HEDG

HEALTH, econometrics AND DATA GROUP

THE UNIVERSITY of York

WP 21/15

Socioeconomic Health Inequalities: Differences Between and Within Individuals

Fiorella Parra-Mujica; Matthew Robson and Richard Cookson

August 2021

http://www.york.ac.uk/economics/postgrad/herc/hedg/wps/

Socioeconomic Health Inequalities: Differences Between and Within Individuals

Fiorella Parra-Mujica[†] Matthew Robson[‡] Richard Cookson[§¶]

August 23, 2021

Abstract

Research on socioeconomic health inequalities has primarily relied on univariate markers of socioeconomic status (SES), measured at one point in time. Using data from the UKHLS dataset (2009–2020), we build an age-conditional multidimensional SES index for the adult population in the UK. By using a "within-between" model we disentangle the relationship between health outcomes and: i) between-individual differences in SES, and ii) within-individual variations of SES across time. Results show that both are positive and highly significant predictors of physical and mental health. However, we find that these relationships are not linear and that within and between effects interact. While higher levels of SES are always associated with an increase in physical health, for mental health, after some point (SES = 0.613), higher average SES is associated with a score decrease. For an individual with the lowest average SES rank we observe a large and significant effect of a within-individual increase in SES on mental health (8.91) and physical health (1.82), however, this within-effect diminishes for those individuals with higher average SES. Individuals with high expected mobility are also found to have significantly better health scores, particularly for mental health.

Keywords: Health Inequality, Socioeconomic Status, Social Gradient, Mundlak

JEL Codes: I14, C23, D63

[†]Corresponding Author. (1) Health Economics Research Centre, Nuffield Department of Population Health, University of Oxford, Oxford, UK. (2) Department of Health Sciences, University of York, UK. Email: fiorella.parramujica@ndph.ox.ac.uk.

[‡]Department of Health Sciences, University of York, UK. Email: matthew.robson@york.ac.uk. Website: https://mrobson92.com.

[§]Centre for Health Economics, University of York, UK. Email: richard.cookson@york.ac.uk.

[¶]We would like to thank Jonathan Bradshaw, Andrew Jones, Tim Doran, Jan Abel Olsen and Nigel Rice for their helpful comments and discussion. The study was funded by the Wellcome Trust (205427/Z/16/Z). All errors remain our own.

1 Introduction

Individuals with lower socioeconomic status (SES) present higher morbidity and mortality rates, a phenomenon that has drawn the attention of both policy makers and academics (Elo, 2009; Cutler, Lleras-Muney, and Vogl, 2012; Hoffmann, Kröger, and Pakpahan, 2018; Galama and Kippersluis, 2018). In the fields of public health, epidemiology and sociology, health inequalities are primarily assessed by analysing the differences between social groups using population level data. Economists, on the other hand, primarily rely on proxies of SES, such as education, occupation or income, and then apply this unidimensional measure to analyse inter-individual inequality (Stockwell et al., 1975; Marmot, Rose, et al., 1978; Pappas et al., 1993; Case et al., 2002; Nazroo, 2003). The health inequalities literature mainly falls within either of these approaches, a separation that misses out on several nuances in the SES-health relationship. First, SES is a multidimensional concept which cannot be fully captured through a proxy. Second, SES is not static but evolves through the life course. Third, there is a distinction between: i) how differences in SES between individuals are associated to differences in health, and ii) how individual's fluctuations in SES over time are associated with changes in health.

In this paper, we integrate these approaches to address the evolving and multidimensional nature of SES, and analyse the nuances of its relationship with physical and mental health across time (Dutton and Levine, 1989; Adler, Boyce, et al., 1994; Savage et al., 2013; Savage, 2015). Instead of relying on a unidimensional measure of SES, we develop a multidimensional measure which is estimated conditionally on age and comprises of both "sticky" variables (such as wealth, education) and "fluid" variables (e.g. income, occupation) (Lee and Jackson, 2017; Hoffmann, Kröger, and Pakpahan, 2018). By using a "within-between" model we are able to disentangle the relationship between health outcomes and: i) between-individual differences in SES, and ii) within-individual variations of SES across time. Moreover, by decomposing current SES, we are able to uncover the relationship between socioeconomic mobility and health.

We exploit the panel structure of the UK Household Longitudinal Study, between 2009 and 2020, to observe individual's health and socioeconomic status over time. We use the SF-12 Mental Composite Scale (MCS) and the SF-12 Physical Composite Scale (PCS) as measures of mental and physical health. We use Principal Component Analysis (PCA) to incorporate multidimensional attributes of SES and derive a single, continuous, numerical index, which determines the relative SES of an individual within their age group. The index is contingent upon both continuous variables (e.g. income) and more discrete class distinctions (e.g. NS-SEC). Although imperfect, this data-driven approach incorporates an array of variables such as income, labour market status, housing and qualifications.

Together, these methods allow us to reach three objectives: 1) to capture the several dimensions of SES and mobility; 2) to disentangle the relationship between physical and mental health outcomes and: i) between-individual differences in SES, and ii) within-individual variations of SES across time; and 3) to compare the results of this multidimensional measure with the standard univariate measures equivalised income, occupation, and education.

Our research provides insights on the sources of the social gradient of health, necessary for the design of policy interventions to eliminate it. In the first place, it provides a broad evaluation of the aspects that shape SES and have a potential impact on health through adulthood. Furthermore, it contributes to a better understanding of the social factors behind healthier (or unhealthier) aging and health deterioration. We also contribute to the discussion on how social mobility and stressful factors of social environments influence physical and mental health through adulthood.

The remainder of the paper is structured as follows. Section 2 introduces the literature on socioeconomic status and its relationship with health. Section 3 outlines the methods used for the PCA and the "within-between" specification, whilst Section 4 describes the data. Section 5 presents the results, where special attention is paid to the difference between the "within" and "between" estimators, the role of mobility, and the comparison of our multidimensional index to other univariate proxies of SES . Section 6 provides a discussion, and Section 7 concludes.

2 Conceptual Framework

2.1 Socioeconomic status (SES) and Health

Socioeconomic health inequalities have long caught the attention of scholars and policy makers. As early as 1842, Chadwick (1842) concluded that a socioeconomic gradient in health was caused by the vast differences in living conditions among different social classes, classified by occupation. The Black Report (1980) and work by Marmot, Shipley and Rose (1984) also found significant differences in mortality across social classes, in detriment of the lowest social classes (Department of Health and Social Security, 1980; Marmot, Shipley, et al., 1984; Feinstein, 1993). More recently, studies have confirmed that individuals with lower socioeconomic status (SES) present higher morbidity and mortality rates (Elo, 2009; Cutler, Lleras-Muney, and Vogl, 2012; Hoffmann, Kröger, and Pakpahan, 2018; Galama and Kippersluis, 2018).

SES comprises current (and past) income, wealth, education, occupation, neighbourhood, and social networks, factors which relate to health through several mechanisms (Cutler and Lleras-Muney, 2010; Mosby, 2013; Savage et al., 2013; Lee and Jackson, 2017; Hoffmann, Kröger, and Pakpahan, 2018) and cause socioeconomic gradients in health. Furthermore, although socioeconomic status has a historical component based on the accumulation of advantages (or disadvantages) over generations, SES evolves through the life course due to personal and general circumstances (Biressi and Nunn, 2013; Savage, 2015). Hence, the different aspects of SES are interrelated through the life course: for instance, a good education during childhood might influence a person's adult occupation and income (Galobardes et al., 2006).

However, these aspects of SES are not usually considered when analysing socioeconomic inequalities in health. Economists usually apply unidimensional and time-invariant proxies of SES, such as education, occupation or income to analyse inter-individual inequality (e.g. bottom vs. top of the distribution) (Ettner, 1996; Hauck and Rice, 2004; Wiggins et al., 2004; Frijters et al., 2005; Apouey and Geoffard, 2013; Davillas, Andrew Jones, and

Benzeval, 2019). On the other hand, sociologists and epidemiologists focus on systematic disparities, or how disadvantage is concentrated in certain groups (defined by social class, gender, ethnicity, etc.) which display worse morbidity and mortality outcomes (Adler and Stewart, 2010; Eyal, 2018). Using the quantitative tools provided by economics while acknowledging the multi-dimensional and evolving nature of SES can lead to a deeper understanding of the relationship between, and inequalities in, socioeconomic status and health.

This paper integrates concepts and methods from economics, epidemiology and sociology, as a means to a powerful framework for the analysis of the socioeconomic status, mobility and health. We consider "fluid" variables, such as income and labour market status, alongside more "sticky" crystallised variables, such as educational attainment. This methodology is intended to be pragmatic and flexible; allowing for variables to be included if and when they are available. On the other hand, the decomposition of SES into component parts sheds a light on the effect that the SES dynamics have on health through time.

2.2 Socioeconomic Status Over Time

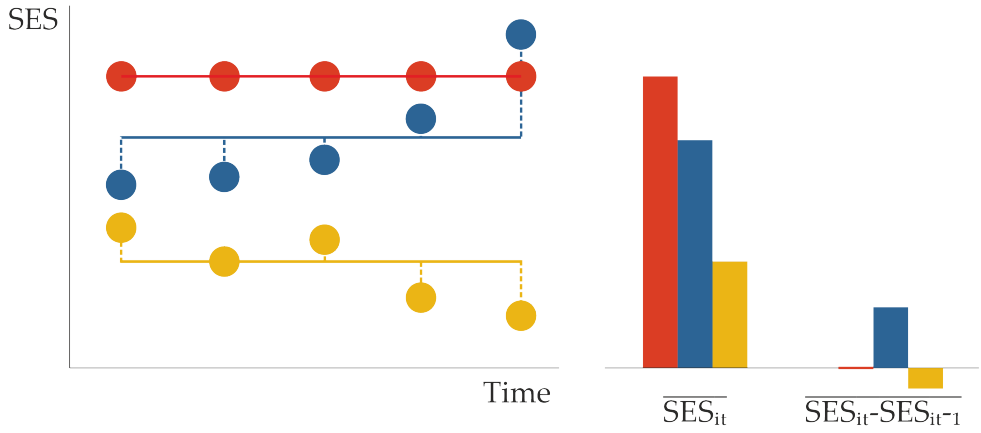
To disentangle the relation between health and socioeconomic status we use the panel structure of the data to decompose SES into component parts, in two ways. Figure 1 illustrates the concepts behind our approach.

Imagine we have three individuals: red, blue and yellow. Socioeconomic status, SES_{it} is observed for these individuals, i, across multiple time periods, t. This is plotted in the left panel of Figure 1. We observe differences *between* individuals, as well as changes *within* individual over time. Average levels, variation and trends all differ, and may be important in determining socioeconomic health inequalities.

Our first approach decomposes socioeconomic status into average levels and within-individual differences:

$$SES_{it} = \underbrace{\overline{SES_{it}}}_{\text{Between}} + \underbrace{(SES_{it} - \overline{SES_{it}})}_{\text{Within}} \tag{1}$$

Figure 1: Socioeconomic Status Through Time



The average level of SES is calculated for each individual: $\overline{SES_{it}} = \frac{1}{T}\sum_t^T SES_{it}$. This level is shown as the horizontal line in the left panel and the first bar chart on the right. As shown, red has the highest average level of SES and yellow the least. Within-individual differences are illustrated by the dashed vertical lines, as: $SES_{it} - \overline{SES_{it}}$. This shows the extent and direction of change in an individual's current SES with respect to their average. Here, red is shown to have no within-individual differences, maintaining the same level of SES throughout, while both blue's and yellow's SES changes over time.

One limitation with this approach, is that it does not consider the trend of SES over time. It might be important that individual's SES increases or decreases over time. To explore this, our second approach first decomposes current socioeconomic status into *lagged* SES and

mobility, the difference between current and lagged SES, and then further decomposes these terms into average and within levels:

$$\begin{aligned} SES_{it} &= \underbrace{SES_{it-1}}_{\text{Lag}} + \underbrace{(SES_{it} - SES_{it-1})}_{\text{Mobility}} \\ &= \underbrace{\overline{SES_{it-1}}}_{\text{Average-Lag}} + \underbrace{(SES_{it-1} - \overline{SES_{it-1}})}_{\text{Within-Lag}} \\ &\quad + \underbrace{\overline{(SES_{it} - SES_{it-1})}}_{\text{Expected-Mobility}} + \underbrace{\overline{((SES_{it} - SES_{it-1}) - \overline{(SES_{it} - SES_{it-1})})}}_{\text{Within-Mobility}} \end{aligned} \tag{2}$$

Of primary interest are the third and fourth terms. The third term, is a measure of *expected* socioeconomic mobility. This is illustrated in the right bar chart, where red has no mobility; blue has positive expected mobility, as their SES increases year on year; while yellow has negative expected mobility. The fourth term shows a within-individual deviation from their expected trajectory. For example, blue in the last period has a significantly large increase in SES than is typical.

With the first decomposition, we can explore whether it is *between* individual differences in average socioeconomic status ($\overline{SES_{it}}$) or *within* individual changes in socioeconomic status ($SES_{it} - \overline{SES_{it}}$) which have the strongest relationship with health. With the second decomposition, we can identify if individuals with higher *expected socioeconomic mobility*, or have particularly steep within-individual improvements, have higher levels of health.

3 Methods

3.1 Socioeconomic Status (SES) Index

Principal component analysis (PCA) will be used for identifying the latent construct of SES. This method transforms a number of correlated variables into a (smaller) number of uncorrelated principal components. The first component is the linear combination of

variables that retains the most of the variation present in all of the original data (Vyas and Kumaranayake, 2006; Jolliffe and Cadima, 2016).

The relation between each individual characteristic and overall SES is estimated by using the weights from the first component of the PCA. The higher the (positive) weight, the more having that characteristic denotes an increase in SES. Hence, it is the combination and interaction of an array of characteristics which will determine each individual's SES. This method derives a single (continuous) numerical index, which determines the relative SES of an individual within the sample. The index is contingent upon both continuous variables (i.e. income) and more discrete class distinctions (i.e. NS-SEC).

Importantly, SES is estimated conditional upon age group. This separation allows for characteristics to have different weights throughout the life course. Therefore, this index more closely resembles life-time potential SES, as some factors are more important than others to determine SES at certain stages of life. Moreover, the composite SES includes "sticky" variables such as house ownership, most recent occupation and highest qualification along with more "fluid" variables such as income and current labour market status.

The component score, a continuous variable, will be used as the construct of SES. The component score is obtained from the product of the weight of each variable (components' loading) and the variable's value, which then is aggregated. Higher component scores denote higher SES in each age group. Results are then ranked, conditional upon the age group, from lowest to highest, and normalised to a 0 to 1 scale.

In order to compare results using the composite SES index with alternative univariate measures of socioeconomic status (e.g. income, occupation, education, home ownership), similar age-conditional rank transformations are conducted. Providing an age-conditional normalised rank from 0 to 1 for each alternative measure.

3.2 Within-Between Model

We use a "within-between" specification, with lagged health outcomes, to model socioeconomic health inequalities. A reformulation of the Mundlak model (Mundlak, 1978), the

"within-between" model aims to retain the flexibility of Random Effects (RE) models, whilst reducing the concerns of bias that Fixed Effects (FE) models are able to address. While RE models allow for the analysis of group-level characteristics, its assumption of covariate exogeneity is often unfeasible and leads to biased estimators. FE models offer unbiased estimators, but are more limited for research trying to analyse group-level time-invariant characteristics and cannot always answer questions relevant for equity research. Hence, a "within-between" specification provides an alternative, evidence-based approach to endogeneity which enables the decomposition of regressors into their time-varying and time-invariant components (Bell, Fairbrother, et al., 2019). The "within-between" model provides the same results as the FE model for the within effect, but retains the between effect which cannot be measured with a FE model (Dieleman and Templin, 2014; Bell, Fairbrother, et al., 2019; Bell and K. Jones, 2015). It accounts for the correlation between the group-level effects and the explanatory variables by including the group mean in the regression (Dieleman and Templin, 2014).

Our "within-between" model is specified as follows:[1]

$$h_{it} = \boldsymbol{\beta}'\overline{x}_i + \boldsymbol{\gamma}'(x_{it} - \overline{x}_i) + \boldsymbol{\delta}'h_{it-1} + \boldsymbol{\eta}'c_i + (u_i + e_{it}) \tag{3}$$

Where h_{it} is the dependent variable, the health outcome, for individual i at time t. x_{it} is the set of observed variables of interest, separated into: $\overline{x}_i$, individual i's average over the sample period, and $(x_{it} - \overline{x}_i)$, i's difference from their average at time t.[2]

The corresponding coefficients are β, the "between-effects", and γ, the "within-effects". Lagged health variables are represented by h_{it-1} , while c_i are the time-invariant control variables,[3] with δ and η as the corresponding coefficients. The residuals u_i and e_{it} are assumed to have mean zero and be normally distributed.

[1]Throughout we use the *xthybrid* in Stata for estimation (Perales and Schunck, 2020).

[2]Variables include all time varying observable variables: our main SES variables of interest, alongside time varying demographic controls (age, age squared, married, divorced) and an attrition-based control (present in next wave).

[3]These include demographic controls (sex, white and UK born), initial health states (h_{i1}) and attrition controls (wave count and all waves).

Intuitively this approach allows the identification of the "between" effects, as the effect of differences in SES between persons, and the "within" effects as the effect of an increase in an individual's SES, relative to their average SES. Furthermore, the inclusion of a lagged health variable is an attempt to reduce the impact that reverse causality has on the estimates. Eq. 3 controls for the previous years and the "initial" state of mental and physical health.

3.2.1 Relaxing the Functional Form

Our base model assumes linearity of the within and between effects, and does not account for any potential interaction of these effects. To relax these assumptions we formulate two alternative specifications.

The first model incorporates a quadratic term for both within and between effects:

$$h_{it} = \boldsymbol{\beta}'\overline{x}_i + \boldsymbol{\beta_A}'\overline{x}_i^2 + \boldsymbol{\gamma}'(x_{it} - \overline{x}_i) + \boldsymbol{\gamma_A}'(x_{it}^2 - \overline{x}_i^2) + \boldsymbol{\delta}' h_{it-1} + \boldsymbol{\eta}' c_i + (u_i + e_{it}) \tag{4}$$

By adding quadratic terms we can identify non-linearities in the relationships between health and between-individual differences in average SES and within-individual increases in SES. Coefficients β_A and γ_A capture the quadratic terms for the between and within effects, respectively. This uncovers differential marginal effects of an increase in SES.

The second model, allows for an interaction between the two levels:

$$h_{it} = \boldsymbol{\beta}'\overline{x}_i + \boldsymbol{\gamma}'(x_{it} - \overline{x}_i) + \boldsymbol{\gamma_B}'(\overline{x_i}(x_{it} - \overline{x}_i)) + \boldsymbol{\delta}' h_{it-1} + \boldsymbol{\eta}' c_i + (u_i + e_{it}) \tag{5}$$

This allows us to identify if a within-individual increase in SES depends on the average level of SES of individual. In other words, does an increase in within-individual SES for someone with lower average SES have a greater impact than on someone with higher average SES.

3.2.2 Socioeconomic Mobility

The current level of SES can be further decomposed by using lagged SES variables and the change between the current and lagged SES, as shown in Eq. 2, with their respective interpretation.

4 Data

4.1 The UK Household Longitudinal Study

This paper relies on the panel data available in the 10 waves of the UK Household Longitudinal Study (UKHLS). The UKHLS builds on its predecessor, the British Household Panel Survey (BHPS) and includes socioeconomic, demographic and health data of individuals living in private households in Great Britain between 2009 and 2020. Table 1 defines the main variables used: the health outcomes, socioeconomic variables and demographic characteristics.

Table 1: Variable Definitions

Variable	**Definition**
Physical Health	Physical Component Summary SF-12: 0-100
Mental Health	Mental Component Summary SF-12: 0-100
Income	(Log.) Net Household Equivalised Income (GBP 2015 prices, OECD scale)
Savings	Average monthly personal savings, GBP 2015 prices
Education	Highest Level of Education achieved: Degree, A-Level, GCSE, None
Occupation	Most recent occupation: National Statistics socioeconomic Classification
Labour Market Status	Current Labour Market Status: Employed, Unemployed, Retired, Other
Housing Tenure	1 if individual Owns their own house (with or without a mortgage)
Overcrowding	Number of household members divided per the number of rooms
Age	Age in Years
Sex	Sex: 0 if male, 1 if female
Ethnicity	Ethnicity: 0 if non-white, 1 if white
Married	Marital Status: 0 if unmarried, 1 if married
Divorced	Divorced Status: 0 if not divorced, 1 if divorced
Not Born in UK	Location of Birth: 0 if UK, 1 if UK

Health outcomes are derived from the SF-12 questionnaire. The SF-12 is a short generic health-related quality-of-life questionnaire used in large population health surveys, and whose results can reproduce two summary measures: the Physical Component Summary (PCS) and

the Mental Component Summary (MCS) (Vilagut et al., 2013). The questionnaire covers eight main domains of well-being, including: measures of physical functioning, bodily pain, general health, perceptions, social functioning, role limitations due to emotional problems, and mental health (Fleishman et al., 2010). The scores from these answers are used to calculate the PCS and MCS. We should point out that the patient-assessed outcomes of this questionnaire estimate physical function and mental well-being, rather than diagnosed or chronic conditions such as dementia, schizophrenia and diabetes.

The socioeconomic variables are primarily used to construct our multidimensional index of socioeconomic status. These include: household income, monthly savings, highest attained qualification, most recent occupation, current labour market status, housing tenure and overcrowding. Demographic characteristics are primarily included as control variables, with age, gender and ethnicity being used to conduct subgroup analysis.

5 Results

5.1 Descriptive Statistics

Table 2 presents the descriptive statistics across the sample, grouped by age. Physical SF-12 is shown to deteriorate with age, while Mental SF-12 is shown to increase. Higher numbers of females are found in younger age groups, as are non-white individuals and those not born in the UK. Marriage rates increase in midlife, and divorce rates are highest amongst those aged 51-60. Household income and monthly savings increase with age, but fall after 65. Occupation is inverse-U shaped, with the young and old occupying less management and professional roles, and more semi-routine and routine roles. Employment rates increase until mid-age and significantly decrease for the over 65s, who move to retirement. A significant minority, particularly of the youngest group, are classed as other; these are primarily students, those looking after the home and the long-term disabled. Younger age groups generally have higher levels of education and a significant minority of over 65s have no qualifications. Home ownership increases with age and overcrowding generally decreases.

Table 2: Descriptive Statistics by Age Group

	22-30 Mean	**31-40** Mean	**41-50** Mean	**51-60** Mean	**65+** Mean
Health					
- Physical Health (SF-12)	54.18	53.11	51.28	47.85	42.56
- Mental Health (SF-12)	47.54	47.89	48.48	49.84	52.15
Demographic Characteristics					
- Age	26.19	35.69	45.54	57.73	73.97
- Female	0.58	0.58	0.57	0.55	0.54
- White	0.81	0.81	0.85	0.91	0.96
- Married	0.23	0.60	0.64	0.66	0.59
- Divorced	0.00	0.03	0.08	0.12	0.09
- Not born in the UK	0.12	0.18	0.14	0.11	0.08
Income					
- (Log.) Household Eq.Income	7.33	7.36	7.39	7.42	7.28
- Savings (Monthly)	128.22	123.93	133.51	140.24	91.89
Occupation					
- Management and professional	0.35	0.45	0.43	0.39	0.32
- Intermediate	0.17	0.14	0.13	0.13	0.14
- Small employers and own account	0.05	0.08	0.09	0.11	0.11
- Lower supervisory and technical	0.08	0.07	0.07	0.07	0.08
- Semi-routine and routine	0.33	0.25	0.26	0.30	0.33
- Never employed	0.03	0.01	0.01	0.00	0.01
Labour market status					
- Employed	0.75	0.81	0.84	0.64	0.08
- Unemployed	0.07	0.05	0.05	0.04	0.00
- Retired	0.00	0.00	0.00	0.21	0.91
- Other	0.18	0.15	0.11	0.11	0.01
Education					
- Higher degree/ University	0.48	0.51	0.45	0.38	0.26
- A Level	0.27	0.21	0.19	0.18	0.13
- GCSE or other qualification	0.22	0.24	0.31	0.32	0.31
- No Qualification	0.02	0.03	0.05	0.12	0.30
Housing					
- Own (Outright or with mortgage)	0.52	0.66	0.76	0.80	0.82
- Overcrowded	0.73	0.76	0.67	0.50	0.41

Note: Descriptive statistics are shown per age group. Data shown is the complete (pooled) data used in our main analysis, with 42,373 individuals and 208,412 observations.

These differences in health and socioeconomic status across age groups highlight the need to account for age when exploring their association.

5.2 SES Composite Index

Table 3 describes the component loadings of the first principal component for each variable in each of the age groups. The component loadings are the correlations between the variable and the component. Variables that are positively correlated with a higher SES, such as income, savings or higher education, present positive component loadings. On the other hand, variables such as "overcrowding" or "never employed" present negative component

loadings. The linear combination of the component loadings and the variables' values for each individual give the score for each individual (Jolliffe and Cadima, 2016). In this case, a higher score denotes a higher SES. As a result, individuals with higher values for variables such as employment, education, savings or income, and lower values for overcrowding, have higher SES scores. Furthermore, each loading's absolute value represent the relative importance of each factor for determining SES.

Table 3: PCA weighting for SES variables

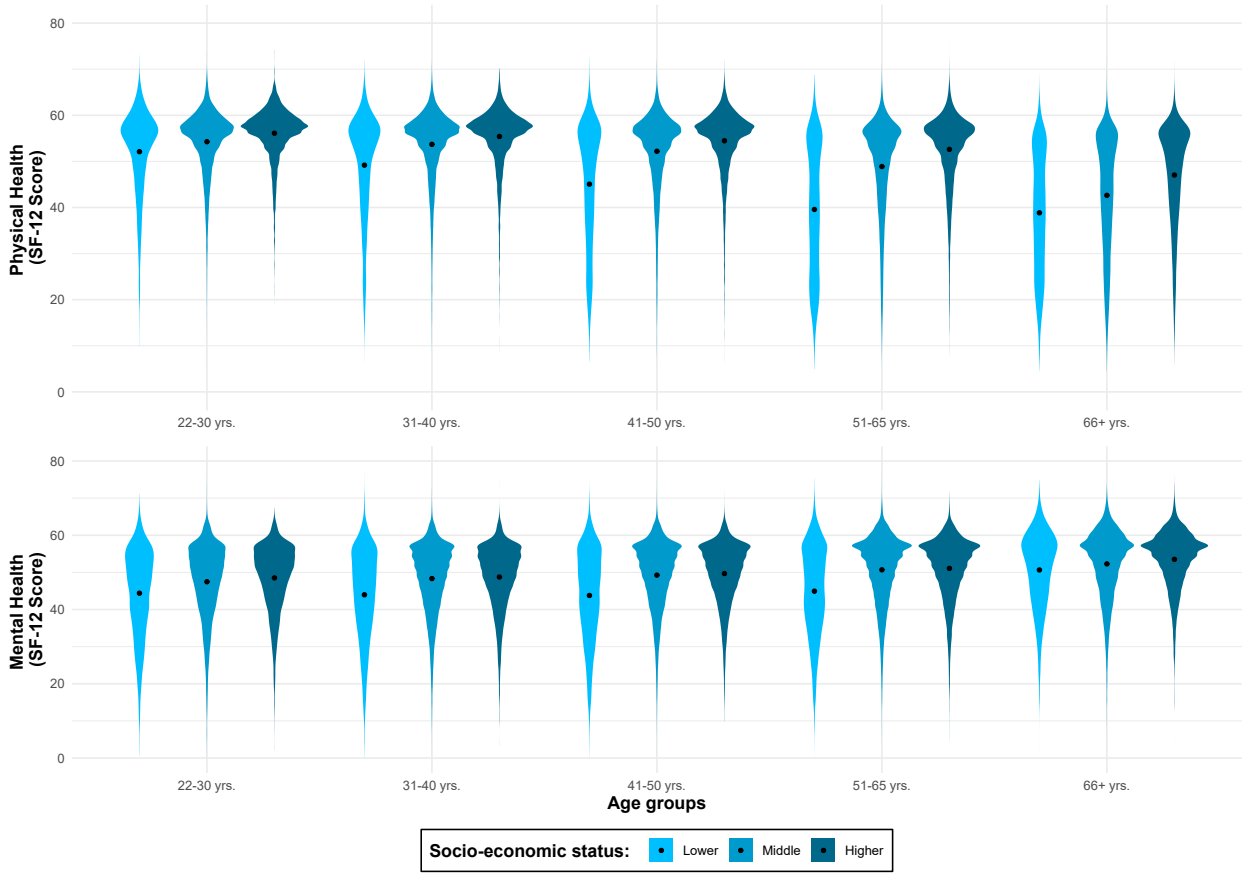
	22-30 Weights	**31-40** Weights	**41-50** Weights	**51-60** Weights	**65+** Weights
Income					
- (Log.) Household Eq.Income	0.354	0.362	0.342	0.339	0.364
- Savings (Monthly)	0.216	0.184	0.162	0.201	0.201
Occupation					
- Management and professional	0.345	0.369	0.373	0.404	0.417
- Intermediate	0.022	-0.030	-0.016	-0.010	0.015
- Small employers and own account	-0.001	-0.031	-0.034	-0.017	0.024
- Lower supervisory and technical	-0.004	-0.053	-0.055	-0.068	-0.078
- Semi-routine and routine	-0.295	-0.315	-0.328	-0.358	-0.378
- Never employed	-0.164	-0.127	-0.105	-0.068	-0.041
Labour market status					
- Employed	0.407	0.346	0.366	0.265	0.211
- Unemployed	-0.210	-0.188	-0.201	-0.146	0.004
- Retired	0.004	-0.003	-0.007	-0.035	-0.194
- Other	-0.318	-0.273	-0.290	-0.262	0.002
Highest qualification					
- Higher degree/ University	0.312	0.338	0.341	0.388	0.411
- A Level	-0.069	-0.068	-0.027	-0.005	0.026
- GCSE or other qualification	-0.244	-0.251	-0.239	-0.193	-0.010
- No Qualification	-0.138	-0.174	-0.205	-0.278	-0.387
Housing					
- Own (Outright or with mortgage)	0.218	0.271	0.292	0.294	0.245
- Overcrowded	-0.236	-0.253	-0.186	-0.176	-0.200
Observations	39921	57658	67653	88342	71052

The PCA results suggest that the relative importance of each variable changes across age groups. In the case of labour market status, it is more important for younger groups to currently be employed than it is for older groups. The variable "savings" increases in importance for Over 50s and under 30s, whereas housing tenure has increasing importance for older groups. Furthermore, regarding occupation, only holding a managerial position increases the SES of individuals between the ages of 30 and 50, whereas for the youngest group either holding an intermediate position or a managerial position is positively correlated

to a higher SES. As younger individuals usually lack the experience to take on management positions, holding an intermediate position is already positive for their SES.

5.3 Socioeconomic Status and Physical and Mental Health

Figure 2: SF-12 Physical and Mental Component Score, by SES and age



For lower 20%, middle 60% and upper 20% SES. All waves. N= 208,412.
The black dot represents the median, the width of each violin represents the density.

Figure 2 (top) illustrates the distribution of the SF-12 Physical Component Summary Scale (PCS) by SES and age group. The median physical score (represented by the black dots in the violin plots) is higher for younger individuals. A clear socioeconomic gradient of physical health is observed at all ages: lower SES groups show poorer physical health than higher SES groups. The median health outcomes of the lower SES groups are distinctly

poorer than those of the middle and higher SES groups. Furthermore, the high variance in physical health results among lower SES groups push down their median health outcomes. These long tails of deteriorated health among low SES groups accentuate starting at mid-life (age 40), whereas among the high SES groups, this health deterioration is only noticeable after age 66.

In the case of mental health (Figure 2, bottom), the social gradient has flattened and is only affected by low SES (poverty). As with physical health, long tails drag down the median of low SES groups, with the nuance that this happens at almost all ages. Finally, Figure 2 reveals that older groups tend to have higher SF12-MCS scores than their younger counterparts.

The distribution of health outcomes by SES and age group, represented by Figure 2, give an overview of the socioeconomic gradient of health. Furthermore, the wide differences between SES are observed notwithstanding a potential "healthy survivor" bias, as individuals with poor health are less likely to survive after 65.

5.4 Within-Between

While the above shows associations between the current level of SES (SES_{it}) and health, it does not uncover the extent to which these associations are driven by differences between individuals or changes within individuals across time. We can decompose SES into average SES and within-individual differences in SES, as explained in Section 2.2, to estimate between and within effects, respectively.[4]

Table 4 presents results from the base case "within-between" model, with Physical SF-12 and Mental SF-12 as the outcome variables. Results show that differences *between* individual's average levels of SES and *within* individual's changes of SES are both positive and highly significant predictors of physical and mental health.

[4]Descriptive statistics on within and between differences of the health and SES variables can be found in Appendix A.1.

Table 4: SES and SF-12 Score

	(1) Physical SF-12 Coef./S.E.	(2) Mental SF-12 Coef./S.E.
SES-Between	2.987***	1.225***
	(0.095)	(0.104)
SES-Within	0.968***	1.192***
	(0.163)	(0.184)
Constant	47.664***	48.767***
	(0.055)	(0.061)
Individuals	42373	42373
Observations	208412	208412
Controls	YES	YES
Log-Likelihood	-693470.0	-718081.8

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The between effect shows the difference in SF-12 scores between an individual who always has the lowest ranked SES ($\overline{SES_{it}} = 0$) to an individual with the highest ($\overline{SES_{it}} = 1$). For physical health, this is 2.99. This shows a steep socioeconomic gradient in physical health between individuals of differing average SES. For mental health, this difference is smaller (1.23), but still shows evidence of a significant socioeconomic gradient in mental health. The within-effect shows the difference in SF-12 score for an individual whose SES rank increases from the lowest to the highest rank. This effect is 0.97 for physical health, and 1.19 for mental health.

Table 5 shows the results of two specifications which allow for an additional flexibility. The first, Eq.4, incorporates quadratic terms of the within and between SES variables, and the second, Eq.5, allows for the interaction between the within and between levels of SES.

Models (1) and (3) show, for physical health, a monotonically increasing effect of average SES with diminishing marginal returns; while for mental health we observe an inverse-U. This implies that while higher levels of SES are associated with an increase in physical health, for mental health after some point ($SES = 0.613$) higher average SES is associated with a decrease in mental health. The results for the within effects have a similar trend for mental health, but are non-significant for physical health.

Table 5: SES and SF-12 Score: Quadratic and Interaction Model

	Physical SF-12		Mental SF-12	
	(1) Quadratic Coef./S.E.	(2) Interaction Coef./S.E.	(3) Quadratic Coef./S.E.	(4) Interaction Coef./S.E.
SES-Between	4.863***	2.987***	7.402***	1.228***
	(0.383)	(0.095)	(0.423)	(0.104)
SES-Between Sq.	-1.837***		-6.042***	
	(0.363)		(0.401)	
SES-Within	1.605***	1.823***	7.173***	6.115***
	(0.474)	(0.466)	(0.536)	(0.527)
SES-Within Sq.	-0.607		-5.695***	
	(0.424)		(0.479)	
Between X Within		-1.548*		-8.907***
		(0.791)		(0.893)
Constant	47.334***	47.664***	47.679***	48.766***
	(0.086)	(0.055)	(0.094)	(0.061)
Individuals	42373	42373	42373	42373
Observations	208412	208412	208412	208412
Controls	YES	YES	YES	YES
Log-Likelihood	-693456.2	-693468.1	-717897.8	-718032.2

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Model (2) shows that within-individual increases in SES are strongly associated with increased physical health for low SES individuals, but not for high SES individuals. Namely, there is large and significant effect of a within-individual increase in SES for an individual with the lowest average SES rank (1.82), but this effect decreases for individuals with higher average SES (down to 0.28 for an individual with the highest average SES).

Similarly, Model (4) shows that a within-individual increase in SES has an extremely large association with an increase in mental health for low SES individuals (6.1), but we actually observe such a strong interaction effect that an increase in within-individual SES for an individual with the highest average SES is actually associated with a reduction in mental health.

5.5 Social Mobility

Table 6 shows that it is not just socioeconomic level differences between individuals what drives health inequalities, but also differences in mobility, particularly for mental health. Individuals who are, on average, highly mobile (i.e., those who expect an increase in their SES from one year to the next) have a significantly higher physical and mental health. For mental health, this association is higher than the association with the average lagged SES (4.55 and 1.13, respectively).

Table 6: SES and SF-12 Score: Lagged SES and Social Mobility

	(1) Physical SF-12 Coef./S.E.	(2) Mental SF-12 Coef./S.E.
SES_{t-1} -Between	2.969***	1.130***
	(0.097)	(0.106)
SES_{t-1} -Within	0.809***	0.456*
	(0.207)	(0.234)
SES Mobility -Between	3.719***	4.545***
	(0.499)	(0.553)
SES Mobility -Within	1.152***	1.680***
	(0.180)	(0.204)
Constant	47.663***	48.778***
	(0.057)	(0.062)
Individuals	40311	40311
Observations	198927	198927
Controls	YES	YES
Log-Likelihood	-662105.4	-685216.8

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

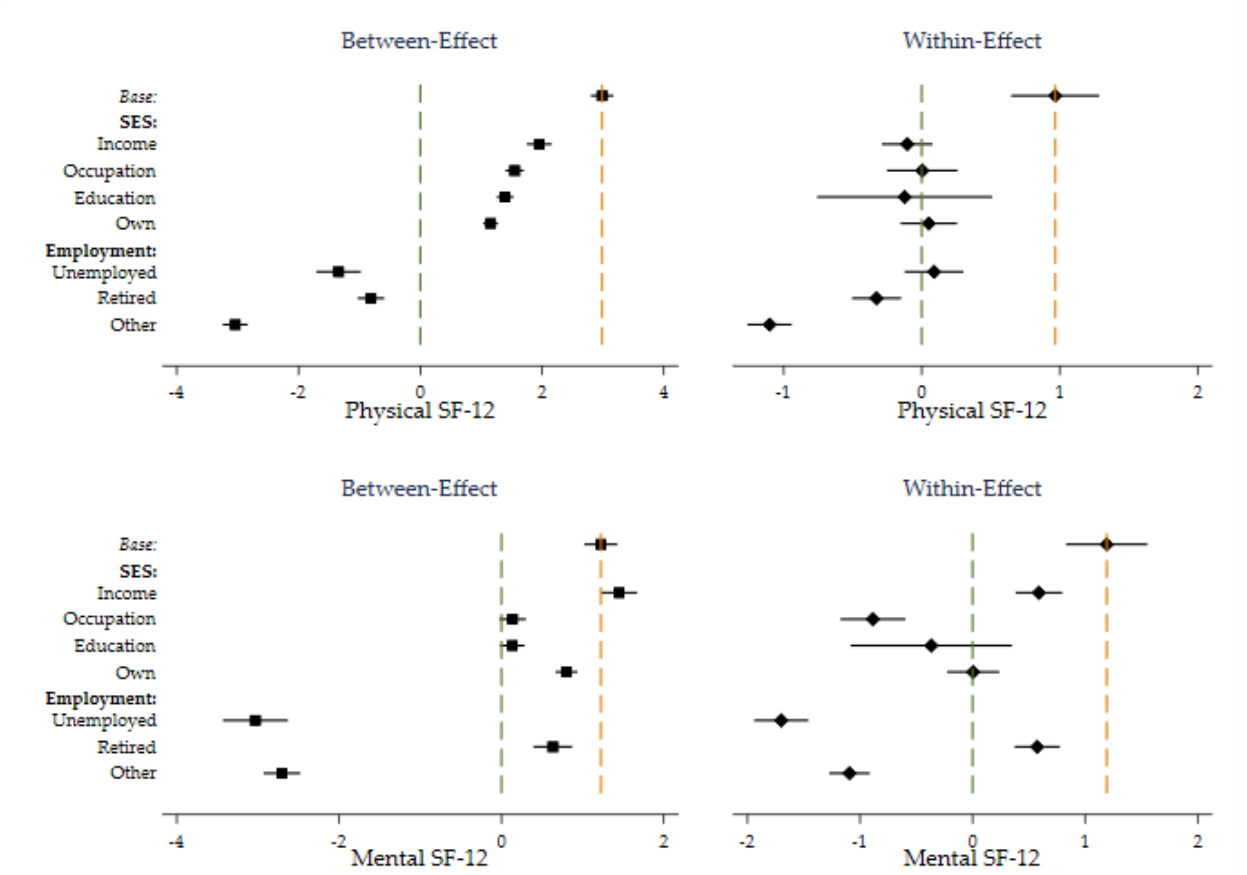
In years where the SES increase from the previous year is particularly high compared to the average mobility, this entails a large and significant effect on both physical (1.15) and mental (1.68) health.

The coefficients of the lagged SES variables shows that there are significant effects for between and within effects for both physical and mental health. The between-effects for the lagged SES are similar to the current SES effects, in Table 4, but the within-effects are smaller.

5.6 Alternative measures of SES

Figure 3 shows coefficient estimates from separate within-between regressions using alternative measures of SES: income, occupation and education ranks (within age groups), a dummy variable Own (or mortgage), and current employment status (represented as three dummy variables, with "Employed" as the reference category). See Appendix A.2 for further details.

Figure 3: Alternative SES Within-Between Forest Plot: Physical and Mental



Note: Point estimates and 95% confidence intervals are shown. The top panel shows results for Physical SF-12, whilst the bottom panel shows Mental SF-12. For SES variables each coefficient is a separate within-between regression, whilst employment coefficients are dummy variables with employed as the reference.

For physical health, the between-effects for all alternative measures of SES are positive and highly significant, but lower than the SES coefficients. Within-effects, on the other hand, are not significant for any of these variables. Therefore, those who, on average, have

higher ranked incomes, occupations, education and home ownership have higher physical health levels, but that within-individual changes in these variables are not associated with physical health. Those who spend more time (on average) unemployed, retired or other, have lower physical health. A within-individual move to retirement or other is associated with significant decreases in physical health, but changes to unemployment do not.

For mental health, whilst there are positive and significant effects of higher income rank and home ownership on mental health, no significant effects are found for higher ranked occupations or qualifications. Within-individual changes in income rank are associated with higher mental health, however, gaining qualifications or home ownership has no significant effect, and gaining a higher ranked occupation actually significantly decreases mental health. For employment status, we observe negative between and within effects from unemployment and other, but an increase in mental health for those who have retired for longer and who become retired.

5.7 Further Analysis

Appendix A.3 provides subgroup analysis, to identify heterogeneity in the results across age groups, sex and ethnicity. Between and within effects on physical health are larger for 41-65 year old's, and between-effects on mental health are larger for 22-30 years old's. Between-effects are significantly greater for men, for physical health, but greater for women for mental health. Within-effects are similar for men and women for physical health, but are higher (insignificant) for men for mental health. Larger (but insignificantly so) between-effects are found for non-white populations for both mental and physical health.

Appendix A.4 additionally shows results for alternative health variables: chronic disease prevalence and self-assessed heath. In both cases we observe similarly significant associations for both within and between SES effects.

Appendix A.5 provides sensitivity analysis where we change the sample, according to worries of attrition in the panel. The results appear robust to these alternative samples:

between and within effects are positive, highly significant and of a similar magnitude across all models.

6 Discussion

We find a steep socioeconomic gradient in physical health, and inverted-U gradient for mental health in the UK. Between-individual differences had stronger relationships with health outcomes than within-individuals variations did. This could imply that the "sticky" component of SES (e.g. childhood conditions, education) have a stronger impact on health outcomes than other "fluid" variables (e.g. current income). For instance, an individual of low SES may win the lottery and dramatically increase their income (causing a major jump in their "within" variable), but this might not cause an instant improvement in their physical health. This result is aligned to the growing body of literature suggesting that long term SES has a persistent relationship with health (Baum and Ruhm, 2009; Contoyannis and Li, 2011; Davillas and Andrew Jones, 2020). Interventions aimed at long-term life quality and financial security might be more protective of health than policies limited to provide protection against shocks.

Within-effects are also found to interact with an individual's average level of SES. For low average SES individuals, a within-individual increase in SES is associated with a large increase in physical and mental health. However, this within-effect decelerates for physical health as the average level of SES increases, and in the case of mental health, the within-effect becomes negative. It is, therefore, low SES individuals that benefit the most from changes in SES. By including social mobility in our specification (Eq. 2), we additionally found that individuals with higher expected mobility (between-effect) have significantly higher levels of both physical and mental health. Significant positive effects of a particularly steep increase in SES over the last year (within-effect), were also found, but with lower magnitudes than the between effects. In both cases, these effects are particularly strong for mental health.

Furthermore, the composite SES indicator comprises variables of livelihood, among other long-term ("sticky") and short-term ("fluid") variables. In this sense, the PCA offers a flexible and adaptive method that minimises information loss (Jolliffe and Cadima, 2016). We confirmed that the relative importance of each variable changes for different age groups. Though, overall, individuals with higher values for variables such as employment, education, savings or income, and lower values for overcrowding, have higher SES scores, the importance of each of these variables varies across different age groups. As a result, variables such as "labour market status" are less important for the eldest groups, while variables such as "savings" or "occupation" are more important for them. This provides further evidence on the ever-evolving nature of SES.

The use of different measures of SES uncovered more nuance in the relationship between socioeconomic status and health. These depend on the proxy of SES considered, on whether these are within or between differences and on the health outcome variable. Although the between individuals' results are positive and highly significant using proxies of SES, they are lower than the multidimensional SES coefficients. The within-effects are not significant for either of the proxies. Hence, using a unidimensional and time-invariant measure of SES for all ages might not capture its multi-dimensional ever-evolving nature, in line with the work of a number of authors (Kaplan, 1999; Miech and Hauser, 2001; Adler, Boyce, et al., 1994; Adler and Stewart, 2010).

We compare our findings to those of two selected studies. Hoffmann, Kröger, and Geyer (2019) analyses the relationship between health and SES using structural equation models and panel data from 10 European countries. They find SES affects health, specially in the transition from adulthood to old age. They separately use a latent variable for SES, a material wealth variable, and occupational skill level, but find no substantial differences between using either. Foverskov and Holm (2016) uses dynamic fixed-effects regression models to estimate the associations of SES and health over time in the British Household Panel Survey. They find social inequality in health using cross-sectional models, as we do. However, the authors do not find any influence of SES on health over a one to five year period, and

conclude that there is not a causal relationship between SES and health, but that health and SES are shaped by dynamics and influences in place before the respondents turn 30 years old. Though we do not claim causality, we do find a strong association between changes in SES and health across adulthood.

This study does not explain causality but analyses inference, for which the "within-between estimator" provides flexibility along with unbiased estimates (Dieleman and Templin, 2014). Although for large samples like the UKHLS a fixed effects specification is advised, this method is a better fit for our research objectives and provides with similar estimates than FE but with a reduced use of the degrees of freedom (Dieleman and Templin, 2014; Bell, Fairbrother, et al., 2019). Additionally, we partially address reverse causality by controlling for past health and first period health. A further limitation emerges, as the PCA can only be performed among individuals that have non-missing values for all the variables considered, the loss of observations is an additional limitation and a possible cause of attrition bias (A. Jones and Wildman, 2008). This methodology imposes a trade-off between the number of variables to include in the PCA and the number of observations. Healthy survivor bias is another limitation, since individuals with very poor initial health are more likely to drop out of the sample (Contoyannis, Andrew Jones, et al., 2004).

Future research could include variables on childhood circumstances in the composite index, as our study was unable to include them, due to a large number of missing observations. Moreover, given the panel nature of the UKHLS, it would be useful to perform an analysis by generations. Finally, the violin plots in Figure 2 could be the starting point for further research on the factors behind the premature health deterioration of lower SES individuals and policies for healthier aging, both physically and mentally.

7 Conclusion

This paper provides further characterization of the relationship between SES and physical and mental health. Apart from confirming the evidence of a positive association between

socioeconomic status and health, we find diminishing marginal effects of SES on physical health and an inverted-U shaped behaviour with mental health. Additionally, respondents with lower average SES are found to have higher within-individual associations with physical and mental health, and therefore benefited the most from positive changes in their SES. High mobility individuals who increase their SES from one year to the next have significantly higher physical and mental health.

Secondly, this paper demonstrates that the several dimensions comprised by SES evolve and change their relative importance through the life course, which calls for considering multi-variate time-variant measures of SES along with a longitudinal analysis of socioeconomic gradients of health. Exploring patterns across time would not only provide insightful results, but would also allow to control for health dynamics and possible lagged and cumulative effects.

Thirdly, this paper provides evidence in favour of a link between changes in SES across the life course and health outcomes. For physical and mental health outcomes, there is a significant positive association between "within" divergences from the average SES and health. This suggests that the multidimensional SES index can pick up changes "within" individuals that other SES proxies (e.g.income, education, occupation) cannot, and so time series associations are stronger with the multidimensional SES rank.

Interventions aimed at long-term aspects of life quality and financial security could potentially be more protective of health than policies limited to provide protection against shocks, and might hide the key to healthier aging.

References

Adler, Nancy E., T. Boyce, et al. (1994). "Socioeconomic status and health. The challenge of the gradient." eng. In: *The American psychologist* 49(1), pp. 15–24. DOI: 10.1037//0003-066x.49.1.15.

Adler, Nancy E. and Judith Stewart (2010). "Health disparities across the lifespan: Meaning, methods, and mechanisms". In: *Annals of the New York Academy of Sciences* 1186, pp. 5–23. DOI: 10.1111/j.1749-6632.2009.05337.x.

Apouey, Bénédicte and Pierre Geoffard (2013). "Family income and child health in the UK". In: *Journal of Health Economics* 32(4), pp. 715–727.

Baum, Charles L. and Christopher J. Ruhm (2009). "Age, socioeconomic status and obesity growth". In: *Journal of Health Economics* 28(3), pp. 635–648. DOI: https://doi.org/10.1016/j.jhealeco.2009.01.004.

Bell, Andrew, Malcolm Fairbrother, and Kelvyn Jones (2019). "Fixed and random effects models: making an informed choice". In: *Quality and Quantity* 53(2), pp. 1051–1074. DOI: 10.1007/s11135-018-0802-x.

Bell, Andrew and Kelvyn Jones (2015). "Explaining Fixed Effects: Random Effects modelling of Time-Series Cross-Sectional and Panel Data". English. In: *Political Science Research and Methods* 3(1), pp. 133–153. DOI: 10.1017/psrm.2014.7.

Biressi, Anita and Heather Nunn (2013). *Class and Contemporary British Culture:* Palgrave Macmillan.

Case, Anne, Darren Lubotsky, and Christina Paxson (2002). "Economic Status and Health in Childhood: The Origins of the Gradient". In: *American Economic Review* 92(5), pp. 1308–1334. DOI: 10.1257/000282802762024520.

Chadwick, Edwin (1842). *The Report from the Poor Law Commissioners on an Inquiry into the Sanitary Conditions of the Labouring Population of Great Britain.*

Contoyannis, Paul, Andrew Jones, and Nigel Rice (2004). "The dynamics of health in the british household panel survey". In: *Journal of Applied Econometrics* 19(4), pp. 473–503. DOI: 10.1002/jae.755.

Contoyannis, Paul and Jinhu Li (2011). "The evolution of health outcomes from childhood to adolescence". In: *Journal of Health Economics* 30(1), pp. 11–32. DOI: https://doi.org/10.1016/j.jhealeco.2010.10.007.

Cutler, David M. and Adriana Lleras-Muney (2010). "Understanding differences in health behaviors by education". In: *Journal of Health Economics* 29(1), pp. 1–28. DOI: 10.1016/j.jhealeco.2009.10.003.

Cutler, David M., Adriana Lleras-Muney, and Tom Vogl (2012). *Socioeconomic Status and Health : Dimensions and Mechanisms Socioeconomic Status and Health : Dimensions and Mechanisms*. June 2020, pp. 1–49. DOI: 10.1093/oxfordhb/9780199238828.013.0007.

Davillas, Apostolos and Andrew Jones (2020). "Ex ante inequality of opportunity in health, decomposition and distributional analysis of biomarkers". In: *Journal of Health Economics* 69, p. 102251. DOI: 10.1016/j.jhealeco.2019.102251.

Davillas, Apostolos, Andrew Jones, and Michaela Benzeval (2019). "Chapter 22 - The Income-Health Gradient: Evidence From Self-Reported Health and Biomarkers in Understanding Society". In: *Panel Data Econometrics Empirical Applications*. Ed. by Mike Tsionas. Academic Press, pp. 709–741. DOI: https://doi.org/10.1016/B978-0-12-815859-3.00022-6.

Department of Health and Social Security (1980). *Inequalities in health: Report of a research working group*.

Dieleman, Joseph L. and Tara Templin (2014). "Random-effects, fixed-effects and the within-between specification for clustered data in observational health studies: A simulation study". In: *PLoS ONE* 9(10). DOI: 10.1371/journal.pone.0110257.

Dutton, D.B. and S Levine (1989). "Socioeconomic status and health: Overview, methodological critique, and reformulation". In: *Pathways to Health*. Ed. by JP Bunker, DS Gomby, and BH Kehrer. The Henry J. Kaiser Family Foundation: Menlo Park, CA, pp. 29–69.

Elo, Irma T (2009). "Social Class Differentials in Health and Mortality: Patterns and Explanations in Comparative Perspective". In: *Annual Review of Sociology* 35(1), pp. 553–572. DOI: 10.1146/annurev-soc-070308-115929.

Ettner, Susan L (1996). "New evidence on the relationship between income and health". In: *Journal of Health Economics* 15(1), pp. 67–85. DOI: https://doi.org/10.1016/0167-6296(95)00032-1.

Eyal, Nir (2018). "Inequality in political philosophy and in epidemiology: A remarriage". In: *Journal of Applied Philosophy* 35(1), pp. 149–167. DOI: 10.1111/japp.12150.

Feinstein, Jonathan S (1993). "The Relationship between Socioeconomic Status and Health: A Review of the Literature". In: *The Milbank Quarterly* 71(2), pp. 279–322. DOI: 10.2307/3350401.

Fleishman, John A, Alfredo J Selim, and Lewis E Kazis (2010). "Deriving SF-12v2 Physical and Mental Health Summary Scores: A Comparison of Different Scoring Algorithms". In: *Quality of Life Research* 19(2), pp. 231–241.

Foverskov, Else and Anders Holm (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." In: *Social Science & Medicine (1982)* 150, pp. 172–183. DOI: 10.1016/j.socscimed.2015.12.021.

Frijters, Paul, John P. Haisken-DeNew, and Michael A. Shields (2005). "The causal effect of income on health: Evidence from German reunification". In: *Journal of Health Economics* 24(5), pp. 997–1017. DOI: https://doi.org/10.1016/j.jhealeco.2005.01.004.

Galama, Titus J and Hans van Kippersluis (2018). "A Theory of Socio-economic Disparities in Health over the Life Cycle". In: *The Economic Journal* 129(617), pp. 338–374. DOI: 10.1111/ecoj.12577.

Galobardes, Bruna et al. (2006). "Indicators of socioeconomic position (part 1)". eng. In: *Journal of epidemiology and community health* 60(1), pp. 7–12. DOI: 10.1136/jech.2004.023531.

Hauck, Katharina and Nigel Rice (2004). "A longitudinal analysis of mental health mobility in Britain". In: *Health Economics* 13(10), pp. 981–1001. DOI: https://doi.org/10.1002/hec.939.

Hoffmann, R., H. Kröger, and S. Geyer (2019). "Social Causation Versus Health Selection in the Life Course: Does Their Relative Importance Differ by Dimension of SES?" In: *Social Indicators Research* 141(3), pp. 1341–1367. DOI: 10.1007/s11205-018-1871-x.

Hoffmann, R., H. Kröger, and E. Pakpahan (2018). "Pathways between socioeconomic status and health: Does health selection or social causation dominate in Europe?" In: *Advances in Life Course Research* 36(April 2017), pp. 23–36. DOI: 10.1016/j.alcr.2018.02.002.

Jolliffe, Ian T. and Jorge Cadima (2016). "Principal component analysis: A review and recent developments". In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 374(2065). DOI: 10.1098/rsta.2015.0202.

Jones, A. and J. Wildman (2008). "Health, income and relative deprivation: Evidence from the BHPS". In: *Journal of Health Economics* 27(2), pp. 308–324. DOI: https://doi.org/10.1016/j.jhealeco.2007.05.007.

Kaplan, George A (1999). "Part III Summary". In: *Annals of the New York Academy of Sciences* 896(1), pp. 116–119. DOI: 10.1111/j.1749-6632.1999.tb08109.x.

Lee, Dohoon and Margot Jackson (2017). "The Simultaneous Effects of Socioeconomic Disadvantage and Child Health on Children's Cognitive Development". In: *Demography* 54(5), pp. 1845–1871. DOI: 10.1007/s13524-017-0605-z.

Marmot, M.G., G. Rose, et al. (1978). "Employment grade and coronary heart disease in British civil servants". In: *Journal of Epidemiology and Community Health* 32(4), pp. 244–249. DOI: 10.1136/jech.32.4.244.

Marmot, M.G., M.J. Shipley, and G. Rose (1984). "Inequalities in death–specific explanations of a general pattern?" In: *Lancet* 1(8384), pp. 1003–6. DOI: 10.1016/s0140-6736(84)92337-7.

Miech, Richard A and Robert M Hauser (2001). "Socioeconomic Status and Health at Midlife: A Comparison of Educational Attainment with Occupation-Based Indicators". In: *Annals of Epidemiology* 11(2), pp. 75–84. DOI: https://doi.org/10.1016/S1047-2797(00)00079-X.

Mosby (2013). *Mosby's dictionary of medicine, nursing & health professions*. English. St. Louis, Mo.

Mundlak, Yair (1978). "On the pooling of time series and cross section data". In: *Econometrica: Journal of the Econometric Society*, pp. 69–85.

Nazroo, James Y (2003). "The structuring of ethnic inequalities in health: economic position, racial discrimination, and racism". eng. In: *American Journal of Public Health* 93(2), pp. 277–284. DOI: 10.2105/ajph.93.2.277.

Pappas, G et al. (1993). "The increasing disparity in mortality between socioeconomic groups in the United States, 1960 and 1986." eng. In: *The New England Journal of Medicine* 329(2), pp. 103–109. DOI: 10.1056/NEJM199307083290207.

Perales, Francisco Paco and Reinhard Schunck (2020). "XTHYBRID: Stata module to estimate hybrid and correlated random effect (Mundlak) models within the framework of generalized linear mixed models (GLMM)". In: *Statistical Software Components*.

Savage, Mike (2015). *Social Class in the 21st century*. Pelican Books.

Savage, Mike et al. (2013). "A New Model of Social Class? Findings from the BBC's Great British Class Survey Experiment". In: *Sociology* 47(2), pp. 219–250. DOI: 10.1177/0038038513481128.

Stockwell, Edward G., Evelyn M. Kitagawa, and Philip M. Hauser (1975). *Differential Mortality in the United States: A Study in Socioeconomic Epidemiology*. DOI: 10.2307/2062204.

Vilagut, Gemma et al. (2013). "The Mental Component of the Short-Form 12 Health Survey (SF-12) as a Measure of Depressive Disorders in the General Population: Results with Three Alternative Scoring Methods". In: *Value in Health* 16(4), pp. 564–573. DOI: https://doi.org/10.1016/j.jval.2013.01.006.

Vyas, Seema and Lilani Kumaranayake (2006). "Constructing socio-economic status indices: How to use principal components analysis". In: *Health Policy and Planning* 21(6), pp. 459–468. DOI: 10.1093/heapol/czl029.

Wiggins, R.D. et al. (2004). "Social position and minor psychiatric morbidity over time in the British Household Panel Survey 1991-1998". In: *Journal of Epidemiology and Community Health* 58(9), pp. 779–787. DOI: 10.1136/jech.2003.015958.

A Appendix

A.1 Within and Between Descriptive Statistics

Descriptive statistics in Table 7 show the mean, standard deviation and 5 to 95 percentile range of our health and SES variables. Overall variation is greater in physical health compared to mental health; this is primarily driven by between-individuals differences. However, there is higher within-individual variation for mental health. For SES rank the majority of the variation is due to between individual differences, we do observe substantial variation in within-individual differences, with those at the 95th percentile having an increase in their rank by 13. Overall social mobility is positive, showing a (small) increase in SES over time. Importantly, we observe variation in expected social mobility, with individuals in the 95th percentile increasing their rank by 6 each year, compared to those in the 5th percentile with a decrease of 4. Within-individual variation is greater: indicating that social mobility is not smooth, individuals rise and fall on their way to the top (or bottom).

Table 7: Within and Between Descriptive Statistics

	Mean	S.D.	P5	P95
Physical SF-12				
-Overall	49.25	11.26	24.38	61.04
-Between	49.25	9.90	26.78	59.21
-Within	0.00	5.37	-9.54	8.17
Mental SF-12				
-Overall	49.42	9.88	30.11	61.56
-Between	49.42	7.84	34.37	59.11
-Within	-0.00	6.02	-10.70	9.17
SES Rank				
-Overall	0.53	0.28	0.07	0.96
-Between	0.53	0.27	0.08	0.93
-Within	0.00	0.08	-0.13	0.13
SES Lagged Diff.				
-Overall	0.01	0.10	-0.14	0.16
-Between	0.01	0.04	-0.04	0.06
-Within	0.00	0.09	-0.14	0.14

N=208,412

A.2 Alternative Measures of SES

Tables 8, 9 and 9 show results from the within-between regressions for alternative measures of socioeconomic status. Income, occupation and education variables are constructed as normalised rank, conditional on age group.[5] This provides a comparable metric as our SES variable. Home ownership is included as a dummy. Results for these variables are shown in Table 8 for physical health and Table 9 for mental health.

Table 8: SES and Physical SF-12: Alternatives to SES

	(1) SES Coef./S.E.	(2) Income Coef./S.E.	(3) Occupation Coef./S.E.	(4) Education Coef./S.E.	(5) Own Coef./S.E.
Between-Effect	2.987***	1.954***	1.550***	1.393***	1.155***
	(0.095)	(0.104)	(0.077)	(0.070)	(0.063)
Within-Effect	0.968***	-0.105	0.004	-0.124	0.051
	(0.163)	(0.094)	(0.130)	(0.323)	(0.105)
Constant	47.664***	48.205***	48.239***	48.366***	48.357***
	(0.055)	(0.060)	(0.055)	(0.050)	(0.053)
Individuals	42373	42373	42373	42373	42373
Observations	208412	208412	208412	208412	208412
Controls	YES	YES	YES	YES	YES
Log-Likelihood	-693470.0	-693806.0	-693780.9	-693787.8	-693814.6

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 10 shows results for current employment status. Here three dummy variables: Unemployed, Retired and Other are used, with Employed as the reference level.

[5]The ranking of occupation and education variables is a rough, parsimonious, approximation. NS-SEC categories were not designed with the intention of a particular ranking, nor are the qualifications directly comparable in terms of rank. However, by ranking them as such we allow for a comparison with income and our SES variable, and our results do not differ much when they are included as separate dummy variables.

Table 9: SES and Mental SF-12: Alternatives to SES

	(1) SES Coef./S.E.	(2) Income Coef./S.E.	(3) Occupation Coef./S.E.	(4) Education Coef./S.E.	(5) Own Coef./S.E.
Between-Effect	1.225***	1.450***	0.137	0.134*	0.802***
	(0.104)	(0.114)	(0.084)	(0.077)	(0.069)
Within-Effect	1.192***	0.590***	-0.886***	-0.367	0.005
	(0.184)	(0.106)	(0.147)	(0.365)	(0.118)
Constant	48.767***	48.651***	49.321***	49.325***	48.805***
	(0.061)	(0.065)	(0.060)	(0.055)	(0.058)
Individuals	42373	42373	42373	42373	42373
Observations	208412	208412	208412	208412	208412
Controls	YES	YES	YES	YES	YES
Log-Likelihood	-718081.8	-718074.6	-718152.3	-718169.7	-718103.5

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 10: SES and Mental SF-12: Alternatives to SES

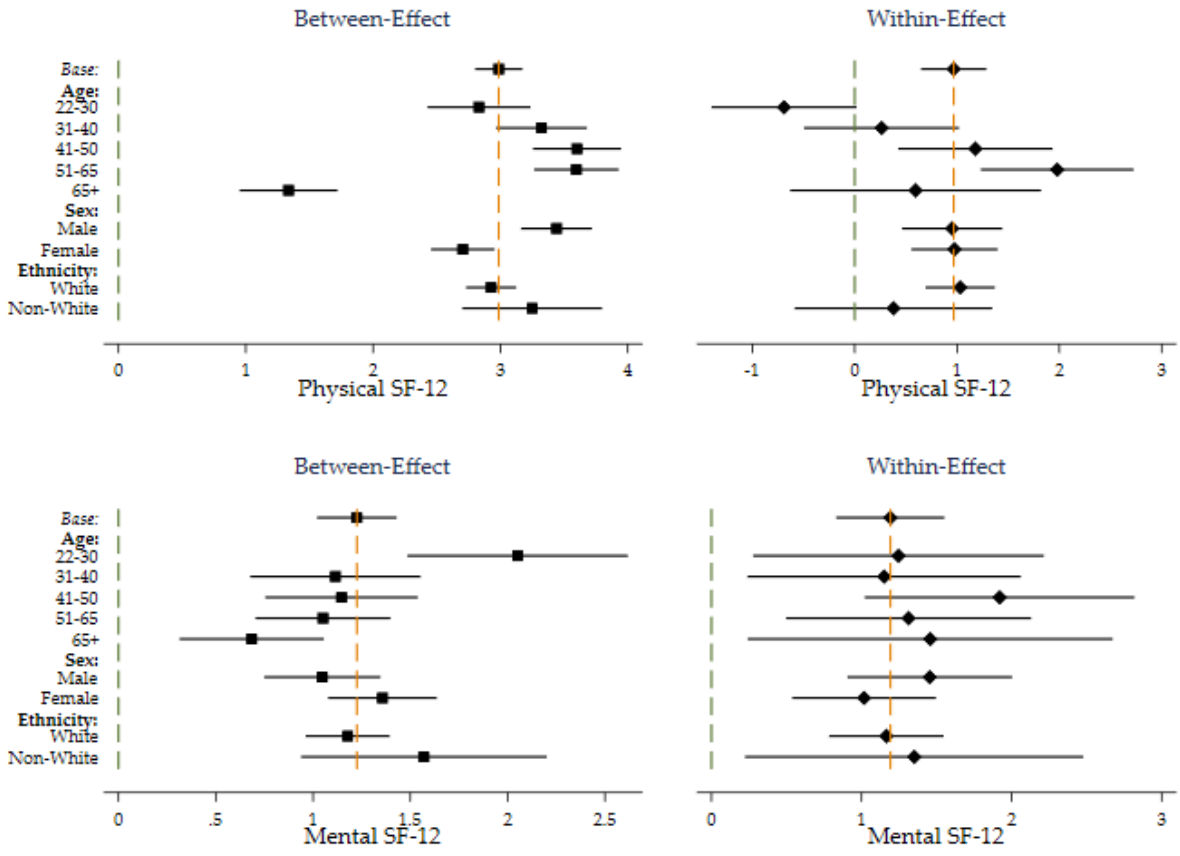
	(1) Physical SF-12 Coef./S.E.	(2) Mental SF-12 Coef./S.E.
Between-Unemployed	-1.342***	-3.031***
	(0.186)	(0.204)
Between-Retired	-0.810***	0.634***
	(0.112)	(0.122)
Between-Other	-3.041***	-2.704***
	(0.105)	(0.115)
Within-Unemployed	0.089	-1.700***
	(0.109)	(0.123)
Within-Retired	-0.327***	0.574***
	(0.091)	(0.102)
Within-Other	-1.103***	-1.093***
	(0.082)	(0.092)
Constant	49.781***	49.607***
	(0.041)	(0.045)
Individuals	42373	42373
Observations	208412	208412
Controls	YES	YES
Log-Likelihood	-693434.0	-717581.0

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A.3 Subgroup Analysis

To identify heterogeneity in results, subgroup analysis is performed across age groups, sex and ethnicity. Results are summarised in Figure 4. For physical health, in the top panel, results show that between-effects are positive and highly significant across all age groups, ranging from 1.336 to 3.602. The largest coefficients are for the mid age group 41-50 and 51-65, whilst the lowest are amongst the over 65s. Within-effects are less robust, with only the 41-50 and 51-65 groups having positive and significant within-effects, 1.179 and 1.979 respectively. For mental health, we observe positive and significant coefficients across all age groups. However, we observe an age related trend in between-effects, with the 22-30 years with coefficients of 2.051 down to 0.685 for the over 65s. There are no significant differences across age groups for within-effects.

Figure 4: Subgroup Within-Between Forest Plot: Physical and Mental



When splitting the sample by sex, we observe highly significant and positive coefficients for both between and within effects, across physical and mental health. Between effects are higher for males (3.44) for physical health compared to females (2.70), but there is little difference for within-effects. For mental health, females (1.36) have a slightly higher between-effect than men (1.05), while men (1.46) have a bigger within-effect than women (1.02). Indeed, the finding of larger between-effects compared to within-effects is reversed for men for mental health.

Finally, we divide the sample by ethnicity: white and non-white. For physical health, we observe larger between effects for non-white populations (3.25) compared to white populations (2.92), and while within-effects are still significant for the white subgroup (1.03) the coefficient is insignificant for the non-white subgroup (0.38). For mental health, coefficients are more similar, but with the non-white population seeing higher between and within coefficients.

A.4 Alternative Measures of Health

While the SF-12 provides a measures of physical and mental health, alternative health variables are available within the BHPS-UKHLS survey. Here two further variables: Disease Prevalence[6] and Self-assessed health[7] are further analysed. The former is an objective measure of long-term health conditions, available in the UKHLS waves, while the latter is an subjective measure of perceived health status, available in BHPS.

Table 11: SES and Alternative Health Measures

	(1) Disease Prevalence Coef./S.E.	(2) Self-Assessed Health Coef./S.E.
Between-Effect	-0.075***	0.210***
	(0.003)	(0.007)
Within-Effect	-0.032***	0.069***
	(0.007)	(0.010)
Constant	0.148***	0.583***
	(0.001)	(0.004)
Individuals	51350	15706
Observations	260440	103381
Controls	YES	YES
Log-Likelihood	-57794.7	-44092.8

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

We run sensitivity analysis, using our base within-between specification, adapted for these alternative health variables. Table 11 shows these results. As, with SF-12 Physical and Mental health we observe that both between and within effects are highly significant. We observe lower rates of disease prevalence and higher levels of self-assessed health amongst individuals with higher average SES. Similar differences, with lesser magnitudes are found for within-individual differences in SES.

[6] *Disease prevalence* is a binary variable indicating whether or not an individual has one or more long-term condition. In the survey, respondents were asked if a doctor or other health professional have ever diagnosed them with any of the health conditions from a list. The list includes: asthma, arthritis, congestive heart failure, coronary heart disease, angina, heart attack or myocardial infarction, stroke, emphysema, hyperthyroidism or over-active thyroid, hypothyroidism or under-active thyroid, chronic bronchitis, any kind of liver condition, cancer or malignancy, diabetes, epilepsy, high blood pressure, clinical depression, other long-standing/chronic condition, multiple sclerosis and HIV.

[7] *Self-assessed Health* (SAH) refers to the perceived health status of an individual compared to people of their own age. This variable was dichotomised by assigning a value of one to those that consider themselves to be in good to excellent health, and zero otherwise. Self-assessed health is a widely used predictor of subsequent use of medical care (Contoyannis, Andrew Jones, et al., 2004).

A.5 Sample Sensitivity

To evaluate potential biases in our results from sample attrition we run our within-between model across two alternative samples. The full sample is compared to one which excludes individuals not observed in the first wave of UKHLS (Wave 1) and where only individuals who were present in all 10 waves of UKHLS (Balanced) are compared.

Table 12: SES and SF-12 Score: Sample Sensitivity

	Physical SF-12			Mental SF-12		
	(1) Full Coef./S.E.	(2) Wave 1 Coef./S.E.	(3) Balanced Coef./S.E.	(4) Full Coef./S.E.	(5) Wave 1 Coef./S.E.	(6) Balanced Coef./S.E.
Between-Effect	2.987***	3.145***	3.305***	1.225***	1.240***	0.989***
	(0.095)	(0.111)	(0.161)	(0.104)	(0.121)	(0.170)
Within-Effect	0.968***	0.961***	1.072***	1.192***	1.082***	0.733***
	(0.163)	(0.188)	(0.230)	(0.184)	(0.210)	(0.255)
Constant	47.664***	47.460***	47.594***	48.767***	48.875***	49.370***
	(0.055)	(0.065)	(0.098)	(0.061)	(0.071)	(0.103)
Individuals	42373	31127	12239	42373	31127	12239
Observations	208412	157654	93551	208412	157654	93551
Controls	YES	YES	YES	YES	YES	YES
Log-Likelihood	-693470.0	-525802.0	-308740.5	-718081.8	-542731.1	-318136.2

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$