, 2007).

The ability to measure and analyse processes over a long period of time is the main reason for our adoption of the life course perspective in health research here to study the interplay between SES and health (Blane, Netuveli, & Stone, 2007). We do not claim to reveal causal effects as they are understood in the potential outcome or counterfactual framework, e.g. what happens to the SES of an average individual after a change in health or vice versa (Rubin, 2005). Throughout the text we use the terms *effect* and *influence* in terms of explanatory power. Modelling two processes simultaneously, we can reveal the relative explanatory power of each pathway by comparing standardised coefficients.

We estimate the parameters of a structural equation model (Bollen, 1989; Pakpahan, Kröger, & Hoffmann, 2015) that represent the relative importance of causation and selection in different stages of the life course (see Fig. 1). We model SES and health at three different ages as latent variables. Our structural equation model estimates the paths' parameters between these latent variables, while education is treated as an observed variable. Another difference between education and the other dimensions of SES is that, logically, the former can no longer be influenced by health after young adulthood. This is why we do not include education in our measure of SES. To include a fixed element in the measure of SES would distort our comparison between causation and selection. At the same time, education is an important mediator between childhood conditions and later life. Both features of education are reflected in our model.

The parameters are estimated using the Full Information Maximum Likelihood method (FIML), which allows non-normally distributed variables and includes persons with item non-response in the analysis. We give standardised coefficients in a uniform value range of −1 to 1, in order to make them comparable across paths and models.

Our model estimates the correlation between SES and health in childhood that can be jointly influenced by common unobserved background factors, e.g. genetic factors or unobserved characteristics of the family. Consequently, we address the common background factors mentioned above to the extent that such factors create a correlation between health and SES in childhood. The actual path parameters can be divided in two groups: first, the autoregressive parameters showing the effect of SES at t_1 on SES at t_2 (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 working ages) can be subdivided into direct and indirect effects, the latter being mediated by education. Direct and indirect effects can be

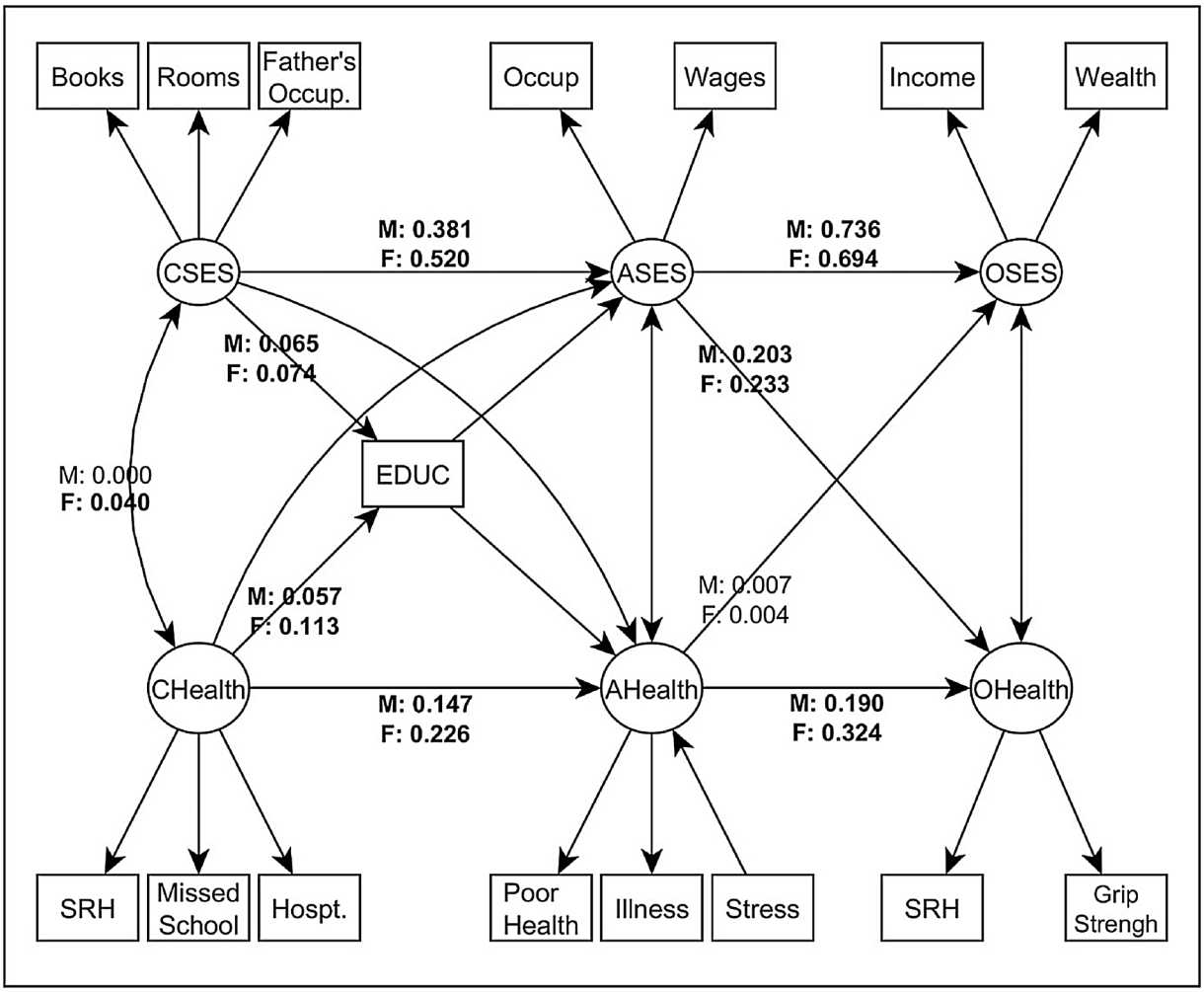

Fig. 1. Structural equation model for bi-directed relationships between SES and health over the life course, with standardised coefficients for all countries combined.
Note: a) countries = Austria, Germany, Netherlands, France, Switzerland, Belgium, Sweden, Denmark, Spain, Italy; C = childhood (0–15); A = adulthood (30–49); O = old age (50–90); M/F = men/women; b) bold numbers = statistically significant ($p < 0.05$, two-tailed); c) the cross-lagged effects between childhood and adulthood are the total effects (direct + indirect)

summed up to the total effect.

We calculated the models on three levels: first, one joint analysis of a pooled sample of all countries; second, countries grouped according to their region in Europe; third, each country separately (only shown in Appendix A). The regions roughly reflect different welfare state models according to a classification by Ferrera (1996), differentiating between Scandinavian (Denmark, Sweden), Southern (Italy, Spain) and Bismarckian welfare systems (Austria, Belgium, France, Germany, Netherlands and Switzerland). When the magnitude of health inequalities is compared between countries within a relatively homogenous cultural setting such as the EU or the USA, it is often assumed that international differences can be attributed to social or cultural factors within these countries, for example attitudes towards inequality or welfare state regulations (Eikemo, Huisman, Bambra, & Kunst, 2008; Hoffmann, 2011). In our analysis of the interplay between SES and health, we adopt a similar approach, assuming that observed differences between European regions can be attributed to human factors that, in principle, could be changed. We consider this to be an explorative approach that illustrates the heterogeneity within Europe, but we do not claim to test particular welfare-state specific hypotheses.

We study an aggregated dataset of 10 countries mainly because a complex structural equation model only produces meaningful results with a large sample. We acknowledge the remaining heterogeneity in our country groups, because national welfare systems are unique, complex systems that depend on historical settings and developments. We use country dummies in our European and regional models to control for these unobserved country differences. All models are calculated separately for men and women. Age at interview, also reflecting the birth cohort, is a control variable in the statistical models, but shows no significant effect on the results. All data preparation was done using Stata 14.1 and user-written software packages (Kröger, 2015b). All analyses are performed in Mplus 7.4 (Muthen & Muthen, 2015), with extraction of results done in R (Hallquist & Wiley, 2011).

3. Results

Results from the structural equation models are shown in Fig. 1 (as a graphical illustration of the model and with results for all countries together), in Table 2 (all coefficients and standard errors), and in Fig. 2 (only results that are relevant for our main question). Factor loadings for the six measurement models (for all countries combined) can be found in Appendix B.

As expected, the correlation between SES in childhood (CSES) and health in childhood (CHEALTH) is positive in most groups – higher SES is correlated with better health. However, the coefficient is only statistically significant for women in all countries combined (0.040). This low correlation exists not only between the latent variables, but also between the observed indicators for CSES and CHEALTH (not shown); the correlation increases during the life course. All autoregressive coefficients for SES in both phases of the life course are statistically significant, and range between 0.270 from CSES to ASES and 0.926 from ASES to OSES (both for men in Western countries). A value of 0.270 means that, for example, one standard deviation increase in CSES predicts a 0.270 standard deviation increase in ASES. Autoregressive parameters for health are less consistent and less often statistically significant. These results show that SES and health depend on their prior status, but this path dependency is stronger for SES than for health.

Our main question can be answered by comparing the coefficients that represent each of the two pathways in the model (causation and selection). In the first phase, this comparison is more complex, because the model estimates direct effects between SES and health and indirect effects that are mediated by education. The results for the indirect effects show strong effects between education and SES in both directions: For all countries combined, the effect of CSES on education is 0.508 and 0.501 for men and women respectively. The effect of education on ASES is 0.422 and 0.467, respectively. There are no statistically significant effects of CHEALTH on education, and only some minor effects of education on AHEALTH (for all countries combined, 0.079 and 0.088 for men and women respectively). This shows that education has some importance for social causation between childhood and adulthood as a mediator between SES and health. This small mediated effect is still stronger than the direct effects (all those not mediated by education). At the same time, education is not affected by CHEALTH, which means that education does not mediate health selection between childhood and adulthood. In other words, we cannot confirm previous findings that poor health in childhood affects adult SES through educational attainment (Haas, 2006).

The two rows 'Total' in Table 2 show the sum of the direct and indirect effect that is also displayed in Fig. 1 for all countries combined. As expected, all but one of 16 coefficients in these two rows are positive, which means that higher SES leads to better health and better health leads to higher SES. Out of 16 coefficients, 12 are statistically significant, and range between 0.057 and 0.140 (classifiable as relatively small effects) not showing systematic differences between countries or gender. To summarise phase 1 of the life course: both causation and selection are of equally minor importance, and selection seems to be slightly more important for women than for men.

In phase 2, as expected, the coefficients for causation are positive. They range between 0.066 and 0.233 and they are statistically significant. None of the 8 coefficients for selection are statistically significant. In addition, the coefficients for selection are much smaller than for causation. Comparing phase 1 and phase 2: in phase 1, causation and selection start with low values of equal size. When turning to phase 2, the amount of causation substantially increases, while selection decreases, which results in causation being more important than selection. In this regard, we consider the effects in phase 1 to be small, relative to the coefficients for causation in phase 2. This overall result can also be seen in Fig. 2, where we only give the relevant coefficients for causation and selection, including a direct test for their difference. In almost all regions, causation plays a more important role than selection, albeit only in phase 2. The *p*-values of these tests range from 0.000 to 0.110. Some inconsistencies between regions should be acknowledged: Firstly, men in Southern countries in phase 1 are the only example where selection is statistically significantly higher than causation; secondly, men and women in Northern countries in phase 2 show negative selection coefficients in phase 2. Besides these exceptions our general findings apply to all regions and both men and women.

4. Discussion

This study showed in a comprehensive life course perspective that, firstly, although SES and health in childhood are not strongly correlated, the correlation grows with increasing age; secondly, that SES and health during the life course depend substantially on their prior status; finally, that in the transition from childhood to working age, the social causation path was as important as the health selection path, while in the transition from working age to old age, causation was much more important than selection. This finding needs to be interpreted together with the finding that the autoregressive parameters for SES are much higher than those for health. If SES is more determined by its own past, it is consequently also less dependent on other influences, including health. The degree to which SES is self-dependent and stable, or conversely susceptible to health, depends on the stability of the society as a whole, with its social relations and institutions (Dannefer & Kelley-Moore, 2009). 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.

We did not find 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

Table 2
Results from structural equation models on the relationship between SES and health during the life course.

					Male				Female			
			Parameter		West	North	South	All	West	North	South	All
Correlation			CSES ↔ CHEALTH	Coef.	−0.016	0.022	0.016	0.000	0.028	0.076	0.034	**0.040**
				S.E.	0.025	0.044	0.040	0.020	0.027	0.043	0.036	0.020
			ASES ↔ AHEALTH	Coef.	0.112	0.095	−0.039	**0.213**	0.192	0.089	−0.168	0.101
				S.E.	0.062	0.068	0.088	0.071	0.108	0.078	0.108	0.055
			OSES ↔ OHEALTH	Coef.	**0.312**	**0.319**	0.071	**0.294**	**0.286**	0.067	0.331	**0.329**
				S.E.	0.091	0126	0.048	0.077	0.082	0.123	0.351	0.087
Phase 1	Autoregression		CSES → ASES	Coef.	**0.270**	**0.324**	**0.518**	**0.381**	**0.631**	**0.310**	**0.492**	**0.520**
				S.E.	0.031	0.041	0.068	0.270	0.039	0.044	0.076	0.026
			CHEALTH → AHEALTH	Coef.	**0.144**	**0.105**	**−0.162**	**0.147**	**0.223**	0.132	0.048	**0.226**
				S.E.	0.032	0.055	0.064	0.025	0.045	0.092	0.090	0.034
	Causation	Indirect 1	CSES → EDUC	Coef.	**0.490**	**0.468**	**0.622**	**0.508**	**0.487**	**0.517**	**0.597**	**0.501**
				S.E.	0.017	0.025	0.024	0.013	0.015	0.022	0.022	0.011
		Indirect 2	EDUC → AHEALTH	Coef.	**0.053**	**0.146**	**−0.058**	**0.079**	**0.069**	0.050	**0.088**	**0.088**
				S.E.	0.036	0.057	0.047	0.028	0.031	0.073	0.035	0.025
		Indirect 1 × Indirect 2	(CSES → EDUC) × (EDUC → AHEALTH)	Coef.	0.026	**0.069**	−0.036	0.040	**0.033**	0.026	**0.052**	**0.044**
				S.E.	0.018	0.027	0.029	0.014	0.015	0.038	0.021	0.013
		Direct	CSES → AHEALTH	Coef.	0.029	0.028	−0.015	0.025	0.040	0.063	0.018	0.030
				S.E.	0.036	0.051	0.046	0.029	0.034	0.041	0.038	0.028
		Total	CSES → AHEALTH	Coef.	0.054	**0.096**	−0.052	**0.065**	**0.073**	**0.089**	**0.071**	**0.074**
				S.E.	0.028	0.041	0.031	0.022	0.027	0.041	0.031	0.022
	Selection	Indirect 1	CHEALTH → EDUC	Coef.	0.018	0.021	0.058	0.009	0.005	0.004	−0.003	0.008
				S.E.	0.019	0.034	0.030	0.014	0.019	0.030	0.040	0.012
		Indirect 2	EDUC → ASES	Coef.	**0.355**	**0.505**	**0.454**	**0.422**	**0.392**	**0.578**	**0.466**	**0.467**
				S.E.	0.029	0.035	0.066	0.025	0.033	0.036	0.064	0.023
		Indirect 1 × Indirect 2	(CHEALTH → EDUC) × (EDUC → ASES)	Coef.	0.006	0.010	0.026	0.004	0.002	0.002	−0.001	0.004
				S.E.	0.007	0.017	0.014	0.006	0.007	0.017	0.018	0.006
		Direct	CHEALTH → ASES	Coef.	**0.019**	**0.087**	**0.114**	**0.053**	**0.041**	**0.099**	**0.115**	**0.109**
				S.E.	0.022	0.039	0.045	0.019	0.030	0.040	0.057	0.022
		Total	CHEALTH → ASES	Coef.	**0.026**	**0.097**	**0.140**	**0.057**	**0.043**	**0.101**	**0.113**	**0.113**
				S.E.	0.024	0.042	0.048	0.020	0.032	0.045	0.057	0.024
Phase 2	Autoregression		ASES → OSES	Coef.	**0.926**	**0.496**	**0.387**	**0.736**	**0.549**	**0.701**	**0.694**	**0.694**
				S.E.	0.095	0.060	0.040	0.066	0.059	0.101	0.245	0.054
			AHEALTH → OHEALTH	Coef.	**0.133**	**0.153**	**−0.171**	**0.190**	**0.270**	**0.050**	**0.178**	**0.324**
				S.E.	0.037	0.079	0.045	0.026	0.049	0.017	0.083	0.037
	Causation		ASES → OHEALTH	Coef.	**0.155**	**0.132**	**0.195**	**0.203**	**0.156**	**0.066**	**0.223**	**0.233**
				S.E.	0.038	0.047	0.029	0.026	0.023	0.017	0.038	0.021
	Selection		AHEALTH → OSES	Coef.	0.055	−0.131	−0.003	0.007	0.079	−0.145	0.058	0.004
				S.E.	0.046	0.128	0.037	0.043	0.042	0.118	0.058	0.042

Note: a) standardised regression coefficients; b) S.E. = standard errors; C = childhood; A = adulthood (30–49); O = old age (50–90); SES = socioeconomic status; West = Austria, Germany, Netherlands, France, Switzerland, Belgium; North = Sweden, Denmark; South = Spain, Italy; Phase 1 = transition from childhood to adulthood; Phase 2 = transition from adulthood to old age; c) 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; d) Statistically significant coefficients ($p < 0.05$, two-tailed) are printed in bold.

on the general level as we study it here. The few exceptions we identified may be due to random variation in the results, and within the scope of this study and with the available data, we are unable to determine factors that could explain the small differences.

Our finding that causation and selection have the same explanatory power in the first phase of the life course gives some support to previous findings that there is relatively high health selection in childhood (Van De Mheen, Stronks, & Mackenbach, 1998) and relatively high social mobility at labour market entry, where health influences occupation (James Smith, 1999). On the other hand, our finding that causation increases and selection decreases towards higher ages is not in line with previous evidence that selection is especially important in older working ages, where many health problems start to become more prevalent (Smith, 1998; Smith, 2003). Our results suggest instead that health selection is not important in old age because the SES of retired persons can no longer easily be influenced by health. For the increasing effect of social causation we can only speculate that this has to do with an increase in social inequality resulting from accumulative processes during the life course (Dannefer, 2003).

To compare our results to existing similar studies, we selected four studies that use similar age groups, indicators, and methods. The first study uses prospective data from the USA, but with an age range from 41 to 88, corresponding only to our second life phase. It shows that both causal directions are present – causation slightly more so than selection – with only small gender differences (Mulatu & Schooler, 2002). Another study with mostly prospective data from the USA in the age range 18 to 65 also shows that health in childhood has no effect on educational achievement. Interestingly, this study finds no evidence for selection (Warren, 2009). Third, Finnish register data in the age range 17 to 66 shows causation to be slightly more important than selection, although the indicator for health is limited to sickness absence from work (Aittomäki, Martikainen, Laaksonen, Lahelma, & Rahkonen, 2012). Fourth, Palloni, Milesi, White, & Turner (2009) use the 1958 British Cohort Study, which starts the observation at age 0 with birth weight and follows the sample from age 7 to age 42 with regular prospective interviews. Palloni et al. do not directly compare social

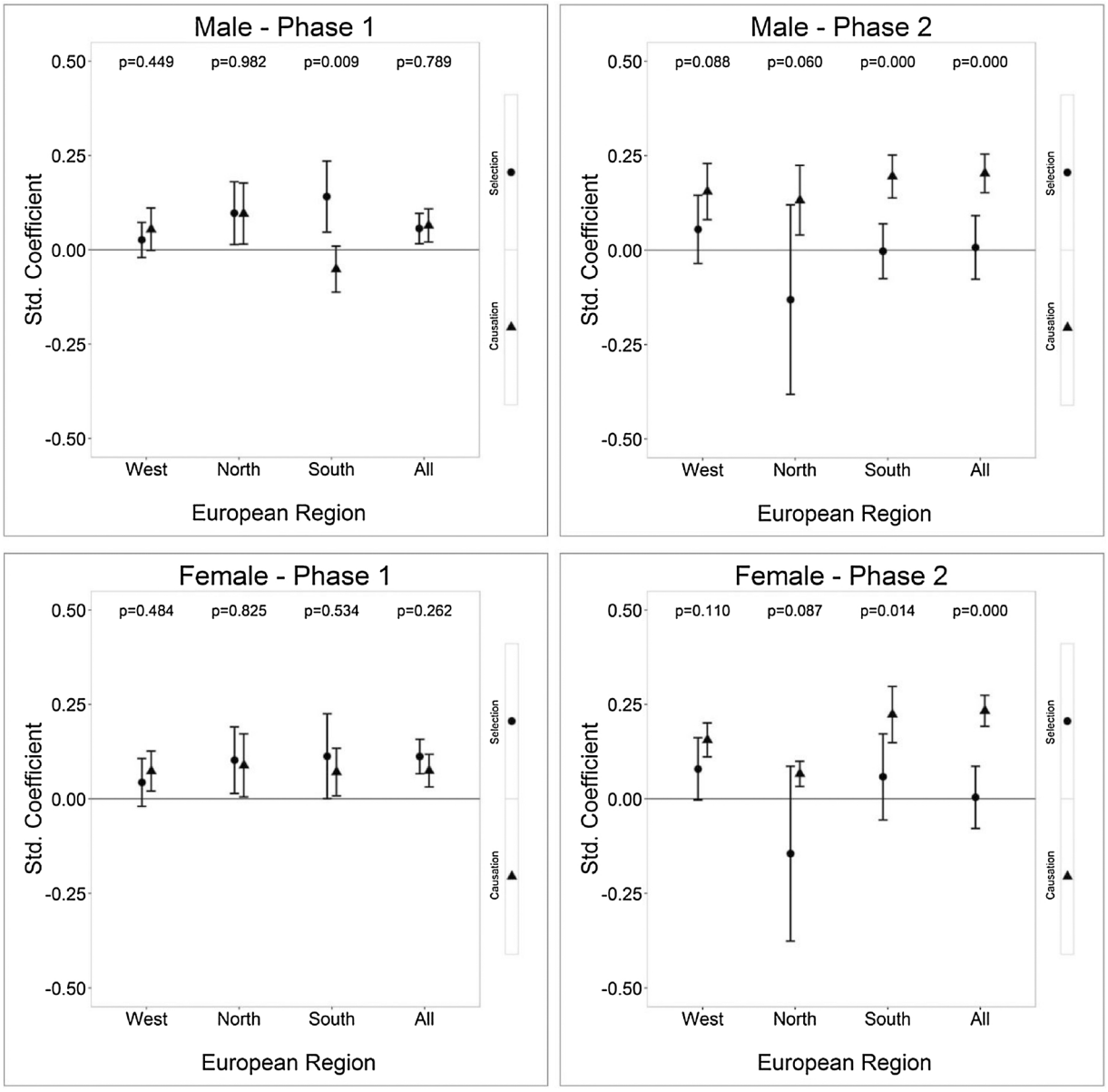

Fig. 2. Relative explanatory power of social causation and health selection, by life course phase, gender, and European region.
Note: a) West = Austria, Germany, Netherlands, France, Switzerland, Belgium; North = Sweden, Denmark; South = Spain, Italy; Phase 1 = transition from childhood to adult age; Phase 2 = transition from adulthood to old age; b) the confidence intervals show whether an estimate is different from zero ($p < 0.05$, two-tailed), while the *p*-values in the graph are from a direct Wald test for difference between the standardised coefficient for causation and selection.

causation and health selection, but they find evidence for the impact of childhood health on adult SES (via cognitive skills) and claim that this accounts for 10–12 percent of adult health inequality.

In the quantitative assessment of the relative importance of causation and selection, the third causal model of indirect selection also needs to be discussed. It assumes that SES and health are determined by common background factors, such as innate cognitive or physical characteristics (Goldman, 2001a), which are genetically determined and which can lead to the development of specific personalities and even concrete preferences in lifestyle (Fuchs, 1982). It is difficult to measure such common background factors and related mechanisms. Any existing variable that has been used in other studies (IQ, non-cognitive traits, school performance, birth weight, height etc.) comes with the problem that it may already depend on prior SES (of the parent). We interpret our result that health and SES in childhood are only weakly correlated – and that stronger correlation only appears with increasing age – as a hint that indirect selection is not of major importance. This differs from several studies that begin observing the relationship between SES and health in mid-life, find little or no mutual effects, and conclude that health inequalities can be explained by indirect selection. The recent study by Foverskov and Holm (2016) is an example, and the authors define as indirect selection everything that happens before age 30. In contrast, an important contribution of our study is that it starts measuring SES and health as early as possible in the life course, in order to attribute as much as possible of their interrelation to either causation or selection, instead of using indirect selection as a black-box or residual causal model that absorbs all interactions before observations began. However, our design cannot address the possibility that common background factors exist, but only affect health and SES later in life, hence not creating any correlation between health and SES in childhood.

Some authors describe indirect selection as being more important than health selection (e.g. Valkonen, 1996), some as being unimportant (Blane, Davey Smith, & Bartley, 1993; Goldman, 2001b). This causal

model is only rarely discussed and tested, perhaps because it is even more difficult to verify than the difference between social causation and health selection. However, the unknown importance of indirect selection has been discussed in relation to the social justice aspect of health inequalities (Fritzell, 2014). One perspective is that the greater the role of indirect selection, the less health inequalities could be considered unfair and thus a concern for social policy (Mackenbach, 2012).

The relative importance of social causation and health selection has implications for the normative assessment of health inequalities: The liberal or meritocratic claim that health selection is less unfair than social causation is based on the assumption that there is a significant amount of individual variation in health that is caused by biology alone and as such unrelated to SES. Instead, we would propose the welfare-state perspective that causation and selection are equally unfair, because both ill health as a consequence of low SES and low SES because of ill health indicate a dysfunction of the social security system that should, in principle, counteract both pathways.

4.1. Strengths and limitations

Our study combines a number of innovative strengths; first, we start early in the life course by measuring the early development of health and SES, gradually proceeding to old age. Second, we include several indicators that are important for a valid measurement of SES. Third, we combine these indicators into measurement models for latent variables, which reduces measurement error. Fourth, we use structural equation models that can simultaneously model two pathways (causation and selection), also taking into account indirect selection to some extent.

On the other hand, some limitations to our approach remain. The high percentages of good health in childhood and adulthood may bias downwards the coefficients for social causation, because − measured in this way − there is only little variance in health to be explained by SES. We concede that a more sensitive measure for adult health with more variance could have resulted in higher coefficients for social causation in Phase 1 if the majority of health differences in this phase are between the very healthy and those who are slightly less healthy. On the other hand, those who are really ill can already be identified through our approach. In this view, the low prevalence of bad health in adulthood is realistic and suggests that differences in childhood SES did not influence health in adulthood to a large extent.

We use retrospective data that, in principle, might be affected by recall bias (Smith & Thomas, 2003; Van De Mheen, Stronks, Looman, & Mackenbach, 1998). Measurement error in childhood variables can reduce the associations between childhood and later life outcomes. However, several studies have shown that the retrospective measurement of health and SES is relatively valid: Haas (2007) has validated retrospective health information from childhood, Garrouste and Paccagnella (2011) and Havari & Mazzonna (2015) provide validation studies of SHARELIFE data, and Kapteyn et al. (2007) compare Swedish SHARELIFE data to administrative data. In addition, latent variables reduce measurement error by using several indicators for a latent concept, where more reliable (objective) indicators complement less reliable ones. The remaining disadvantages of this retrospective data need to be balanced with the fact that it allows the study of longer periods than in previous research based on prospective data (Adams, Hurd, McFadden, Merrill, & Ribeiro, 2003; Stowasser, Heiss, McFadden, & Winter, 2011). As mentioned above, for the study of the relative importance of causation, selection, and indirect selection, it is especially important to commence measurement in childhood as close as possible to the starting point of these causal mechanisms (Heckman, 1981). In this regard, it is noteworthy that conditions in utero, perinatal circumstances, and very early childhood are not covered in our data, although their long-term influence on later life outcomes have been established (Doblhammer, 2004). Likewise, we miss later adolescent and early adulthood years, in which health conditions might affect educational attainment.

4.2. Sensitivity analyses

Our study of the relative importance of social causation and health selection and its change over age is based on a separation of the life course into three broad age groups and two transitions between them, because the amount of data and specific information (and also the needs and the complexity of our statistical model) do not allow for more and smaller age groups. To start the higher age group at age 50+ was a pragmatic choice to make optimal use of the SHARELIFE dataset, and we acknowledge the risk of ignoring interactions between SES and health that happen within this age group, for example at specific older working ages. In order to test whether our results for the wide age group 50–90 differ between people who are still working and retired people, we separated these two groups and present the findings in Appendix C. Unfortunately, the results do not allow further insights and a valid subgroup comparison: In both groups, social causation seems to be much stronger than health selection, but for the small group of people who are still working, the confidence intervals are large and widely overlapping, which indicates instability in this particular structural equation model. We also tested whether our results are sensitive to the starting age of the highest age group, which, among other things, also influences the share of retired persons in the oldest age group. The results in Appendix C reveal no substantial deviation from the original results, which suggests that the age range 50–60 is not substantially different from 60+ in the broad terms of the relative importance of causation and selection.

We explored whether our findings are sensitive to the inclusion of a direct path from childhood to old age, and we found that these paths were not statistically significant and did not change the relative magnitude of social causation and health selection (results not shown). This confirms existing evidence that most of the effects of childhood on old age ('long arm of childhood') are moderated by SES and health in adulthood (Pakpahan, Hoffmann, & Kröger, 2017). Likewise, it can be argued that education not only affects SES and health in adulthood, but also has a long-lasting direct effect on health in old age, because it shapes behavioural patterns and decision-making (Mirowsky & Ross, 2003). We addressed this possibility by adding a direct path from education to old age health, which surprisingly turned out to be negative. Most other coefficients remained the same, except for the effect of adult SES on old age health, which was even higher than before (see Appendix D). We can only speculate that the small but negative effect of education on health in old age occurs because of multicollinearity, due to the fact that occupation and several material status variables, which are normal consequences of education, are correlated. An alternative explanation for the negative effects of SES variables on health is mortality selection, i.e. the fact that from the population subgroup with low SES and poorer health only a few very robust persons survive until old age. This group would be positively selected and may reverse the social gradient. To address this potential explanation, we performed a sensitivity analysis, excluding those above age 75, corresponding to the oldest 20% of the sample (Appendix D). This actually increases the negative direct effect, but the main results concerning social causation and health selection stay the same. While both multicollinearity and health selection may play a role, the latter seems to be less important. Nevertheless, we do not consider the influence of multicollinearity as a bias of our main conclusion regarding the relative importance of causation and selection, because this remained unchanged in all analyses. Modelling an additional direct path from education to old age health yields unreliable results, because the variance of old age health is mostly explained by adult health and adult SES. The total effect of education on health in old age is still positive, and we conclude that the simpler original model is sufficient to compare social causation and health selection, because the positive impact of education on health seems to work mainly via adult health, and not directly. Mortality selection can, in principle, also be responsible for our finding that health and SES in childhood are hardly correlated. However, it is also the case

here that the sensitivity analyses excluding the oldest people do not lead to higher correlations.

We only study the surviving population that might be selected, but we assume that, while selective mortality decreases health inequality in the surviving population – e.g. by poor, unhealthy people dying first – it does not systematically bias the comparison between social causation and health selection. Besides mortality selection, a general health-related participation bias might also apply to the SHARE sample in the sense that less healthy people and people suffering from a mental disorder are also less likely to participate.

Our multiple indicators for SES and health show differences in the percentage of missing values. The highest percentage of missing values in our analysis (50.9 percent) is for the variable 'wages'. The other indicators for adult SES have lower missing values (occupation 17.8 percent, household income 13.0 percent, household wealth 3.5 percent). We performed a sensitivity analysis for wages by excluding the 50.9 percent of the sample that did not answer this question, and found that it does not substantially change the results. Since many of the missing values are from 'housewives', we performed another sensitivity analysis excluding 1241 women and 51 men who reported, for at least 75 percent of the years between ages 30 and 49, mainly having worked in the household; this, too, did not change the results (see Appendix D). This is probably because we always use several indicators to define a latent variable and we use the Full Information Maximum Likelihood method that is the preferable estimation method for dealing with values missing at random.

In conclusion, both social causation and health selection play a part in the creation of health inequalities over the life course. In the second part of the life course, causation is more important than selection. In spite of several mostly data-related limitations, warranting a cautious interpretation of our findings as causal estimates, we contribute to a complex and important debate on two different causal directions by assessing their relative contribution to health inequality. In doing so, we combine a causal question with a perspective on long term life course processes. If both mechanisms contribute to health inequalities, both could also be used as entry points for social policy to reduce health inequalities, in different ways at different life stages.

Funding

This work was supported by a Starting Grant from the European Research Council [grant number 313532]

Acknowledgements

This paper uses data from SHARE wave 4 release 1.1.1, as of March 28th 2013 (DOI: 10.6103/SHARE.w4.111) or SHARE waves 1 and 2 release 2.6.0, as of November 29th 2013 (DOIs: 10.6103/SHARE.w1.260 and 10.6103/SHARE.w2.260) or SHARELIFE release 1.0.0, as of November 24th 2010 (DOI: 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 for a full list of funding institutions).

Appendix A. Country-specific results from structural equation models on the relationship between SES and health during the life course

Country	Gender	CSES → AHEALTH		CHEALTH → ASES		ASES → OHEALTH		AHEALTH → OSES	
		Coef	*S.E*	*Coef*	*S.E*	*Coef*	*S.E*	*Coef*	*S.E*
Austria	Male	0.039	0.072	0.090	0.114	**0.348**	0.086	0.095	0.103
	Female	−0.066	0.114	0.063	0.075	**0.396**	0.109	**−0.422**	0.117
Belgium	Male	0.082	0.075	−0.072	0.061	0.011	0.045	0.118	0.167
	Female	−0.137	0.071	0.048	0.056	**0.135**	0.058	**−0.451**	0.212
France	Male	−0.058	0.140	−0.018	0.026	**0.294**	0.047	0.101	0.321
	Female	**0.333**	0.170	−0.047	0.065	−0.185	0.549	−0.056	0.316
Germany	Male	0.051	0.049	0.024	0.071	**0.319**	0.059	0.006	0.060
	Female	0.015	0.078	0.012	0.050	**0.252**	0.047	−0.052	0.114
Netherlands	Male	**0.141**	0.064	0.086	0.066	**0.219**	0.056	0.154	0.125
	Female	−0.070	0.062	0.117	0.063	0.117	0.061	0.112	0.097
Switzerland	Male	−0.212	0.166	**0.209**	0.079	0.032	0.084	**−0.359**	0.164
	Female	0.097	0.071	−0.017	0.076	0.112	0.071	0.028	0.106
Denmark	Male	−0.051	0.069	**0.117**	0.057	**0.214**	0.042	**0.186**	0.045
	Female	−0.098	0.064	0.097	0.063	**0.188**	0.050	**0.184**	0.070
Sweden	Male	0.096	0.065	0.081	0.064	0.014	0.049	**−0.554**	0.220
	Female	0.043	0.080	**0.200**	0.061	**0.327**	0.079	0.054	0.108
Italy	Male	**0.239**	0.120	**0.143**	0.070	**0.204**	0.077	−0.207	0.216
	Female	0.035	0.052	0.082	0.076	**0.227**	0.042	−0.057	0.118
Spain	Male	0.002	0.039	0.104	0.075	**0.242**	0.044	−0.012	0.071
	Female	**0.118**	0.055	**0.179**	0.073	**0.364**	0.049	0.057	0.118

Appendix B. Factor loading of six measurement models (all countries combined)

Latent Variable	Indicators	Male			Female		
		Coef.	*S.E.*	*p-value*	*Coef.*	*S.E.*	*p-value*
CSES	Number of books	0.808	0.009	0.000	0.769	0.009	0.000
	Rooms	0.417	0.014	0.000	0.450	0.015	0.000
	Father's occupation	0.564	0.010	0.000	0.553	0.010	0.000
CHEALTH	Self-rated health	0.504	0.022	0.000	0.538	0.023	0.000
	Missed school	0.880	0.022	0.000	0.821	0.022	0.000
	Hospitalised	0.833	0.018	0.000	0.826	0.017	0.000
ASES	Occupation	0.494	0.017	0.000	0.638	0.016	0.000
	Average wages	0.666	0.018	0.000	0.443	0.024	0.000
AHEALTH	Percentage of illness	0.517	0.042	0.000	0.463	0.035	0.000
	Percentage of poor health	0.480	0.041	0.000	0.445	0.034	0.000
OSES	Wealth	0.391	0.041	0.000	0.404	0.035	0.000
	Income	0.263	0.024	0.000	0.226	0.022	0.000
OHEALTH	Grip strength	0.698	0.015	0.000	0.641	0.013	0.000
	Self-rated health	0.412	0.012	0.000	0.518	0.012	0.000

Appendix C. Two sensitivity analyses comparable to the main results in Table 2 of the main article:

1. Splitting the sample into retired and working people at the age of the interview (50+).
2. Reducing the sample to people aged 60+ at interview (n = 14,020).

					Retired		Still working		60+	
			Parameter		Male	Female	Male	Female	Male	Female
Correlation			CSES ↔ CHEALTH	Coef.	−0.015	0.036	−0.033	0.125	0.008	0.017
				S.E.	0.032	0.043	0.068	0.078	0.031	0.041
			ASES ↔ AHEALTH	Coef.	0.212	0.045	0.165	0.954	**0.353**	0.026
				S.E.	0.142	0.093	1.386	5.906	0.148	0.089
			OSES ↔ OHEALTH	Coef.	0.373	0.117	0.643	0.079	**0.334**	0.442
				S.E.	0.219	0.104	19.53	0.148	0.116	0.302
Phase 1	Autoregression		CSES → ASES	Coef.	**0.461**	**0.537**	**0.310**	**0.529**	**0.449**	**0.531**
				S.E.	0.047	0.049	0.074	0.065	0.041	0.046
			CHEALTH → AHEALTH	Coef.	0.125	0.160	0.113	0.236	0.100	0.106
				S.E.	0.047	0.049	2.370	0.170	0.049	0.061
	Causation	Indirect 1	CSES → EDUC	Coef.	**0.461**	**0.522**	**0.593**	**0.509**	**0.542**	**0.525**
				S.E.	0.047	0.022	0.106	0.037	0.019	0.022
		Indirect 2	EDUC → AHEALTH	Coef.	**0.125**	**0.110**	0.222	**0.026**	**0.100**	**0.069**
				S.E.	0.047	0.049	1.455	0.048	0.027	0.042
		Indirect 1 × Indirect 2	(CSES → EDUC) × (EDUC → AHEALTH)	Coef.	0.461	0.058	0.132	0.013	**0.054**	0.036
				S.E.	0.047	0.026	0.841	0.057	0.021	0.022
		Direct	CSES → AHEALTH	Coef.	0.125	0.017	−0.264	0.009	−0.010	0.007
				S.E.	0.047	0.043	1.396	0.063	0.061	0.042
		Total	CSES → AHEALTH	Coef.	0.461	0.075	−0.132	0.022	0.044	0.043
				S.E.	0.047	0.033	2.230	0.057	0.047	0.032
	Selection	Indirect 1	CHEALTH → EDUC	Coef.	0.011	−0.047	0.042	**−0.103**	0.015	−0.026
				S.E.	0.022	0.031	0.066	0.052	0.020	0.023
		Indirect 2	EDUC → ASES	Coef.	**0.522**	**0.498**	**0.290**	**0.408**	**0.572**	**0.500**
				S.E.	0.042	0.044	0.074	0.057	0.037	0.044
		Indirect 1 × Indirect 2	(CHEALTH → EDUC) × (EDUC → ASES)	Coef.	0.006	−0.024	0.012	−0.042	0.009	−0.013
				S.E.	0.012	0.015	0.019	0.023	0.012	0.012
		Direct	CHEALTH → ASES	Coef.	0.025	0.046	−0.003	0.064	0.021	0.047
				S.E.	0.031	0.044	0.141	0.070	0.027	0.046
		Total	CHEALTH → ASES	Coef.	0.032	0.023	0.009	0.022	0.030	0.034

			S.E.	0.034	0.045	0.150	0.078	0.031	0.049
Phase 2	Autoregression	ASES → OSES	Coef.	**0.771**	**0.372**	**0.686**	**0.472**	**0.614**	**0.795**
			S.E.	0.138	0.227	0.219	0.083	0.070	0.187
		AHEALTH → OHEALTH	Coef.	**0.154**	**0.253**	0.202	0.478	**0.173**	**0.167**
			S.E.	0.055	0.047	18.13	0.257	0.044	0.059
	Causation	ASES → OHEALTH	Coef.	**0.155**	**0.176**	0.218	**0.317**	**0.187**	**0.226**
			S.E.	0.034	0.035	5.182	0.109	0.030	0.033
	Selection	AHEALTH → OSES	Coef.	−0.006	0.036	0.055	0.078	−0.001	0.061
			S.E.	0.058	0.030	0.837	0.050	0.030	0.051

Note: a) standardised regression coefficients; S.E. = standard errors; C = childhood; A = adulthood (30–49); O = old age (50–90); SES = socioeconomic status; Phase 1 = transition from childhood to adulthood; Phase 2 = transition from adulthood to old age; b) 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; c) Statistically significant coefficients ($p < 0.05$, two-tailed) are printed in bold.

Appendix D. Three sensitivity analyses comparable to the main results in Table 2 of the main article:

1. Adding a direct path from education to old age health (all ages and only 50–74, remaining sample size 7213 men and 8828 women).
2. Excluding people with missing information on wages (remaining sample size 5588 men and 4342 women).
3. Excluding people mostly working in the household (remaining sample size 9045 men and 9890 women).

					Education direct		Educ. direct (ages 50–74)		With Wages		No 'Housewives'	
			Parameter		Male	Female	Male	Female	Male	Female	Male	Female
Correlation			CSES → CHEALTH	Coef.	0.000	**0.041**	0.020	**0.048**	0.007	0.031	0.001	**0.042**
				S.E.	0.020	0.020	0.023	0.022	0.025	0.031	0.020	0.021
			ASES → AHEALTH	Coef.	**0.204**	0.070	0.121	0.040	0.088	0.143	0.208	0.109
				S.E.	0.070	0.060	0.067	0.071	0.054	0.082	0.070	0.056
			OSES → OHEALTH	Coef.	**0.282**	**0.289**	**0.378**	**0.326**	**0.299**	**0.354**	**0.288**	**0.348**
				S.E.	0.078	0.086	0.116	0.100	0.098	0.137	0.075	0.097
Phase 1	Autoregression		CSES → ASES	Coef.	**0.376**	**0.505**	**0.353**	**0.493**	**0.275**	**0.496**	**0.381**	**0.527**
				S.E.	0.028	0.028	0.030	0.034	0.024	0.039	0.027	0.026
			CHEALTH → AHEALTH	Coef.	**0.146**	**0.220**	**0.168**	**0.226**	**0.169**	**0.282**	**0.143**	**0.224**
				S.E.	0.025	0.033	0.029	0.038	0.033	0.059	0.024	0.037
	Causation	Indirect 1	CSES → EDUC	Coef.	**0.508**	**0.501**	**0.523**	**0.517**	**0.519**	**0.543**	**0.507**	**0.504**
				S.E.	0.013	0.011	0.015	0.012	0.016	0.017	0.013	0.012
		Indirect 2	EDUC → AHEALTH	Coef.	**0.083**	**0.110**	**0.080**	**0.116**	**0.071**	0.046	**0.075**	**0.082**
				S.E.	0.028	0.027	0.034	0.029	0.031	0.037	0.027	0.026
		Indirect 1 × Indirect 2	(CSES → EDUC) × (EDUC → AHEALTH)	Coef.	**0.042**	**0.055**	**0.042**	**0.060**	**0.037**	0.025	**0.038**	**0.041**
				S.E.	0.014	0.014	0.018	0.015	0.016	0.020	0.014	0.013
		Direct	CSES → AHEALTH	Coef.	0.021	0.007	0.012	0.002	0.011	0.074	0.026	0.031
				S.E.	0.029	0.029	0.035	0.032	0.039	0.043	0.029	0.029
		Total	CSES → AHEALTH	Coef.	**0.063**	**0.062**	**0.054**	**0.061**	0.047	**0.099**	**0.065**	**0.072**
				S.E.	0.022	0.022	0.026	0.024	0.031	0.034	0.022	0.023
	Selection	Indirect 1	CHEALTH → EDUC	Coef.	0.009	0.007	0.021	0.007	0.005	0.029	0.008	0.008
				S.E.	0.014	0.012	0.016	0.013	0.018	0.021	0.014	0.014
		Indirect 2	EDUC → ASES	Coef.	**0.425**	**0.485**	**0.360**	**0.499**	**0.299**	**0.376**	**0.422**	**0.456**
				S.E.	0.025	0.026	0.028	0.030	0.022	0.034	0.025	0.023
		Indirect 1 × Indirect 2	(CHEALTH → EDUC) × (EDUC → ASES)	Coef.	0.004	0.003	0.008	0.003	0.002	0.011	0.003	0.004
				S.E.	0.006	0.006	0.006	0.007	0.005	0.008	0.006	0.006
		Direct	CHEALTH → ASES	Coef.	**0.056**	**0.129**	**0.085**	**0.173**	0.030	**0.080**	**0.053**	**0.106**
				S.E.	0.019	0.025	0.025	0.029	0.019	0.030	0.019	0.023
		Total	CHEALTH → ASES	Coef.	**0.060**	**0.133**	**0.093**	0.176	**0.031**	**0.091**	**0.056**	**0.110**
				S.E.	0.021	0.027	0.026	0.031	0.020	0.032	0.020	0.024

Phase 2	Autoregression	ASES → OSES	Coef.	**0.737**	**0.689**	**0.773**	**0.680**	**1.114**	**0.733**	**0.731**	**0.681**
			S.E.	0.065	0.054	0.079	0.058	0.118	0.101	0.065	0.55
		AHEALTH → OHEALTH	Coef.	**0.186**	**0.321**	**0.352**	**0.413**	**0.179**	**0.455**	**0.186**	**0.357**
			S.E.	0.026	0.037	0.046	0.050	0.037	0.079	0.026	0.044
	Causation	ASES → OHEALTH	Coef.	**0.261**	**0.356**	**0.560**	**0.551**	**0.245**	**0.286**	**0.201**	**0.231**
			S.E.	0.059	0.054	0.115	0.084	0.046	0.044	0.026	0.022
	Selection	AHEALTH → OSES	Coef.	0.008	0.008	0.013	-0.001	0.049	0.083	0.005	0.003
			S.E.	0.043	0.041	0.051	0.043	0.041	0.045	0.042	0.045
		EDUC → OHEALTH	Coef.	-0.045	-0.119	-0.106	-0.235	–	–	–	–
			S.E.	0.038	0.045	0.066	0.070	–	–	–	–

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. http://dx.doi.org/10.1016/S0304-4076(02)00145-8.

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 & Medicine, 75*(7), 1303–1310. http://dx.doi.org/10.1016/j.socscimed.2012.05.015.

Altzinger, W., Crespo Cuaresma, J., Rumplmaier, B., Sauer, P., & Schneebaum, A. (2015). *Education and social mobility in europe: Levelling the playing field for europe's children and fuelling its economy*.

Börsch-Supan, A., Brandt, M., Hunkler, C., Kneip, T., Korbmacher, J., Malter, F., ... Zuber, S., ... (2013). Data resource profile: The survey of health, ageing and retirement in europe (SHARE). *International Journal of Epidemiology, 42*(4), 992–1001.

Blane, D., Davey Smith, G., & Bartley, M. (1993). Social selection: What does it contribute to social class differences in health? *Sociology of Health and 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. http://dx.doi.org/10.1016/j.respe.2006.12.004.

Blossfeld, H.-P., Golsch, K., & Rohwer, G. (2007). *Event history analysis with stata* (1 ed.). New York: Psychology Press.

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* (pp. 83–109). Netherlands: Springer.

Bollen, K. A. (1989). *Structural equations with latent variables.* New York, NY: Wiley–Interscience.

Case, A., & Deaton, A. S. (2005). *Broken down by work and sex: How our health declines Analyses in the Economics of Aging.* University of Chicago Press185–212.

Chandola, T., Bartley, M., Sacker, A., Jenkinson, C., & Marmot, M. (2003). Health selection in the Whitehall II study UK. *Social Science & Medicine, 56*, 2059–2072.

Cox, D. R. (1992). Causality: Some statistical aspects. *Journal of the Royal Statistical Society. Series A, 155*, 291–301. http://dx.doi.org/10.2307/2982962.

Dannefer, D., & Kelley-Moore, J. A. (2009). *Theorizing the life course: New twists in the path handbook of theories of aging.* New York: Springer389–412.

Dannefer, D. (2003). Cumulative Advantage/Disadvantage and the life course: Cross-Fertilizing age and social science theory. *The Journals of Gerontology, Series B: Social Sciences, 58*(6), S327–337.

Doblhammer, G. (2004). *The late life legacy of very early life.* Berlin, Heidelberg: Springer.

Eikemo, T. A., Huisman, M., Bambra, C., & Kunst, A. E. (2008). Health inequalities according to educational level in different welfare regimes: A comparison of 23 European countries. *Sociology of Health & Illness, 30*(4), 565–582. http://dx.doi.org/10.1111/j. 1467-9566.2007.01073.x.

Elo, I. T. (2009). Social class differentials in health and mortality: Patterns and explanations in comparative perspective. *Annual Review of Sociology, 35*, 553–572. http://dx.doi.org/10.1146/annurev-soc-070308-115929.

Erikson, R., & Goldthorpe, J. H. (1992). *The constant flux. a study of class mobility in industrial societies.* Oxford: Clarendon Press.

Ferrera, M. (1996). The 'Southern model' of welfare in social europe. *Journal of European Social Policy, 6*(1), 17–37. http://dx.doi.org/10.1177/095892879600600102.

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 & Medicine, 150*, 172–183. http://dx.doi.org/10.1016/j.socscimed.2015.12.021.

Fritzell, J. (2014). Health inequality and social justice. In M. Reisch (Ed.). *Routledge international handbook of social justice* (pp. 339–352). London/New York: Routledge.

Fuchs, V. R. (1982). Time preference and Health: An explanatory study. In V. R. Fuchs (Ed.). *Economic aspects of health* (pp. 93–120). Chicago: University of Chicago Press.

Galama, T., & van Kippersluis, H. (2010). *A theory of socioeconomic disparities in health over the life cycle*.

Garrouste, C., & Paccagnella, O. (2011). Three examples of consistency across SHARE and SHARELIFE 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 & Medicine, 127*, 74–82.

Goldman, N. (2001a). Social inequalities in health: Disentangling the underlying mechanisms. *Annals of the New York Academy of Sciences, 954*, 118–139. http://dx.doi.org/10.1111/j. 1749-6632.2001. tb02750.x.

Goldman, N. (2001b). *Mortality differentials: Selection and causation international encyclopedia of social and behavioral sciences.* 10068–10070.

Goldthorpe, J. H. (2010). Analysing social inequality: A critique of two recent contributions from economics and epidemiology. *European Sociological Review, 26*(6), 731–744. http://dx.doi.org/10.1093/esr/jcp046.

Graham, H. (2009). *Understanding health inequalities.* Open University Press.

Haas, S. A. (2006). Health selection and the process of social stratification: The effect of childhood health on socioeconomic attainment. *Journal of Health and Social Behavior, 47*(4), 339–354.

Haas, S. A. (2007). The long-term effects of poor childhood health: An assessment and application of retrospective reports. *Demography, 44*, 113–135.

Hallquist, M., & Wiley, J. (2011). *MplusAutomation: Automating Mplus model estimation and interpretation. R package version 0.6.* Available online at: http://cran.r-project.org/web/packages/MplusAutomation/index.htmlhttp://cran.r-project.org/web/packages/MplusAutomation/index.html.

Havari, E., & Mazzonna, F. (2015). *European Journal of Population, 31*, 233. http://dx.doi.org/10.1007/s10680-014-9332-y.

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*. 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* (pp. 67–92). New York: Aldine de Gruyter.

Hoffmann, R. (2008). *Socioeconomic differences in old age mortality, Vol. 25*. Dordrecht: Springer.

Hoffmann, R. (2011). Socioeconomic inequalities in old-age mortality: A comparison of Denmark and the USA. *Social Science & Medicine, 72*(12), 1986–1992. http://dx.doi.org/10.1016/j.socscimed.2011.04.019.

Hummer, R. A., Rogers, R. G., & Eberstein, I. W. (1998). Sociodemographic differentials in adult mortality: A review of analytic approaches. *Population and Development Review, 24*(3), 553–578.

Huurre, 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. http://dx.doi.org/10.1007/s00127-005-0930-1.

International Labour Organization (2012). *International Standard Classifi cation of Occupations. Structure, group definitions and correspondence tables.* Geneva.

Jylhä, M. (2009). What is self-rated health and why does it predict mortality? Towards a unified conceptual model. *Social Science & Medicine, 69*(3), 307–316.

Kapteyn, A., Ypma, J., & xa, Y. (2007). Measurement error and misclassification: A comparison of survey and administrative data. *Journal of Labor Economics, 25*(3), 513–551. http://dx.doi.org/10.1086/513298.

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. http://dx.doi.org/10.1093/eurpub/ckv111.

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. http://dx.doi.org/10.1016/j.puhe.2015.10.022.

Kröger, H. (2015b). Newspell—Easy management of complex spell data. *Stata Journal, 15*, 155–172.

Kuh, D., Ben-Shlomo, Y., Lynch, J., Hallqvist, J., & Power, C. (2003). Life course epidemiology. *Journal of Epidemiology and Community Health, 57*, 778–783. http://dx.doi.org/10.1136/jech.57.10.778.

Lee, D., & Jackson, M. (2017). The simultaneous effects of socioeconomic disadvantage and child health on children's cognitive development. *Demography, 54*(5), 1845–1871.

Lundberg, O. (1991). Childhood living conditions, health status: And social mobility: A contribution to the health selection debate. *European Sociological Review, 7*, 149–162.
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* jech 2010.112680.
Mackenbach, J. P., Kulhánová, I., Menvielle, G., Bopp, M., Borrell, C., Costa, G., & Kovacs, K. (2015). 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. (2012). The persistence of health inequalities in modern welfare states: The explanation of a paradox. *Social Science & Medicine, 75*(4), 761–769.
Manor, O., Matthews, S., & Power, C. (2003). Health selection: The role of inter- and intra-generational mobility on social inequalities in health. *Social Science & Medicine, 57*(11), 2217–2227. http://dx.doi.org/10.1016/S0277-9536(03)00097-2.
Mayer, K. U. (2009). New directions in life course research. *Annual Review of Sociology, 35*, 413–433. http://dx.doi.org/10.1146/annurev.soc.34.040507.134619.
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. http://dx.doi.org/10.2307/3090243.
Muthen, L. K., & Muthen, B. O. (2015). *Mplus user's guide: Seventh edition.* Los Angeles: Muthén & Muthén.
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. http://dx.doi.org/10.1080/13645579.2015.1091641.
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. http://dx.doi.org/10.1016/j.alcr.2016.10.003.
Palloni, A., White, R. G., & Milesi, C. (2008). *The size of health selection effects.* Retrieved from.
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 & Medicine, 68*, 1574–1582.
Rubin, D. B. (2005). Causal inference using potential outcomes. *Journal of the American Statistical Association, 100*, 322–331. http://dx.doi.org/10.1198/016214504000001880.
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. http://dx.doi.org/10.1111/1467-985x.00257.
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*.
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 investigations in the economics of aging.* University of Chicago Press267–317.
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.
Thorley, C. (2016). *Education, education, mental health: Supporting secondary schools to play a central role in early intervention mental health services.* London: Institute for Public Policy Research (IPPR).
Valkonen, T. (1996). Social inequalities in mortality. *Demographie: analyse et synthese: causes et consequences des evolutions demographiques: actes du Seminaire international "Demography: analysis and synthesis"* (Vol. 3) Sienne, Chartreuse du Pontignano, 22–24 avril.
Van De Mheen, H., Stronks, K., Looman, C. W. N., & Mackenbach, J. P. (1998). Recall bias in self-reported childhood health: Differences by age and educational level. *Sociology of Health & Illness, 20*, 241–254. http://dx.doi.org/10.1111/1467-9566.00099.
Van De Mheen, H. D., Stronks, K., & Mackenbach, J. P. (1998). A lifecourse perspective on socio-economic inequalities in health: The influence of childhood socio-economic conditions and selection processes. *Sociology of Health & Illness, 20*, 754–777. http://dx.doi.org/10.1111/1467-9566.00128.
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. http://dx.doi.org/10.1353/sof.0.0219.
Weiss, C. T. (2012). *Two measures of lifetime resources for Europe using SHARELIFE*.
West, P. (1991). Rethinking the health selection explanation for health inequalities. *Social Science & Medicine, 32*, 373–384. http://dx.doi.org/10.1016/0277-9536(91)90338-D.