

The association between socioeconomic status and C-reactive protein in Bayesian perspective

Alessandro Procopio^{a,b,*}, Robin Samuel^{b,1}

^a Stockholm University, Sweden

^b University of Luxembourg, Luxembourg

1. Introduction

The graded association between socioeconomic status (SES) and mortality risks is an established finding in the biomedical and social scientific literature (Adler & Ostrove, 1999; Brummett et al., 2014; Harris & Schorpp, 2018; House, 2002; Winkleby et al., 1992; Yang et al., 2020). While there is a broad literature that links individuals' social conditions and health outcomes (Adler et al., 1994; Adler & Ostrove, 1999; Clark et al., 2009; Elo, 2009; Ferraro & Shippee, 2009; Marmot et al., 1991; Pudrovska, 2014), the focus on mechanisms driving this relationship is more recent (Elo, 2009; Freese, 2018; Harris & Schorpp, 2018). Particularly, social scientists and epidemiologists have been focusing on the mediatory role of chronic illness and inflammation on the association between individuals' SES and health disparities (Baum et al., 1999; Dowd & Zajacova, 2007; Pudrovska, 2014) in their life course (Ben-Shlomo & Kuh, 2002; Corna, 2013; Ferraro & Shippee, 2009; Hallqvist et al., 2004; Liu et al., 2017; Power et al., 1999; Yang et al., 2020). Although previous studies have focused on measuring the social group differences in health outcomes (Kawachi et al., 2002), little is known regarding the differences among individuals belonging to those social groups. Individual variation within social groups has been significantly understudied compared to the between-group analysis. Within the World Health Organization (WHO), public health scientists have engaged in a heated debate concerning the use of an individual-level approach (Gakidou et al., 2000; Murray et al., 1999), or a group-level comparison strategy (Braveman et al., 2000). After the 2000s, social scientific and public health literature primarily focused on "between group" comparisons (across countries, ethnicities, gender, or social conditions). The within-group aspect of health distribution regained interest among social and public health scientists in the total health inequality theory context (Gakidou & King, 2002). Particularly, the total health inequality theory reframes the previous debate within the WHO where the within and between group comparisons are complementing – not competing – components of inequality. This study

investigates how individuals' SES shapes the distribution of C-reactive protein (CRP). We address the following research questions: How does individuals' SES shape the distribution of C-reactive protein (CRP)? How is the CRP distribution shaped within SES categories? Our main aim is to explore these questions without testing specific hypotheses. CRP is an acute-phase protein produced by hepatocytes as a response of the immune system to acute infection or systemic inflammation. Biomedical studies linked CRP to mortality risks due to cardiovascular diseases (CVDs, Alley et al., 2006; Harris et al., 1999; Laaksonen et al., 2005; McDade et al., 2011). CVDs are one of the leading causes of mortality and morbidity in high-income countries (Brummett et al., 2014; Liu et al., 2017; Mitchell & Aneshensel, 2017). Previous studies have found evidence concerning the socially patterned onset of CVDs, where individuals from a lower SES encountered higher risks of mortality due to CVDs (Goodman et al., 2005; Lubbock et al., 2005; Winkleby et al., 1992), confirming the social causation theory (Link & Phelan, 1995; Phelan et al., 2010). Additionally, CVDs cause a considerable burden on individual and public health. Therefore, biomedical and social science research has striven to understand and identify the principal physiological changes that could signal the onset of CVDs in individuals at the pre-symptom stage of the disease (Davillas et al., 2019; Dowd & Zajacova, 2007; McEwen, 2015; Mitchell & Aneshensel, 2017). Prevention and identification of the population at significant risk of CVDs have been a primary objective for social scientists, public health policy-makers, and epidemiologists (Davillas et al., 2017; Herd et al., 2007; Mitchell & Aneshensel, 2017). In the field of social inequalities in health, accumulating evidence highlights the association between CRP and individuals' social conditions, where individuals in higher SES position have lower levels of CRP, and likely, it might lower the risks of CVDs onset (Brummett et al., 2014; Davillas et al., 2019; Gimeno et al., 2008; Jousilahti et al., 2003; Karimi et al., 2019; Koster et al., 2006; Lubbock et al., 2005; McDade et al., 2011). Empirical evidence has highlighted the importance of health-related behaviors, such as smoking, sedentary lifestyle, and alcohol consumption, as mediating factors in the social

* Corresponding author. Stockholm University, Department of Public Health Sciences, Albanovägen 12, 114 19, Stockholm, Sweden.
E-mail address: alessandro.procopio@su.se (A. Procopio).

¹ University of Luxembourg, Department of Social Sciences, 11 Porte des Sciences, 4366 Esch-sur-Alzette, Luxembourg.

gradient of CRP. In particular, previous studies have shown that individuals in lower social conditions are more prone to engaging in health-risky behaviors, thereby linking SES to CRP levels (Alley et al., 2006; Yang et al., 2020). Additionally, recent studies have demonstrated empirically an interplay between race and SES (Farmer et al., 2022; Farmer et al., 2020; Farmer, Wray, & Haas, 2021).

Our study contributes in two ways to this emergent research strand by first investigating how the social determinants of health get under the skin, addressing the mechanism linking individuals' SES with CRP levels. Second, this study proposes the Bayesian paradigm as an alternative methodological framework to measure the total inequality in CRP, which is our individuals' health indicator. We employed two types of Bayesian Regression Model (BRM) to enhance our understanding of the social determinants of health. Using the BRM, our focus of the empirical analysis shifts from the conventional point estimates (and their statistical significance) to a distribution of likely parameters, allowing the detection of a more comprehensive understanding of the potential impacts that individuals' social conditions exert on their health and well-being. We expanded the analysis range by exploring the intra-group differences in CRP levels between individuals at the same SES level. Consequently, we modeled the standard deviation of CRP distribution as a function of individuals' SES through a Bayesian distributional model (Umlauf & Kneib, 2018). Our study investigates the mediating role of health-related behaviors, focusing on smoking status, physical activity, and Body Mass Index (BMI), and their risk factors. We used data from the wave 2012 (Health Assessment Panel) of the United Kingdom Household Longitudinal Study (UKHLS), a nationally representative longitudinal survey set in the United Kingdom².

To consider the multidimensional characteristics of SES, we included three measures of individuals' social conditions in the empirical analysis: occupational status, educational level, and household income (see Elo, 2009; Goldthorpe, 2010 for a methodological discussion).

2. Data & methods

2.1. Data

The data we used for the empirical analysis comes from the UK Household Longitudinal Study (UKHLS), wave 2 (Health Assessment Panel, 2012). The Understanding Society – UKHLS is a large and representative survey of households sampled in the United Kingdom (UK), Scotland, Wales, and Northern Ireland. Administration of the main survey interested the General Population survey (GPS)³.

The sampling procedure for the GPS consisted of a two-stage step: the first primary sampling unit (PSU) consisted of a sample of postcode sectors, within which the addresses were the sampling units. The UKHLS provides a multi-purpose questionnaire to the respondents, which covers various topics relevant to social research. In 2010 and 2012, alongside the main questionnaire, the survey design included questions on health, and it collected blood samples from the respondents who consented to participate. The eligibility criteria for the respondents to participate in the nurse health assessment included completing the face-to-face interview; being 16 or older; living in England, Scotland, or Wales; completing the questionnaire in English, and not being pregnant (McFall et al., 2012). Individuals with HIV, hepatitis A or B, and clotting or bleeding disorders were excluded. The aim of collecting biospecimens by registered nurses was to gather information on potential health risks. Furthermore, blood sample collection supports genetic analyses and creates a genetic database. The nurse health assessment interested a

subsample of the GPS and included anthropometric measures (such as height, weight, percent body fat, and waist circumference), blood pressure, grip strength, lung function, and blood samples. The basis of the biomarkers selection from the blood samples considers the environmental effect (socioeconomic, physical, or psychosocial), the impact on the biospecimen, its importance to essential health conditions, and the proportion of the population affected by the disease. From the 9896 observations of the initial sample, 521 individuals recorded a missing value of CRP. Thus, we excluded those cases from the study. Individuals with inapplicable values on height and weight measures to calculate individuals' BMI ($N = 283$) were set as missing values and excluded from the statistical analysis. After deleting missing values on the covariates, the size of our analytical sample was $N = 8,960$ ⁴. We conducted a subgroup analysis by testing the difference in observable values of the covariates included in the model. Appendix B provides a summary statistics table on the characteristics of the sample having missing values and results of statistical significance tests between the analytical sample and the missing cases on covariates in models 1 and 2.

2.2. Measures

2.2.1. Dependent variable

The dependent variable of the two BRMs is the recorded CRP in the Wave 2012 of the Understanding Society - UKHLS, measured in mg/L. We considered high-risk CVDs individuals with a CRP level greater than 3 mg/L, which is the cut-off point for defining an individual affected by low-grade inflammation (Brummett et al., 2014; McDade et al., 2011). As the distribution of the recorded values among the individuals in the sample is highly skewed, we log-transformed the variable to normalize the distribution.

2.2.2. Covariates

2.2.2.1. SES measures: To capture the multidimensionality of SES (Elo, 2009; Goldthorpe, 2010), we included in the statistical model three measures of SES: occupational status, education, and income. Occupational status comprises eight categories from the National Statistics Socioeconomic Classification (NS-SEC): large employers & higher management, higher professional, lower management & professional, intermediate, small employers & own account, lower supervisory & technical, semi-routine, routine, and not in LM. The last category includes retired individuals, students, and individuals that are not currently working. The main reason is to compare individuals who have a job (and within those, compare the occupational categories) with individuals who do not currently have a job.

Educational level was measured using five categories: degree, other higher degree, A-level, GCSE, other qualifications, and no qualification.

We have taken the gross household income that was registered the month before the interview concerning the third SES measure. Then, we equalized the scale, dividing the income scale by the equivalence scale set by the OECD, returning the equivalized income scale for the number of household members.

2.2.2.2. Mediating factors: The health behavior of individuals that might influence CRP levels (Dowd et al., 2009; Yang et al., 2020) has been measured by taking into account the level of sports activity (scale from 0 “no activity” to 10 “very active”), the Body Mass Index (BMI) calculated as $\text{weight}/(\text{height}/100)^2$, smoking behavior (current smoker,

² The UKHLS is a continuation and further evolution of the older British Household Panel Study (BHPS).

³ Alongside the GPS, the main survey target sample consisted of three additional components: the Ethnic Minority Boost sample, the former BHPS sample, and the Immigrant and Ethnic Minority Boost sample.

⁴ We conducted a subgroup analysis by testing the difference in observable values of the covariates included in the model. Appendix B provides a summary statistics table on the characteristics of the sample having missing values and results of statistical significance tests of difference between the analytical sample and the individuals with missing values on covariates included in models 1 and 2.

ex-smoker and non-smoker), and self-rated health (as it is associated with CRP levels, see [Shanahan et al., 2014](#)), which have been coded into five categories: excellent, very good, good, fair, and poor health.

2.2.2.3. Controllars: The covariates we included to control the relationship between the SES measures and the CRP levels concern the individuals' sociodemographic characteristics. Among the sociodemographic variables, for the analysis, we considered the age (measured as a continuous scale from 20 to 65 years old) of the individuals, gender, and house ownership (see [McDade et al., 2011](#)) as indicator of wealth.

2.3. Descriptive statistics

[Table 1](#) provides an overview of the descriptive statistics (mean and standard deviations) of the covariates included in the models, while [Table 2](#) provides the correlation matrix among the variables included in the analysis of Understanding Society data.

2.4. Statistical analysis

The empirical analysis provided two Bayesian regression models to draw a posterior distribution from all the possible magnitude effects through Markov Chain Monte Carlo (MCMC) algorithm ([Lynch & Bartlett, 2019](#)). In the context of the Bayesian linear regression model, the posterior distribution of the parameters β_0 (the intercept), β_1 (the slope

of the regression lines), and σ (the standard deviation) are:

$$p(\beta_0, \beta_1, \sigma | D) = \frac{p(D | \beta_0, \beta_1, \sigma) p(\beta_0, \beta_1, \sigma)}{\int \int \int p(D | \beta_0, \beta_1, \sigma) p(\beta_0, \beta_1, \sigma) d\beta_0 d\beta_1 d\sigma} \quad (1)$$

The Bayesian framework aims to infer a likely distribution of the determined parameters β_0 , β_1 , and σ given the observed data D . The prior terms $p(\beta_0, \beta_1, \sigma)$ at the denominator indicate the distribution of credibilities that the specific parameter could take without considering the observed data. The marginal likelihood $p(D | \beta_0, \beta_1, \sigma) p(\beta_0, \beta_1, \sigma)$ informs us about the overall probability of the data $p(D | \theta)$, weighting these probabilities by the strength of their prior likelihood. All models conceive a hierarchical structure of the statistical analysis. That means, within the Bayesian framework, the first step is to apply a determined distribution that could best fit the dependent variable⁵.

All two models apply to the (log) CRP distribution, a t-student distribution. The choice of the t-student concerns the potential outliers present in the observed data, thus providing a more robust model.

Models 1 and 2 set the mean of the dependent variable as a linear function of the aforementioned covariates. The models exploit the hierarchical features of the Bayesian framework by calculating the deviations from the mean (as population-level effects) of the groups outlined by the two categorical variables representing individuals' SES: occupation and education. The specification of Models 1 and 2 take the following form:

$$\begin{aligned} \mu_y &= \beta_0 + \sum_i \beta_i X_i + \beta_j X_j p(\beta_0) = \mathcal{N}(\mu_y, \sigma_y) p(\sigma_{\beta_i}) = \mathcal{C}(0, \sigma_y) p\left(\beta_j\right) \\ &= \mathcal{N}\left(0, 1\right) p\left(\beta_i\right) = \mathcal{N}\left(0, 1\right) p\left(\sigma\right) = \mathcal{N}\left(0, \sigma_y\right) p\left(\nu\right) \\ &= e\left(1 / 29\right) \end{aligned} \quad (2)$$

The prior distributions of Model 1 and Model 2 follow the suggestions from [Gelman \(2006\)](#) and [Kruschke \(2014\)](#). The hierarchical structure of Model 1 allows setting a prior distribution for the deviations from the mean of the SES categorical variables $p(\beta_i)$. Previous studies suggest that the prior distribution for the parameter β_i follows a (half-) Cauchy distribution with shape and scale parameters 0, σ_y . The scale parameter at zero has a twofold function. As the prior distribution should inform the likely parameter regarding the deviations from the mean of the dependent variable of the SES measures occupation and education categories, the average deviation should be 0. Second, for efficiency reasons (i.e., the MCMC would not sample from implausible values), we wanted the values sampled from the MCMC to be not too far from the mean of the dependent variable. The prior distribution for the equalized income (and the other continuous covariates) $p(\beta_j)$ informs model 1 about the more likely parameters (among all the possible in the hyperparameter space) for the slope coefficients. As we have centered the continuous variables as $X_i(c) = X_i - \bar{X}_i$, the prior distribution for the slopes is normally shaped with mean 0 and a standard deviation of 1. Finally, $p(\nu)$ is the exponentially distributed prior for the ν parameter of the t-student distribution. The ν parameter shapes the thickness of the tails of the distributions, thus accommodating the outliers.

Model 2 exploits the unique feature of the Bayesian framework, that

⁵ The choice of the distribution is dependent upon the type of variable the outcome: for instance, the Poisson distribution applies with count data; for instance, if the analysis aims to model a dichotomous variable, then the Bayesian framework will start with a binomial distribution.

Table 1

Summary Statistics of the dependent variable and the covariates. Age, Income, BMI, and Sport Activity are z-standardized for computational efficiency of the MCMC algorithm.

Variables	Mean	St. Dev
CRP (log scale)	0.444	1.093
Sport Activity	3.398	2.896
Age	52.069	16.713
Income	1978.723	1534.189
BMI	27.960	5.096
Occupation		
Large employers & higher management	0.032	0.176
Higher Professionals	0.044	0.205
Lower management & professional	0.162	0.369
Intermediate	0.079	0.269
Small employers & own account	0.055	0.228
Lower supervisory & technical	0.044	0.204
Semi routine	0.095	0.293
Routine	0.055	0.228
Not in LM	0.435	0.496
Education		
Degree	0.221	0.415
Other higher degree	0.137	0.344
A-level etc	0.182	0.386
GCSE etc	0.197	0.398
Other qualification	0.117	0.322
No qualification	0.145	0.353
Gender		
Male	0.441	0.497
Female	0.559	0.497
House Ownership		
Owned	0.762	0.426
Rent	0.225	0.418
Other	0.013	0.112
Self-rated Health		
Excellent	0.154	0.361
Very good	0.355	0.478
Good	0.287	0.452
Fair	0.154	0.361
Poor	0.050	0.219
Smoking Behavior		
Smoker	0.197	0.398
Ex smoker	0.403	0.491
Non smoker	0.400	0.490

Table 2
Correlation matrix of the covariates included in the analysis of Understanding Society data.

	CRP	Occupation	Education	Gender	House Ownership	SRH	Smoking	Sport Activity	Age	Income
CRP										
Occupation	0.15***									
Education	0.18***	0.44***								
Gender	0.07***	0.05***	0.05***							
House Ownership	0.09***	0.14***	0.16***	0.04***						
Srh	0.22***	0.27***	0.26***	0.00	0.15***					
Smoking	-0.09***	-0.09***	-0.15***	0.07***	-0.21***	-0.17***				
Sport Activity	-0.22***	-0.22***	-0.25***	-0.11***	-0.09***	-0.37***	0.10***			
Age	0.16***	0.39***	0.28***	-0.04***	-0.17***	0.21***	0.04***	-0.30***		
Income	-0.12***	-0.41***	-0.33***	-0.05***	-0.18***	-0.20***	0.11***	0.17***	-0.11***	
BMI	0.37***	0.06***	0.12***	0.00	0.06***	0.23***	0.02*	-0.19***	0.12***	-0.08***

is, the possibility to also model the variance across groups through distributional models. The difference between Model 1 and Model 2 is, thus, in Model 1, only the mean parameter can depend on predictors while σ_y is assumed to be constant across observations. Conversely, in Model 2, both μ_y and σ_y can be objective of the statistical modeling, thus relaxing the assumption of homogeneity of variance. Importantly, Model 2 identifies within-group differences in CRP distribution, i.e., estimating the differences in CRP for individuals at the same level of SES. The specification of Model 2 can be formalized as follows:

$$\begin{aligned}
 \mu_y &= \beta_0 + \sum_i \beta_i X_i + \beta_j X_j \ln(\sigma_y) = \beta_0 + \sum_i \beta_i X_i + \beta_j X_j p(\sigma_\sigma) \\
 &= \mathcal{N}(0, \sigma_y) p(\beta_0) = \mathcal{N}(\mu_y, \sigma_y) p\left(\beta_j\right) = \mathcal{N}\left(0, 1\right) p\left(\beta_i\right) \\
 &= \mathcal{N}\left(0, 1\right) p\left(\sigma\right) = \mathcal{N}\left(0, \sigma_y\right) p\left(\nu\right) = \epsilon\left(1 / 29\right)
 \end{aligned} \quad (3)$$

In Model 2, we also apply the same linear function for the log-transformed standard deviation of the dependent variable⁶. The two new prior distributions are $p(\sigma_\sigma)$ and $p(\sigma_{\beta_i})$. The prior distribution $p(\sigma_\sigma)$ defines our expectations for the standard deviation of the σ distribution of the dependent variable. The prior $p(\sigma_{\beta_i})$ defines the *a priori* distribution for the deviations relative to the categories of occupational status and education. The Bayesian Regression models were performed through R, using as a backend the Stan program. Specifically, Stan uses a particular algorithm of MCMC, defined as Hamiltonian Markov Chain (HMC)⁷.

We performed four MCMC chains that included 2000 iterations. We set the burn-in (i.e., the number of initial iterations not considered due to strong autocorrelation) as the first 1000 iterations (Kruschke, 2014).

3. Results

3.1. Model 1

Table 3 shows the summary of the findings by Model 1. We focus on the results concerning the effects of occupation, education, and income on the distribution of log-CRP. Fig. 1 shows the posterior distributions of contrasts against the reference category (i.e., large employers & higher management) drawn by the MCMC for the concerns about occupational status. Under each distribution, the dot represents the expected value (i.

Table 3
Bayesian Regression results, assuming homogeneity of variance.

Parameter	μ	σ	2.5%	50%	97.5%
Intercept	0.379	0.060	0.262	0.378	0.499
Gender: Female					
Gender (Female)	0.150	0.021	0.110	0.150	0.194
House Ownership: Owner					
Rent	0.062	0.027	0.009	0.062	0.116
Other	0.063	0.089	-0.109	0.062	0.240
SRH: Excellent					
Very good	0.079	0.031	0.019	0.079	0.137
Good	0.084	0.033	0.018	0.084	0.146
Fair	0.198	0.039	0.120	0.198	0.273
Poor	0.333	0.056	0.226	0.333	0.436
Smoking: Smoker					
Ex smoker	-0.221	0.030	-0.282	-0.221	-0.161
Non smoker	-0.257	0.030	-0.317	-0.257	-0.199
Sport Activity	-0.075	0.012	-0.098	-0.075	-0.052
Age	0.085	0.013	0.058	0.085	0.111
Income	-0.015	0.011	-0.036	-0.015	0.006
BMI	0.386	0.010	0.366	0.386	0.407
Std Education (Intercept)	0.087	0.056	0.030	0.073	0.222
Std Occupation (Intercept)	0.055	0.027	0.015	0.050	0.121
σ	0.876	0.011	0.855	0.876	0.898
ν	9.747	0.891	8.220	9.688	11.766
Education, Intercepts					
Degree	-0.057	0.046	-0.155	-0.054	0.023
Other higher degree	-0.038	0.048	-0.141	-0.035	0.046
A-level	0.003	0.046	-0.094	0.005	0.090
GCSE	-0.034	0.045	-0.134	-0.031	0.048
Other qualification	0.038	0.048	-0.058	0.038	0.131
No qualification	0.086	0.048	-0.005	0.084	0.182
Occupation, Intercepts					
Large employers.& higher management	-0.043	0.046	-0.146	-0.039	0.034
Higher professional	0.006	0.038	-0.069	0.005	0.085
Lower management & professional	0.0001	0.031	-0.060	-0.000	0.063
Intermediate	-0.023	0.034	-0.096	-0.022	0.041
Small employers & own account	-0.052	0.040	-0.138	-0.048	0.014
Lower supervisory & technical	0.007	0.038	-0.065	0.005	0.085
Semi-routine	0.030	0.035	-0.031	0.028	0.103
Routine	0.026	0.038	-0.043	0.023	0.107
Not in LM	0.052	0.031	-0.002	0.051	0.120

e., the mean), while the thicker lines represent around 65% of the posteriors' probability density function (PDF). The dashed black line serves as a reference for null-divergence of the distribution toward the grand mean of the dependent variable. Fig. 1 shows that small employers, in comparison with the other categories, have a lower concentration of CRP. Surprisingly, the lower categories of occupational status (lower supervisory and technical staff, semi-routine, routine, and individuals not in the labor market) do not show remarkable differences.

Regarding the educational level and CRP concentration, the second

⁶ We have transformed the standard deviation as the Stan Development Team recommended.

⁷ More precisely, Stan's specification of the HMC is the No-U-Turn Shape (NUTS) algorithm.

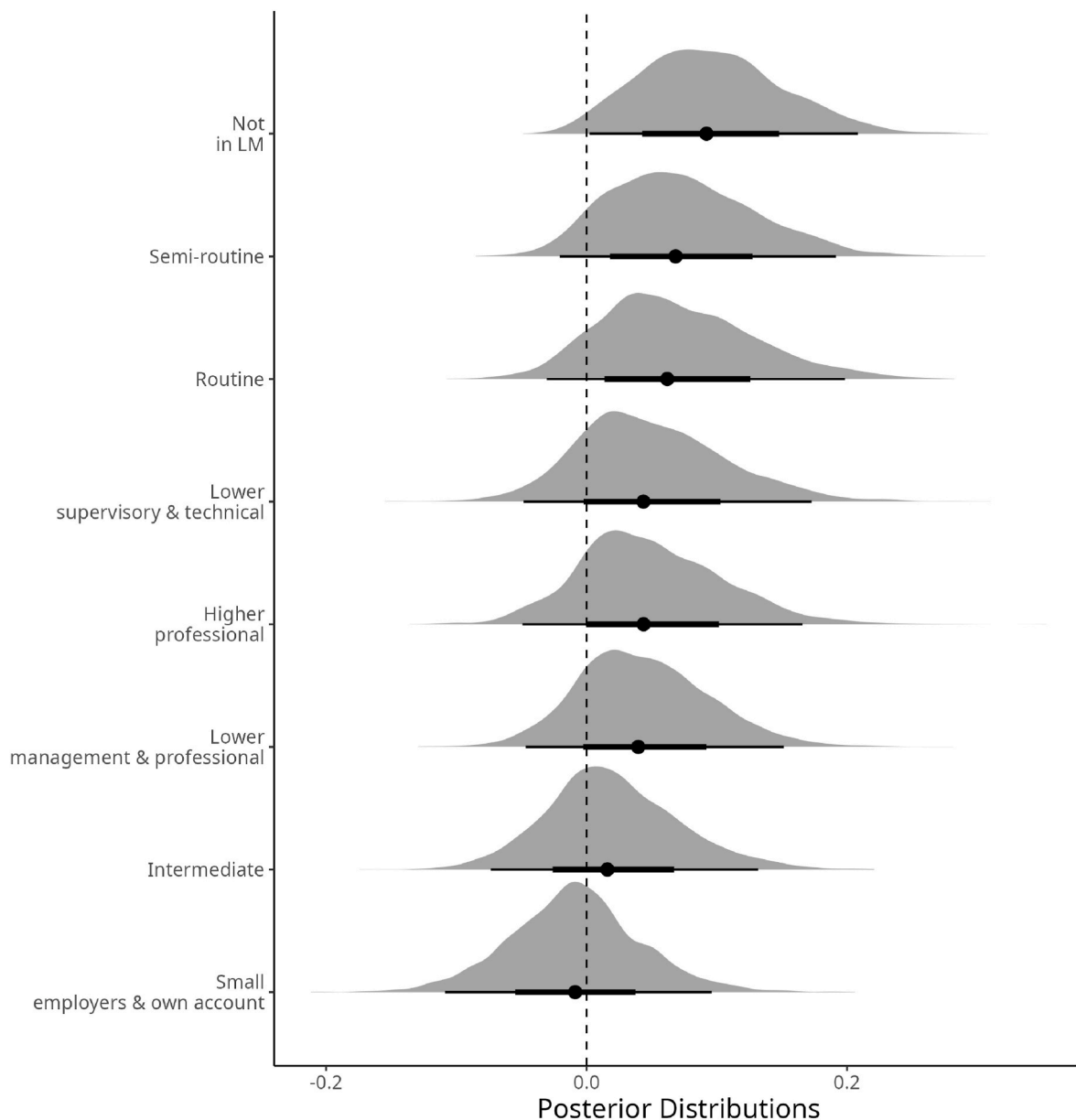


Fig. 1. Posterior Distributions of the likely deviations σ_{β_i} from the mean according to occupational status.

SES measure shows a higher level of CRP for individuals with no educational qualification concerning the referent category (i.e., individuals with educational degrees), as depicted in Fig. 2. Conversely, individuals with higher educational degrees (e.g., with GCSE and higher degrees) are more likely to have lower CRP concentrations in the blood.

To depict the relationship between household income (equivalized) and levels of CRP, we make use of scatterplots where the x-axis represents the levels of household income (in the centered scale) and the levels of (log) CRP on the y-axis. Fig. 3 shows this relationship. Conversely, to the traditional frequentist approach, it is possible to visualize different likely regression lines in the Bayesian framework, as the MCMC samples from the posterior distribution of likely regression slopes. In Fig. 3, we show 20 possible regression lines assessing the relationship between income and CRP.

From Fig. 3, it is possible to recognize a general negative relationship between equivalized income and levels of CRP. Thus, economic resources could slightly impact the risks of CVDs measured through CRP. However, the slopes of the regression lines are not relatively steep, suggesting a mild negative relationship of -0.02 (i.e., the mean of the

income posterior distribution presented in Table 3). From Fig. 3, it seems that Model 1 is affected by the leverage effect.

3.2. Model 2

In Model 2, we relaxed the homogeneity of variance assumption among the categories of occupational status and educational levels, thus predicting the mean of the dependent variable μ and its standard deviation σ . This subsection shows the results from modeling the σ parameter. As in the previous section, we first provide a tabular format of the results provided by Model 2 as in Table 4. The table makes it possible to find the results in modeling the parameter μ and the σ parameters.

Here, we focus on the results of fitting the model in the σ of the dependent variable. Similar to the section dedicated to the results of Model 1, we started by presenting the posterior distributions related to the individuals' occupational status. Fig. 4 shows the posterior distributions for each category of occupational status and how they deviate according to the scale of the standard deviation of the CRP levels.

When the model relaxes the assumption of homogeneity of variance,

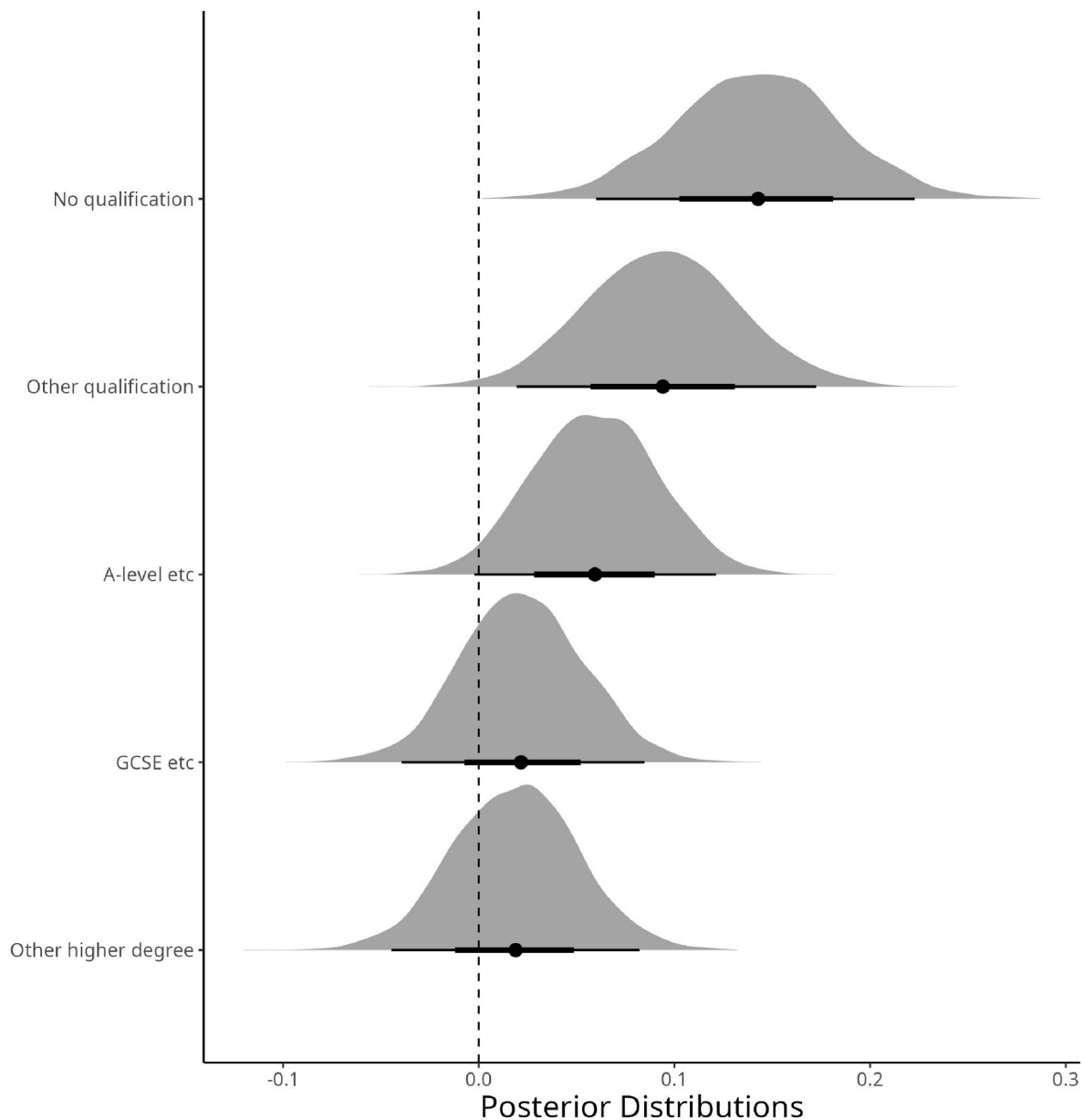


Fig. 2. Posterior Distributions of the likely deviations from the mean according to Educational levels. σ_{β_i} .

the findings suggest a rather clear variance between the large employers and higher management category and individuals not in the labor market. However, the in-between categories show a similar pattern.

Moving to the effects of educational status on the variance between categories on the variance of CRP observed through the data, Fig. 5 shows the findings of Model 2 focused on individuals' education.

From Fig. 5, the findings of Model 2 show a surprising similarity between the educational levels, with the notable exception of the individuals with an A-level. In fact, from Model 2, the variance for the individuals with an A-level educational degree is stretched toward the right.

Similar to Model 1, to present the findings from Model 2 concerning the relationship between equivalized income and logged CRP, Fig. 6 depicts 20 sampled regression lines from the posterior distribution of the regression slopes. However, the intercepts and slopes are computed according to σ_y in this case. From Fig. 6, findings suggested an even weaker relationship between equivalized income and logged CRP, with the model's homogeneity of variance assumption relaxed.

In the Supplementary Materials (Appendix A), we provide the trace

and the autocorrelation plots. These are valuable tools to assess the Bayesian inference's validity and the algorithm's correct sampling from the posterior distributions of the specific parameters.

4. Discussion & conclusions

A main current challenge in social stratification and health inequalities is the assessment of the mechanisms through which the well-documented link between deprived social conditions and health worsening takes form. The rising availability of social data that includes biological information about individuals is rising as a propitious track of research, and it has a possible double effect. First, social scientists acquire additional information potentially helpful to studying the mechanisms through which the social gradient of health occurs and socioeconomic status gets under the skin. Second, sociologists inform the biomedical literature on the importance of individuals' social and economic environments for health levels. This study investigated one of the potential mechanisms that health inequalities generate. Indeed, this study aimed to shed light on the connection between socioeconomic

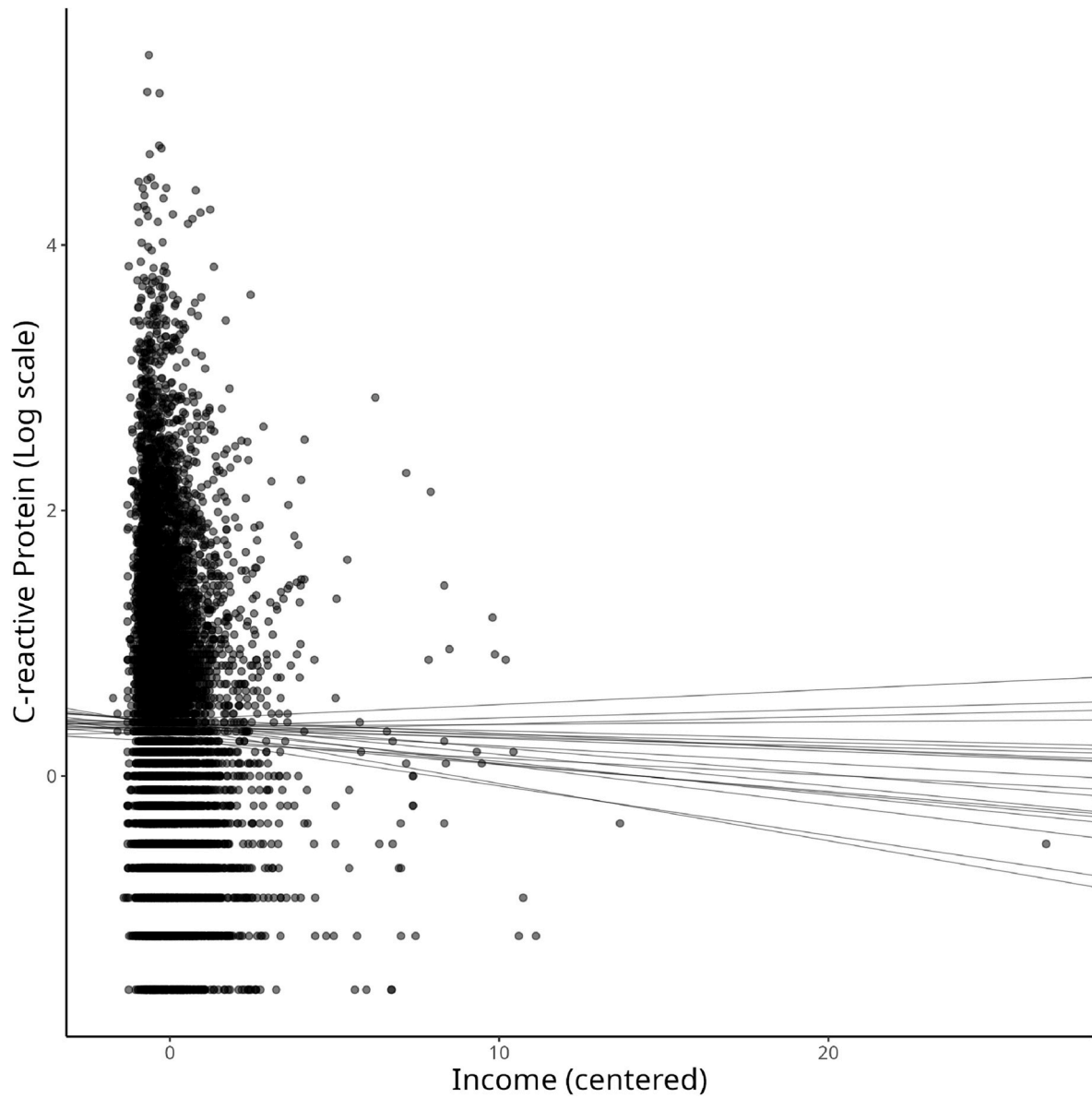


Fig. 3. Plot of the income distribution (on the x-axis) and log-CRP (on the y-axis) and model fit of 20 possible regression lines sampled from the posterior distribution β Income of Model 1.

Table 4

Bayesian Regression results, relaxing the assumption of homogeneity of variance.

Parameter	μ	σ	2.5%	50%	97.5%
Intercept	0.370	0.066	0.253	0.372	0.485
σ Intercept	-0.155	0.041	-0.235	-0.155	-0.075
Gender: Female					
Gender (female)	0.159	0.020	0.118	0.159	0.197
σ Gender (female)	0.022	0.018	-0.013	0.022	0.057
House Ownership: Owner					
Rent	0.059	0.028	0.005	0.059	0.112
Other	0.058	0.091	-0.118	0.058	0.229
σ Rent	0.012	0.023	-0.031	0.012	0.057
σ Other	0.008	0.080	-0.142	0.006	0.167
SRH: Excellent					
Very good	0.083	0.030	0.025	0.082	0.142
Good	0.090	0.032	0.026	0.090	0.152
Fair	0.203	0.039	0.126	0.204	0.279
Poor	0.338	0.057	0.227	0.337	0.450
σ Very good	0.047	0.027	-0.004	0.048	0.099
σ Good	0.056	0.028	0.001	0.056	0.114
σ Fair	0.095	0.033	0.031	0.094	0.161
σ Poor	0.064	0.047	-0.026	0.062	0.155
Smoking: Smoker					
Ex smoker	-0.220	0.031	-0.280	-0.220	-0.161
Non smoker	-0.261	0.031	-0.320	-0.261	-0.199
σ Ex smoker	-0.058	0.025	-0.107	-0.058	-0.009
σ Non smoker	-0.089	0.026	-0.138	-0.089	-0.040
Sport Activity	-0.058	0.025	-0.107	-0.058	-0.009
σ Sport Activity	-0.089	0.026	-0.138	-0.089	-0.040
Age	0.080	0.014	0.053	0.080	0.106
σ Age	-0.027	0.011	-0.049	-0.027	-0.005
Income	-0.015	0.011	-0.036	-0.015	0.007
σ Income	-0.004	0.010	-0.023	-0.004	0.015
BMI	0.386	0.011	0.365	0.385	0.407
σ BMI	-0.040	0.009	-0.057	-0.039	-0.022
Std Education (Intercept)	0.079	0.053	0.026	0.067	0.205
Std Occupation (Intercept)	0.057	0.027	0.019	0.053	0.124
σ Std Education (Intercept)	0.024	0.020	0.001	0.020	0.077
σ Std Occupation (Intercept)	0.052	0.024	0.021	0.048	0.111
ν	10.357	1.044	8.552	10.272	12.606
Education Intercepts					
Degree	-0.052	0.052	-0.142	-0.051	0.030
Other higher degree	-0.032	0.052	-0.126	-0.032	0.051
A-level	0.006	0.052	-0.083	0.004	0.096
GCSE	-0.029	0.051	-0.114	-0.029	0.055
Other qualification	0.034	0.055	-0.053	0.031	0.134
No qualification	0.080	0.056	-0.006	0.075	0.184
Occupation Intercepts					
Large employers & higher management	-0.043	0.045	-0.142	-0.039	0.033
Higher professional	0.004	0.040	-0.075	0.003	0.088
Lower management & professional	-0.000	0.030	-0.060	-0.001	0.061
Intermediate	-0.024	0.035	-0.098	-0.022	0.042
Small employers & own account	-0.056	0.038	-0.137	-0.054	0.010
Lower supervisory & technical	0.007	0.040	-0.071	0.005	0.090
Semi-routine	0.029	0.036	-0.036	0.026	0.105
Routine	0.027	0.039	-0.043	0.024	0.109
Not in LM	0.058	0.032	0.000	0.056	0.125
σ Education, Intercepts					
Degree	-0.006	0.018	-0.046	-0.003	0.027
Other higher degree	-0.007	0.019	-0.051	-0.003	0.027
A-level	0.007	0.018	-0.026	0.005	0.047
GCSE	-0.015	0.020	-0.062	-0.011	0.016
Other qualification	0.017	0.022	-0.014	0.012	0.070
No qualification	0.004	0.019	-0.033	0.002	0.047
σ Occupation, Intercepts					
Large employers & higher management	-0.020	0.037	-0.096	-0.018	0.048
Higher professional	0.037	0.037	-0.027	0.034	0.116
Lower management & professional	-0.028	0.028	-0.085	-0.027	0.026

Table 4 (continued)

Parameter	μ	σ	2.5%	50%	97.5%
Intermediate	-0.007	0.032	-0.068	-0.007	0.058
Small employers & own account	-0.057	0.038	-0.136	-0.055	0.009
Lower supervisory & technical	0.011	0.035	-0.055	0.010	0.083
Semi-routine	0.016	0.031	-0.042	0.015	0.082
Routine	-0.009	0.034	-0.076	-0.009	0.055
Not in LM	0.057	0.027	0.008	0.056	0.115

status (SES), and levels of low-grade systemic inflammation, measured through C-reactive protein (CRP). Biomedical evidence have linked high levels of CRP to increased mortality risks due to cardiovascular diseases (CVDs). We consider this pathway particularly important as mortality due to CVDs represents the first cause of death among individuals in developed countries (Brummett et al., 2014). We used data from the Understanding Society - the United Kingdom Household Longitudinal Study (UKHLS), a representative sample of the population living in England, Scotland, and Wales. In 2012, the UKHLS collected blood samples from voluntary individuals to collect markers of socially relevant health risk factors and diseases. To capture the multifaceted characteristics of individuals' SES, we included three measures of social conditions in the statistical analysis: occupational status, educational levels, and equivalized household income. This study provides a Bayesian framework to assess the pathway that links individuals' SES, mortality risks, and levels of CRP. The first Bayesian regression model (BRM) provides posterior distributions of likely effect magnitude parameters assuming homogeneity of variance across occupational and educational groups. The second step of the statistical analysis deploys a distributional model, relaxing the assumption of homogeneity of variance through modeling the standard deviation of CRP along its mean. Generally, all models show no inference problems, meaning the posterior distributions computed by both models represent all possible distributions. The concern of the findings of Model 1 is that the educational gradient is the most vital determinant of the risk of low-grade inflammation, while equivalized income is the weakest among the three SES measures. According to Davillas et al. (2017), one potential explanation is that better-educated individuals tend to pursue a healthier lifestyle and be more aware of health risks. The relationship between occupational status and levels of CRP seems to be polarized. Indeed, the findings suggest a homogeneity of CRP levels and lowest levels between individuals with the highest occupational status and small employers. That means they do not deviate remarkably from the grand mean of the dependent variable among all the other categories. While the first findings are coherent with previous literature (Marmot et al., 1991), we strongly suggest that future research deepen scientific knowledge of small employers' concerns. Regarding the equivalized income measure, some lessons can be inferred from both the methodological and substantial perspectives: the methodological conclusion is that even in the Bayesian framework, the regression model could suffer from leverage effects. The significant lesson is that economic resources are the weakest social determinant of CRP.

The findings presented in the previous section suggest that health-related behaviors are important in shaping the relationship between SES and CRP. In fact, the posterior distributions obtained by the BRM concerning smoking behavior show a significant increase in CRP levels for smokers compared with ex- and nonsmokers and a strong impact of BMI on CRP. The harmful effect of BMI is consistent within SES categories, as indicated by the low variance coefficients (see Model 2).

Model 2 provides exciting results for how individuals differ within each category of the SES measures. Indeed, the findings suggest a strong cohesion in the highest class of occupational status, meaning that inequalities in health are evident even when we consider the within-variance. Surprisingly, individuals with the same educational level are not very dissimilar, except for individuals with an A-level degree.

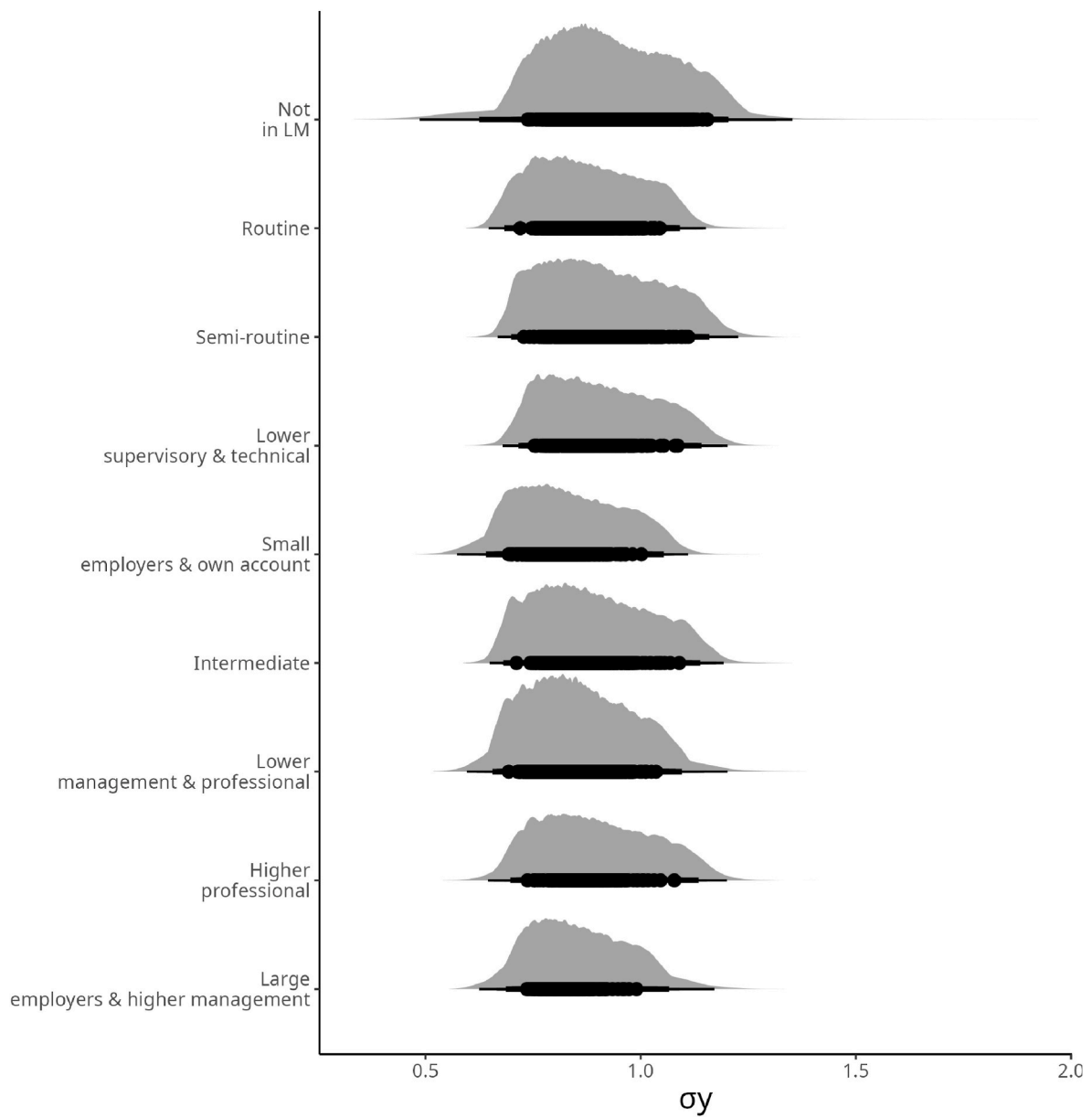


Fig. 4. Posterior Distributions of the likely deviations from the mean according to Occupational status. σ occupation on the σ_y .

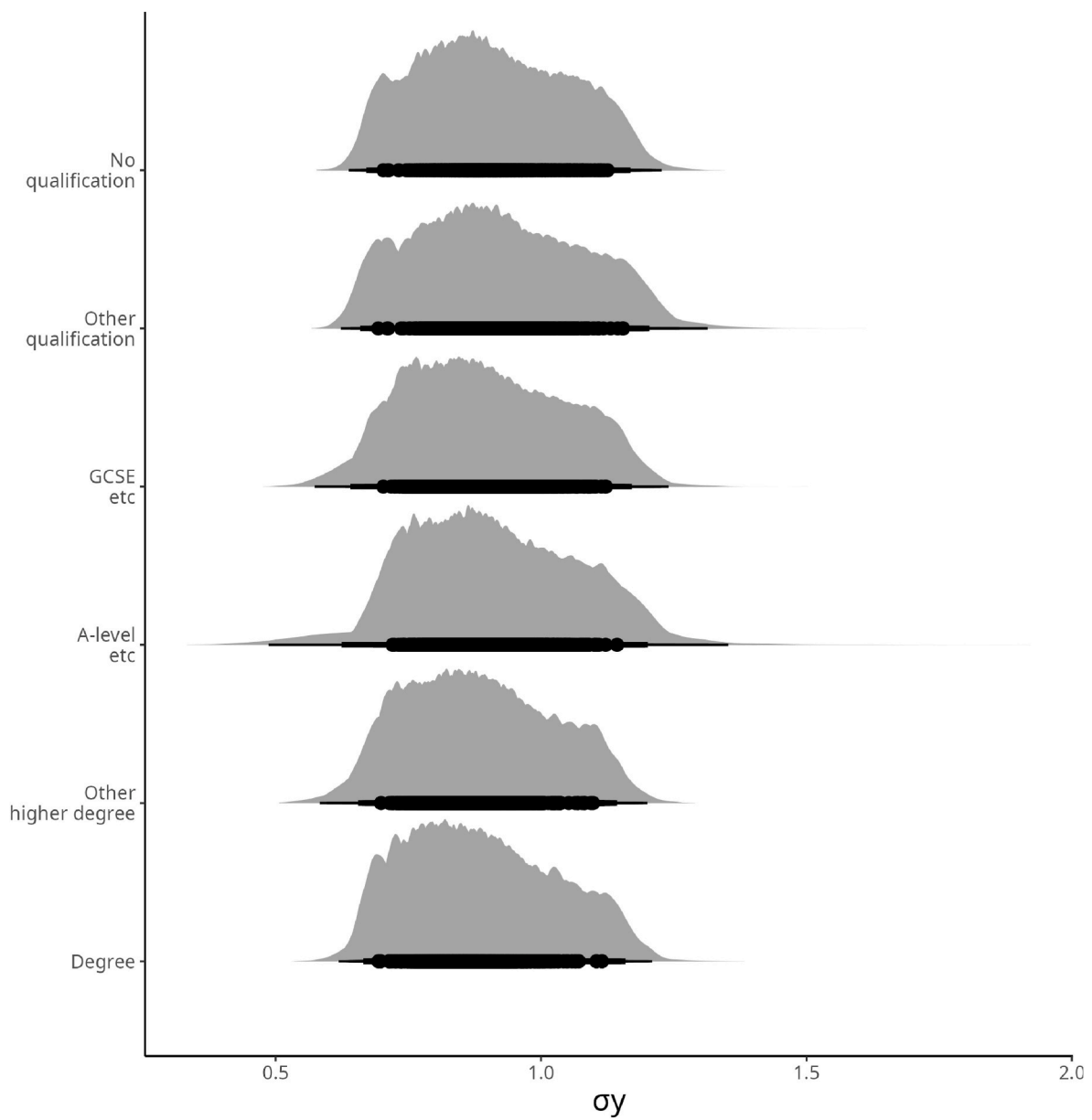


Fig. 5. Posterior Distributions of the likely deviations from the mean according to Educational levels. σ Education on the σ_y .

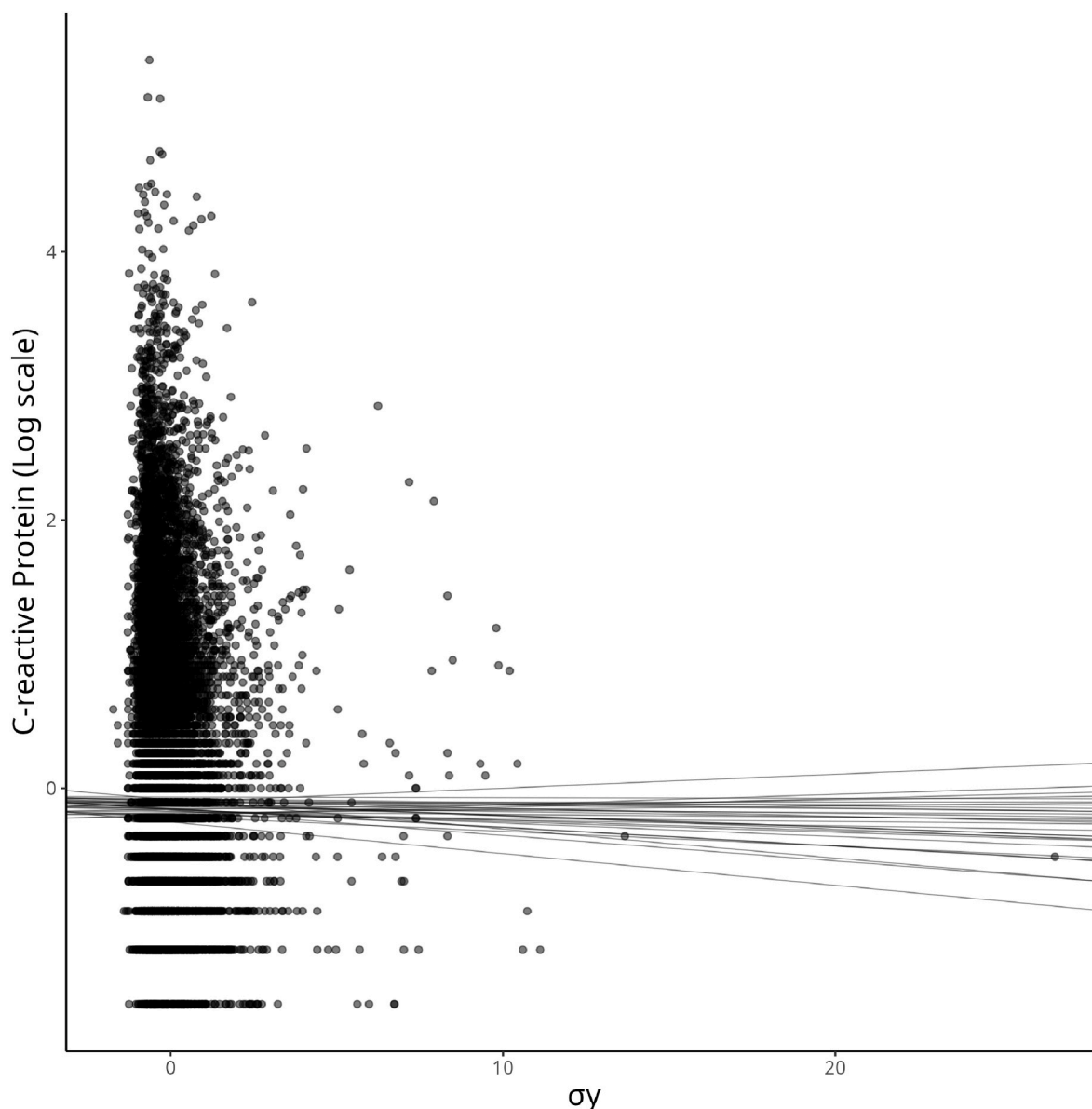


Fig. 6. Plot of the income distribution (on the x-axis) and log-CRP (on the y-axis) and model fit of 20 possible regression lines sampled from the posterior distribution β Income of Model 2.

It is worth noting that the meaning of income may change during the different stages of the life course we considered in the analysis. Therefore, we suggest investigating this aspect further by, for example, considering a more homogenous sample in terms of age.

This study has some limitations that future researchers should consider. The first limitation is the lack of direct comparison between the posterior distributions explicitly drawn for income, education, and occupational status categories. Future research could address this problem ideally from a Bayesian perspective. The second limitation concerns the model specification. Due to the already complex Bayesian models, we were unable to test the models for nonlinearities in age patterns and CVD risks through CRP. We believe it could be interesting to see whether the aging process could take other, more complex ways. The third limitation concerns the omission of other potential explanatory factors in the regression model, such as the presence of comorbidities, family's disease history, the type of individuals' diet, their living location, and their ethnicity. Future research could expand the dimensions of the regression model by including these relevant factors.

Furthermore, the study's findings are based on data from 2012.

While certain conclusions may retain their validity, more recent data sources would enable an assessment of the potential impact of the COVID-19 pandemic. The final limitation concerns robustness checks. Indeed, we specified the models only with this set of prior distributions. Hence, we invite the social scientific community to use our results for better-refined models coherently with the Bayesian philosophy. The Understanding Society data provides the empirical researcher with a wide range of biological markers of (ab-) normal physiological functioning. Future research could potentially exploit the biological information collected by empirically testing the social gradient of health using composite measures such as Allostatic Load. Future researchers could take advantage of this first Bayesian implementation and use the results we provided as a starting point to define a theoretically guided model to advance our knowledge concerning the social gradient of mortality risks.

Author statement

Procopio Alessandro: Conceptualization, Formal Analysis,

Methodology, Writing –Original Draft, Robin Samuel: Conceptualization, Validation, Writing – Review & Editing, Supervision, Planning.

Declaration of competing interest

The authors declare that they have no known competing financial

interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Appendix A. Trace and Autocorrelation Plots of the MCMC produced by the Bayesian Regression Models

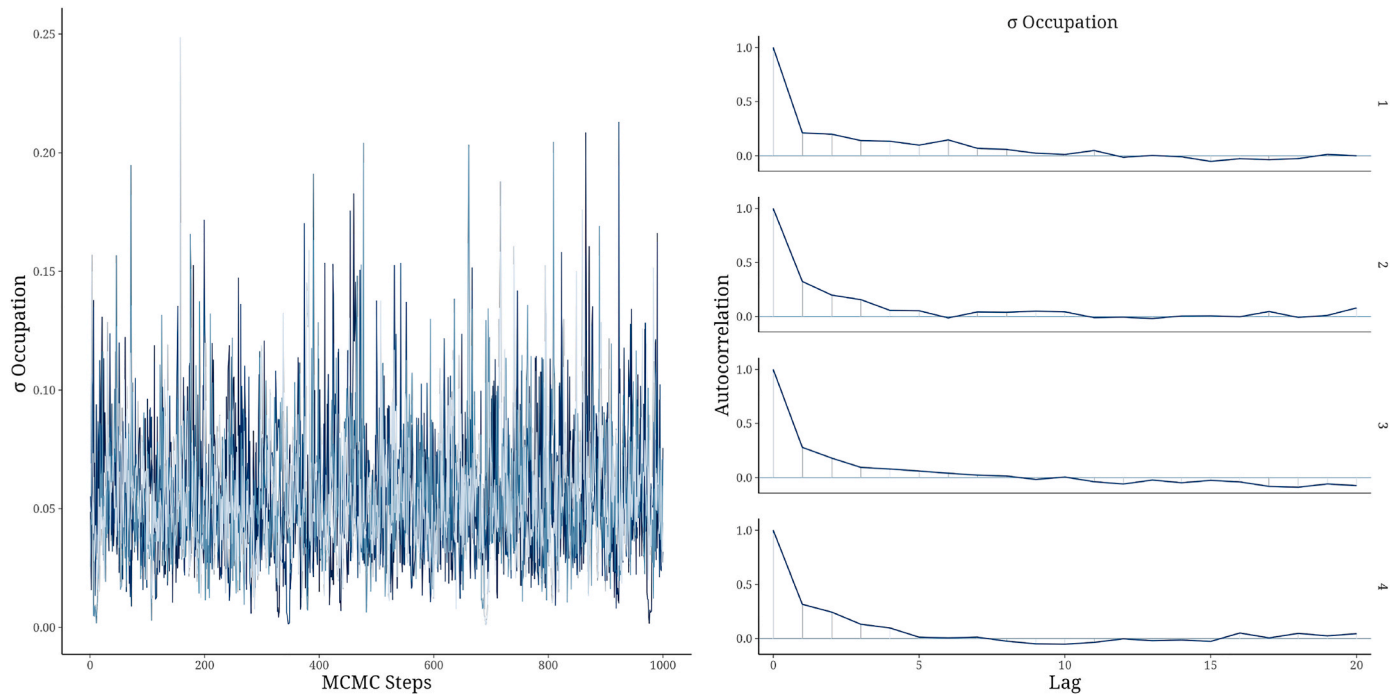


Fig. 7. Panel (a): Trace plot of the occupational status parameter σ_{β_i} . Panel (b): Autocorrelation plot of σ_{β_i} by the number of chains of the MCMC..

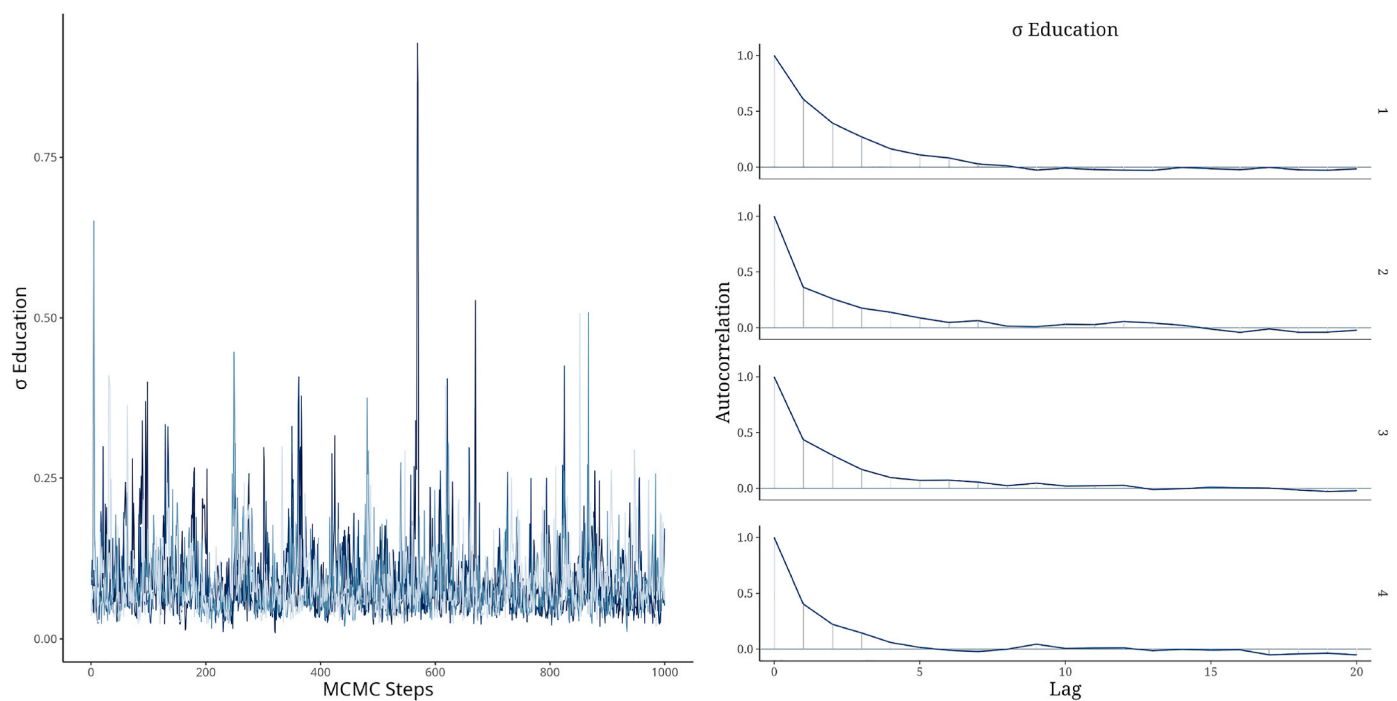


Fig. 8. Panel (a): Trace plot of the educational attainment parameter σ_{β_i} . Panel (b): Autocorrelation plot of σ_{β_i} by the number of chains of the MCMC..

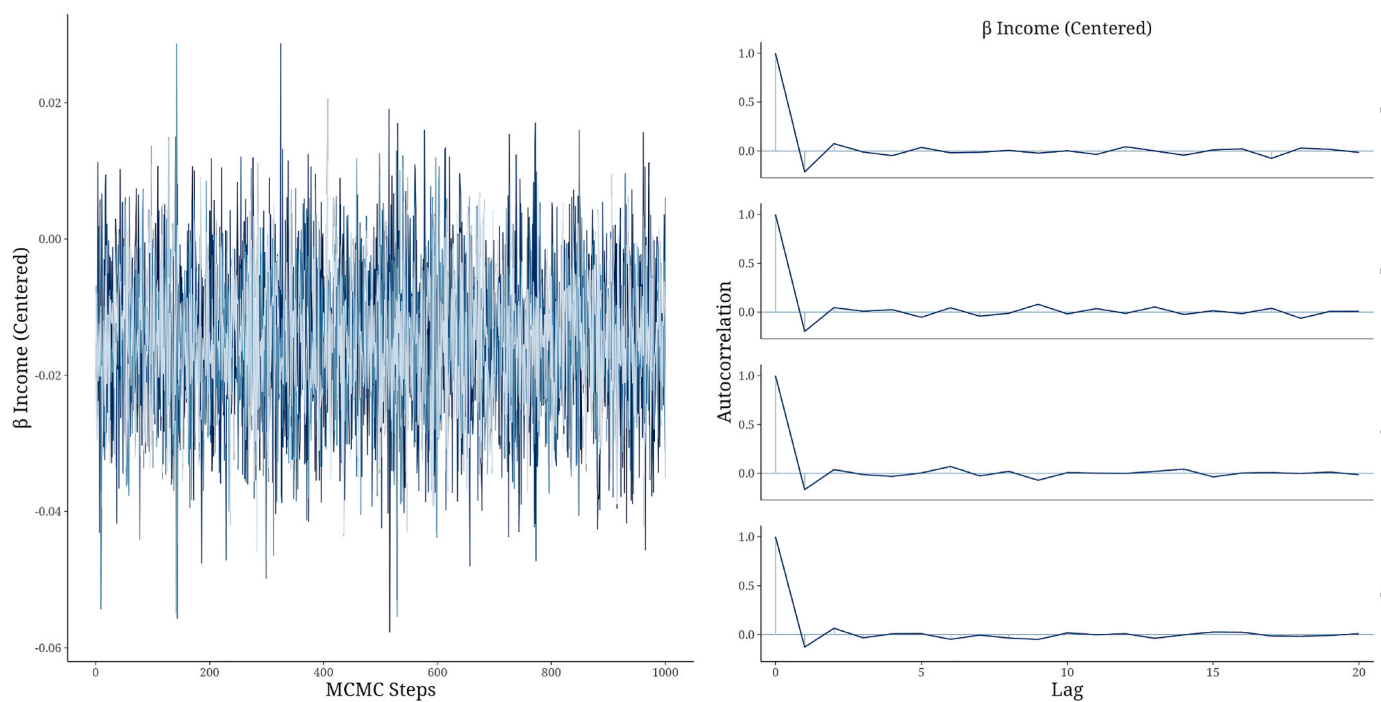


Fig. 9. Panel (a): Trace plot of the β income parameter. Panel (b): Autocorrelation plot of β by the number of chains of the MCMC. Model 2.

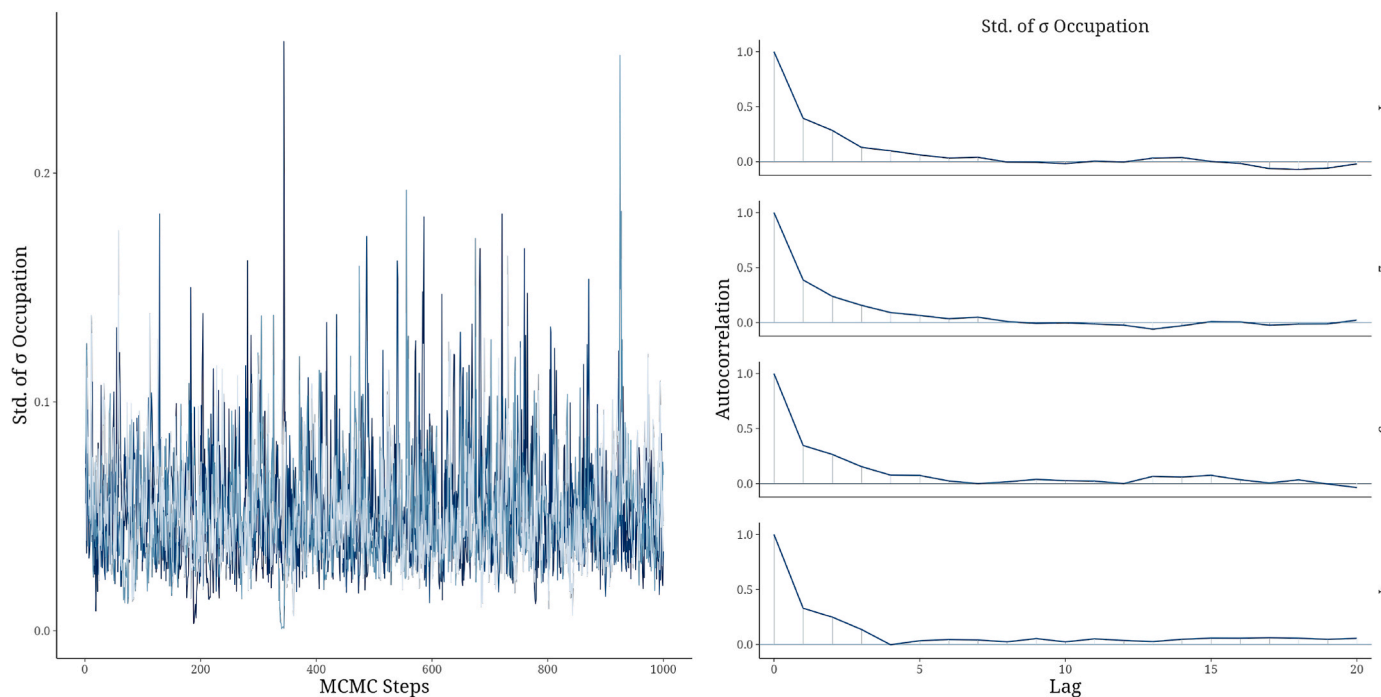


Fig. 10. Panel (a): Trace plot of the occupational status parameter σ_{β_i} on σ_γ . Panel (b): Autocorrelation plot of σ_{β_i} by the number of chains of the MCMC..

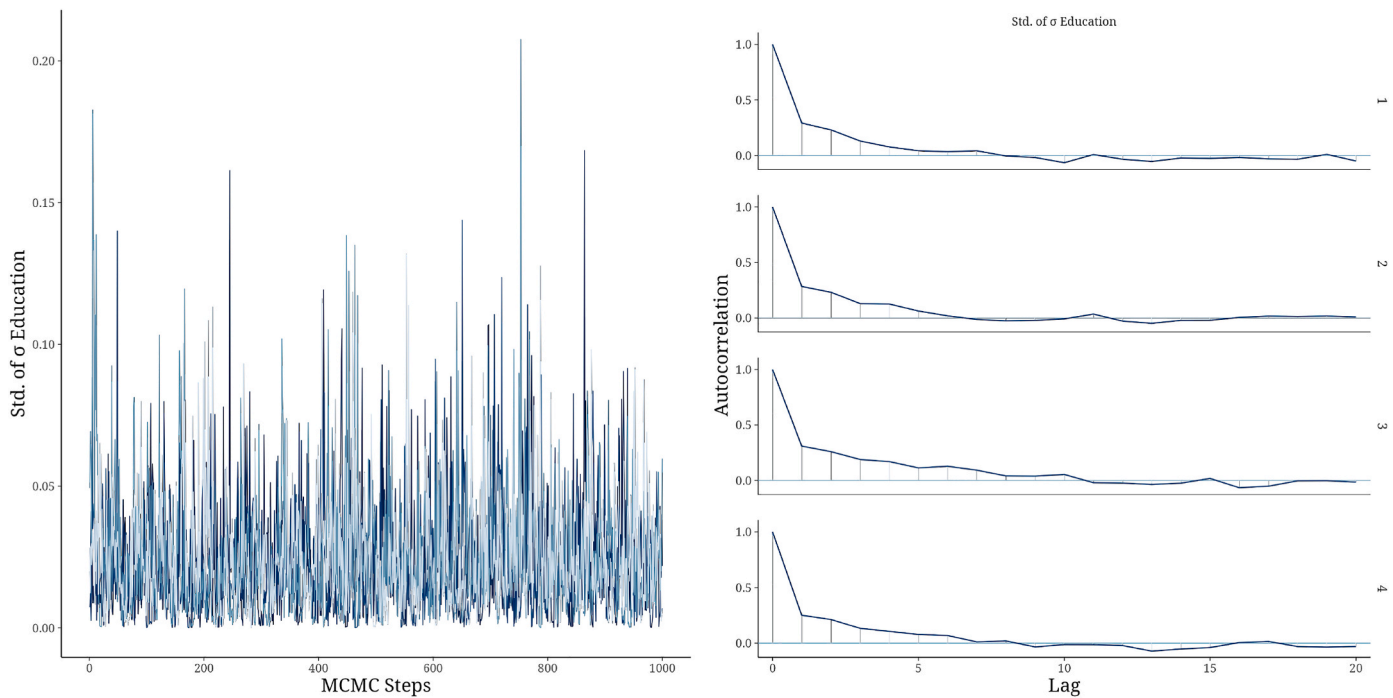


Fig. 11. Panel (a): Trace plot of the educational attainment parameter σ_{β_i} on σ_{γ} . Panel (b): Autocorrelation plot of σ_{β_i} by the number of chains of the MCMC..

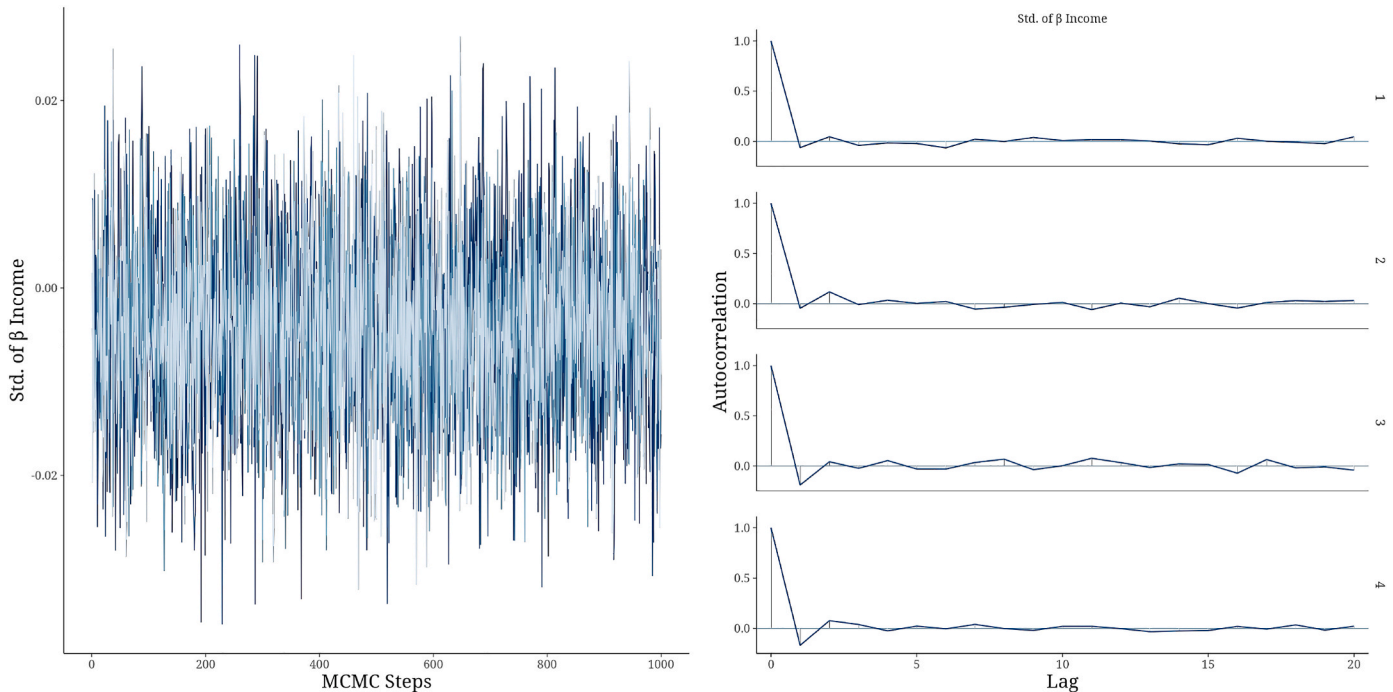


Fig. 12. Panel (a): Trace plot of the β income parameter on σ_{γ} . Panel (b): Autocorrelation plot of β by the number of chains of the MCMC..

Appendix B. - Results of the subgroup analysis on missing values

Table 5

Summary statistics of the covariates used in the analysis by missing indicator variables on CRP, BMI, and on the covariates. Table 1 reports the percentages and proportion confidence intervals for the categorical variables. For the continuous variables, Table 1 reports the mean; confidence intervals of the mean are in parentheses.

	Missing values on CRP (N = 521)	Missing values on BMI (N = 283)	Missing Values overall (N = 936)	Analytical Sample (N = 8960)
--	---------------------------------	---------------------------------	----------------------------------	------------------------------

Occupation (%)

(continued on next page)

Table 5 (continued)

	Missing values on CRP (N = 521)	Missing values on BMI (N = 283)	Missing Values overall (N = 936)	Analytical Sample (N = 8960)
Large employers & higher management	3.07 (0.018, 0.049)	1.06 (0.002, 0.031)	2.32 (0.014, 0.035)	3.13 (0.028, 0.035)
Higher professional	5.56 (0.038, 0.079)	3.91 (0.020, 0.069)	4.76 (0.035, 0.064)	4.42 (0.040, 0.048)
Lower management & professional	17.27 (0.141, 0.208)	10.32 (0.070, 0.145)	13.62 (0.115, 0.160)	16.00 (0.153, 0.167)
Intermediate	8.06 (0.059, 0.107)	5.33 (0.030, 0.087)	6.53 (0.050, 0.083)	7.74 (0.072, 0.083)
Small employers & own account	6.14 (0.042, 0.086)	7.11 (0.044, 0.108)	5.30 (0.044, 0.076)	5.54 (0.051, 0.060)
Lower supervisory & technical	4.03 (0.025, 0.061)	2.49 (0.010, 0.051)	3.32 (0.023, 0.047)	4.25 (0.039, 0.047)
Semi-routine	8.44 (0.062, 0.112)	7.47 (0.047, 0.112)	8.30 (0.066, 0.103)	9.34 (0.088, 0.099)
Routine	6.52 (0.046, 0.090)	4.27 (0.022, 0.073)	6.20 (0.047, 0.080)	5.54 (0.051, 0.060)
Not in LM	40.88 (0.366, 0.452)	58.0 (0.520, 0.638)	49.05 (0.458, 0.524)	43.99 (0.430, 0.450)
Education (%)				
Degree	28.51 (0.246, 0.327)	17.43 (0.132, 0.224)	25.754 (0.229, 0.288)	22.41 (0.216, 0.233)
Other higher degree	12.47 (0.097, 0.157)	10.32 (0.070, 0.145)	11.48 (0.094, 0.138)	13.50 (0.128, 0.142)
A-level etc	20.0 (0.166, 0.238)	14.23 (0.104, 0.189)	17.51 (0.150, 0.202)	18.17 (0.174, 0.190)
GCSE etc	18.21 (0.149, 0.219)	21.70 (0.170, 0.270)	19.83 (0.172, 0.227)	19.70 (0.189, 0.205)
Other qualification	11.28 (0.087, 0.144)	11.74 (0.082, 0.161)	10.78 (0.088, 0.131)	11.65 (0.110, 0.123)
No qualification	9.50 (0.071, 0.124)	24.55 (0.196, 0.300)	14.61 (0.123, 0.172)	14.54 (0.139, 0.153)
Gender (%)				
Male	46.25 (0.419, 0.506)	48.41 (0.425, 0.544)	47.32 (0.441, 0.506)	44.44 (0.435, 0.454)
Female	53.74 (0.494, 0.581)	51.59 (0.456, 0.575)	52.67 (0.494, 0.559)	55.55 (0.546, 0.565)
House Ownership (%)				
Owned	74.51 (0.705, 0.782)	62.89 (0.570, 0.685)	69.89 (0.668, 0.729)	75.63 (0.748, 0.765)
Rent	23.55 (0.200, 0.274)	35.33 (0.298, 0.412)	28.35 (0.254, 0.314)	23.06 (0.222, 0.239)
Other	1.93 (0.009, 0.035)	1.76 (0.006, 0.041)	1.75 (0.010, 0.028)	1.30 (0.011, 0.016)
Self-rated Health (%)				
Excellent	21.30 (0.179, 0.251)	7.77 (0.049, 0.115)	18.15 (0.157, 0.208)	15.67 (0.150, 0.164)
Very good	38.38 (0.342, 0.427)	21.55 (0.169, 0.268)	32.65 (0.296, 0.358)	35.21 (0.343, 0.362)
Good	22.07 (0.186, 0.259)	0.269 (0.218, 0.324)	24.16 (0.214, 0.271)	28.27 (0.274, 0.292)
Fair	12.28 (0.096, 0.154)	26.85 (0.208, 0.313)	16.21 (0.139, 0.187)	15.43 (0.147, 0.162)
Poor	5.95 (0.041, 0.083)	18.02 (0.137, 0.230)	8.80 (0.071, 0.108)	5.39 (0.050, 0.059)
Smoking (%)				
Smoker	18.81 (0.155, 0.224)	19.08 (0.147, 0.242)	18.18 (0.158, 0.208)	19.53 (0.188, 0.203)
Ex-smoker	36.27 (0.321, 0.406)	45.23 (0.393, 0.512)	37.75 (0.346, 0.409)	40.09 (0.391, 0.411)
Non-smoker	44.91 (0.406, 0.493)	35.68 (0.301, 0.416)	44.06 (0.409, 0.473)	0.404 (0.394, 0.413)
Sport Activity				
Mean (CI)	3.988 (3.725, 4.252)	2.198 (1.889, 2.507)	3.543 (3.346, 3.740)	3.412 (3.355, 3.469)
Age				
Mean (CI)	45.965 (44.390, 47.541)	56.360 (54.209, 58.512)	47.349 (46.096, 48.603)	51.622 (51.286, 51.958)
Equivalized Income				
Mean (CI)	2.145.61 (1977.29, 313.94)	1.708.57 (1539.69, 1877.46)	1.975.45 (1863.66, 2087.23)	1.978.41 (1947.77, 2010.05)
BMI				
Mean (CI)	25.144 (24.670, 25.618)	27.783 (27.680, 27.886)		

Table 6

Results of subgroup analysis concerning missing values, compared to analytical sample. Categorical variables report the Chi-squared association statistic value. Continuous variables report the t-student statistic value. Degrees of freedom are in parentheses.

	Missing values on CRP (N = 521)	Missing values on BMI (N = 283)	Missing Values overall (N = 936)
Occupation	5.32 (8)	30.38*** (8)	17.71* (8)
Education	20.18*** (5)	28.34*** (5)	8.37 (5)
Gender	0.65819 (1)	16.957 (1)	3.36 (1)
House Ownership	1.77 (2)	25.76*** (2)	18.10 (2)
SRH	24.06*** (4)	133.45*** (4)	34.62*** (4)
Smoking Behavior	4.95 (2)	3.53 (2)	5.85*** (2)
Age	7.28*** (570.55)	−4.41*** (296.36)	7.12*** (1082.8)
Sport Activity	−4.43*** (573.1)	7.83*** (302.6)	−1.37 (1110.6)
Equivalized Income	−2.03* (555.9)	3.18*** (300.58)	0.05 (1089.5)
BMI	11.28*** (551.27)		

References

- Adler, N. E., Boyce, T., Chesney, M. A., Cohen, S., Folkman, S., Kahn, R. L., & Syme, L. S. (1994). Socioeconomic status and health: The challenge of the gradient. *American Psychologist*, 49(1), 15–24.
- Adler, N. E., & Ostrove, J. M. (1999). Socioeconomic status and health: What we know and what we don't. *Annals of the New York Academy of Sciences*, 896, 3–15.
- Alley, D. E., Seeman, T. E., Ki Kim, J., Karlamangla, A., Hu, P., & Crimmins, E. M. (2006). Socioeconomic status and C-reactive protein levels in the US population: Nhanes IV. *Brain, Behavior, and Immunity*, 20(5), 498–504.
- Baum, A., Garofalo, J. P., & Yali, A. M. (1999). Socioeconomic status and chronic stress. Does stress account for SES effects on health? *Annals of the New York Academy of Sciences*, 896, 131–144.
- Ben-Shlomo, Y., & Kuh, D. (2002). A life course approach to chronic disease epidemiology: Conceptual models, empirical challenges and interdisciplinary perspectives. *International Journal of Epidemiology*, 31(2), 285–293.
- Braveman, P., Krieger, N., & Lynch, J. (2000). Health inequalities and social inequalities in health. *Bulletin of the World Health Organization*, 78(2), 232–235.
- Brummett, B. H., Babyak, M. A., Singh, A., Jiang, R., Williams, R. B., Harris, K. M., & Siegler, I. C. (2014). Socioeconomic indices as independent correlates of C-reactive protein in the national longitudinal study of adolescent health. *Psychosomatic Medicine*, 75(9), 882–893.
- Clark, A. M., DesMeules, M., Luo, W., Duncan, A. S., & Wielgosz, A. (2009). Socioeconomic status and cardiovascular disease: Risks and implications for care. *Nature Reviews Cardiology*, 6(11), 712–722.
- Corna, L. M. (2013). A life course perspective on socioeconomic inequalities in health: A critical review of conceptual frameworks. *Advances in Life Course Research*, 18(2), 150–159.
- Davillas, A., Benzeval, M., & Kumari, M. (2017). Socio-economic inequalities in C-reactive protein and fibrinogen across the adult age span: Findings from Understanding Society. *Scientific Reports*, 7(1), 1–13.
- Davillas, A., Jones, A. M., & Benzeval, M. (2019). The income-health gradient: Evidence from self-reported health and biomarkers in understanding society. In *Panel data econometrics: Empirical applications* (pp. 709–741). Elsevier Inc.
- Dowd, J. B., & Zajacova, A. (2007). Does the predictive power of self-rated health for subsequent mortality risk vary by socioeconomic status in the US? *International Journal of Epidemiology*, 36(6), 1214–1221.
- Dowd, J. B., Zajacova, A., & Aiello, A. (2009). Early origins of health disparities: Burden of infection, health, and socioeconomic status in U.S. children. *Social Science & Medicine*, 68(4), 699–707.
- Elo, I. T. (2009). Social class differentials in health and mortality: Patterns and explanations in comparative perspective. *Annual Review of Sociology*, 35, 553–572.
- Farmer, H. R., Thomas Tobin, C. S., & Thorpe, R. J. (2022). Correlates of elevated c-reactive protein among black older adults: Evidence from the health and retirement study. *The Journals of Gerontology: Serie Bibliographique*, 77(11), 1964–1977.
- Farmer, H. R., Wray, L. A., & Haas, S. A. (2021). Race, gender, and socioeconomic variations in c-reactive protein using the health and retirement study. *The Journals of Gerontology: Serie Bibliographique*, 76(3), 583–595.
- Farmer, H. R., Wray, L. A., Xian, Y., Xu, H., Pagidipati, N., Peterson, E. D., & Dupre, M. E. (2020). Racial differences in elevated c-reactive protein among us older adults. *Journal of the American Geriatrics Society*, 68(2), 362–369.
- Ferraro, K. F., & Shippee, T. P. (2009). Aging and cumulative inequality: How does inequality get under the skin? *The Gerontologist*, 49(3), 333–343.
- Freese, J. (2018). The arrival of social science genomics. *Contemporary Sociology*, 47(5), 524–536.
- Gakidou, E., & King, G. (2002). Measuring total health inequality: Adding individual variation to group-level differences. *International Journal for Equity in Health*, 1(1), 1–12.
- Gakidou, E. E., Murray, C. J. L., & Frenk, J. (2000). A framework for measuring health inequality.
- Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models (comment on article by browne and draper). *Bayesian analysis*, 1(3), 515–534.
- Gimeno, D., Ferrie, J. E., Elovainio, M., Pulkki-Raback, L., Keltikangas-Jarvinen, L., Eklund, C., Hurme, M., Lehtimäki, T., Marniemi, J., Viikari, J. S., Raitakari, O. T., & Kivimäki, M. (2008). When do social inequalities in C-reactive protein start? A life course perspective from conception to adulthood in the cardiovascular risk in young Finns study. *International Journal of Epidemiology*, 37(2), 290–298.
- Goldthorpe, J. H. (2010). Analysing social inequality : A critique of two recent contributions from economics and epidemiology. *European Sociological Review*, 26(6), 731–744.
- Goodman, E., McEwen, B. S., Huang, B., Dolan, L. M., & Adler, N. E. (2005). Social inequalities in biomarkers of cardiovascular risk in adolescence. *Psychosomatic Medicine*, 67(1), 9–15.
- Hallqvist, J., Lynch, J., Bartley, M., Lang, T., & Blane, D. (2004). Can we disentangle life course processes of accumulation, critical period and social mobility? An analysis of disadvantaged socio-economic positions and myocardial infarction in the stockholm heart epidemiology program. *Social Science & Medicine*, 58(8), 1555–1562.
- Harris, T. B., Ferrucci, L., Tracy, R. P., Corti, M. C., Wacholder, S., Ettinger, W. H., Heimovitz, H., Cohen, H. J., & Wallace, R. (1999). Associations of elevated interleukin-6 and C-reactive protein levels with mortality in the elderly. *The American Journal of Medicine*, 106(5), 506–512.
- Harris, K. M., & Schorpp, K. M. (2018). Integrating biomarkers in social stratification and health research. *Annual Review of Sociology*, 44(1), 361–386.
- Herd, P., Goesling, B., & House, J. S. (2007). Socioeconomic position and health: The differential effects of education versus income on the onset versus progression of health problems. *Journal of Health and Social Behavior*, 48(3), 223–238.
- House, J. S. (2002). Understanding social factors and inequalities in health: 20th century progress and 21st century prospects. *Journal of Health and Social Behavior*, 43(2), 125–142.
- Jousilahti, P., Salomaa, V., Rasi, V., Vahtera, E., & Palosuo, T. (2003). Association of markers of systemic inflammation, C reactive protein, serum amyloid A, and fibrinogen, with socioeconomic status. *Journal of Epidemiology & Community Health*, 57, 730–733.
- Karimi, M., Castagné, R., Delpierre, C., Albertus, G., Berger, E., Vineis, P., Kumari, M., Kelly-Irving, M., & Chadeau-Hyam, M. (2019). Early-life inequalities and biological ageing: A multisystem biological health score approach in Understanding Society. *Journal of Epidemiology & Community Health*, 73(8), 693–702.
- Kawachi, I., Subramanian, S. V., & Almeida-Filho, N. (2002). A glossary for health inequalities.
- Koster, A., Bosma, H., Penninx, B. W. J. H., Newman, A. B., Harris, T. B., Eijk, J. T. M. V., Kempen, G. I. J. M., Simonsick, E. M., Johnson, K. C., Rooks, R. N., Ayonayon, H. N., Rubin, S. M., & Kritchevsky, S. B. (2006). Association of inflammatory markers with socioeconomic status. *Journal of Gerontology*, 61(3), 284–290.
- Kruschke, J. (2014). *Doing bayesian data analysis: A tutorial with R, jags, and stan*. Academic Press.
- Laaksonen, D. E., Niskanen, L., Nyyssönen, K., Punnonen, K., Tuomainen, T. P., & Salonen, J. T. (2005). C-Reactive protein in the prediction of cardiovascular and

- overall mortality in middle-aged men: A population-based cohort study. *European Heart Journal*, 26(17), 1783–1789.
- Link, B., & Phelan, J. (1995). Social conditions as fundamental causes of health inequalities. *Journal of Health and Social Behavior*, 35, 80–94.
- Liu, R. S., Aiello, A. E., Mensah, F. K., Gasser, C. E., Rueb, K., Cordell, B., Juonala, M., Wake, M., & Burgner, D. P. (2017). Socioeconomic status in childhood and C reactive protein in adulthood: A systematic review and meta-analysis. *Journal of Epidemiology & Community Health*, 71(8), 817–826.
- Lubbock, L. A., Goh, A., Ali, S., Ritchie, J., & Whooley, M. A. (2005). Relation of low socioeconomic status to C-reactive protein in patients with coronary heart disease (from the Heart and Soul study). *The American Journal of Cardiology*, 96(11), 1506–1511.
- Lynch, S. M., & Bartlett, B. (2019). Bayesian statistics in sociology: Past, present, and future. *Annual Review of Sociology*, 45, 47–68.
- Marmot, M. G., Stansfeld, S., Patel, C., North, F., Head, J., White, I., Brunner, E., Feeney, A., Marmot, M. G., & Smith, G. D. (1991). Health inequalities among British civil servants: The whitehall II study. *The Lancet*, 337(8754), 1387–1393.
- McDade, T. W., Lindau, S. T., & Wroblewski, K. (2011). Predictors of C-reactive protein in the national social life, health, and aging project. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 66 B(1), 129–136.
- McEwen, B. S. (2015). Biomarkers for assessing population and individual health and disease related to stress and adaptation. *Metabolism, Clinical and Experimental*, 64(3), S2–S10.
- McFall, S. L., Booker, C., Burton, J., & Conolly, A. (2012). *Implementing the biosocial component of understanding society – nurse collection of biomeasures and anne conolly nurse collection of biomeasures*.
- Mitchell, U. A., & Aneshensel, C. S. (2017). Social inequalities in inflammation: Age variations in older persons. *Journal of Aging and Health*, 29(5), 769–787.
- Murray, C. J., Gakidou, E. E., & Frenk, J. (1999). Health inequalities and social group differences: What should we measure? *Bulletin of the World Health Organization*, 77(7), 537.
- Phelan, J. C., Link, B. G., & Tehranifar, P. (2010). Social conditions as fundamental causes of health inequalities: Theory, evidence, and policy implications. *Journal of Health and Social Behavior*, 51, 28–40.
- Power, C., Manor, O., & Matthews, S. (1999). The duration and timing of exposure: Effects of socioeconomic environment on adult health. *American Journal of Public Health*, 89(7), 1059–1065.
- Pudrovska, T. (2014). Early-life socioeconomic status and mortality at three life course stages: An increasing within-cohort inequality. *Journal of Health and Social Behavior*, 55(2), 181–195.
- Shanahan, L., Freeman, J., & Bauldry, S. (2014). Is very high C-reactive protein in young adults associated with indicators of chronic disease risk? *Psychoneuroendocrinology*, 40(1), 76–85.
- Umlauf, N., & Kneib, T. (2018). A primer on Bayesian distributional regression. *Statistical Modelling*, 18(3–4), 219–247.
- Winkleby, M. A., Jatulis, D. E., Frank, E., & Fortmann, S. P. (1992). Socioeconomic status and health: How education, income, and occupation contribute to risk factors for cardiovascular disease. *American Journal of Public Health*, 82(6), 816–820.
- Yang, Y. C., Schorpp, K., Boen, C., Johnson, M., & Harris, K. M. (2020). Socioeconomic status and biological risks for health and illness across the life course. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 75(3), 613–624.