HEDG
HEALTH, ECONOMETRICS AND DATA GROUP
THE UNIVERSITY of York

WP 17/04

The income-health gradient:

Evidence from self-reported health and biomarkers using longitudinal data on income

Apostolos Davillas; Andrew M. Jones and Michaela Benzeval

March 2017

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

The income-health gradient: Evidence from self-reported health and biomarkers using longitudinal data on income

Apostolos Davillas
Institute for Social and Economic Research, University of Essex
Andrew M Jones
Department of Economics and Related Studies, University of York
Centre for Health Economics, Monash University
Department of Economics, University of Bergen
Michaela Benzeval
Institute for Social and Economic Research, University of Essex

Abstract

This paper adds to the literature on the income-health gradient by exploring the association between short- and long-term income and a wide set of self-reported health measures and objective nurse-administered and blood-based biomarkers as well as employing estimation techniques that allow for analysis "beyond the mean" and accounting for unobserved heterogeneity. The income-health gradients are greater in magnitude in case of long-run rather than cross-sectional income measures. Unconditional quantile regressions reveal that the differences between the long-run and the short-run income gradients are more evident towards the tails of the distributions, where both higher risk of illnesses and steeper income gradients are observed. A two-step estimator, involving a fixed-effects income model at the first stage, shows that the individual-specific selection effects have a systematic impact in the long-run income gradients in self-reported health but not in biomarkers, highlighting the importance of reporting error in self-reported health.

Keywords: biomarkers; health inequalities; panel data; Understanding Society.

JEL codes: C1, C5, I14.

Understanding Society is an initiative funded by the Economic and Social Research Council and various Government Departments, with scientific leadership by the Institute for Social and Economic Research, University of Essex, and survey delivery by NatCen Social Research and Kantar Public. The research data are distributed by the UK Data Service. We are grateful to the Economic and Social Research Council for financial support for this research via project "How can biomarkers and genetics improve our understanding of society and health?" (award no. ES/M008592/1). The funders, data creators and UK Data Service have no responsibility for the contents of this paper.

1. Introduction

Economic studies that aim to explore the health-income gradient face a number of challenges (for example, Benzeval and Judge, 2001; Chou et al., 2004; Deaton and Paxson, 1998; Ettner, 1996; Jones and Wildman, 2008; Van Doorslaer and Jones, 2003). Firstly, self-assessed health (SAH) measures are indirect indicators of underling health, which may be subject to misreporting and are associated with comparability problems at both the individual level and among countries (Bago d'Uva et al. 2008; Jürges 2007, 2008). If the reporting error is randomly distributed, this might not be an issue. However, reporting bias has been shown to vary systematically with income and other socioeconomic characteristics that are often used to explore health inequalities, which raises doubts about the robustness of studies based on self-reported health indicators (Crossley and Kennedy, 2002; Dowd and Zajacova, 2010; Ziebarth, 2010). The same holds true for other self-reported health measures, such as functional limitations, self-reported diagnosis of chronic conditions and self-administered well-being measures (Baker et al., 2004; Daltroy et al., 1999; Johnston et al., 2009; Powdthavee, 2010).

Secondly, self-reported health indicators and other health proxies do not give information about the pathways through which economic conditions get "under the skin", and may miss important information about pre-symptom stages. Identifying the role of income in physiological processes that occur before a disease or condition manifests may be particularly important for better understanding the link between income and health (Dowd et al., 2009; Jürges et al., 2013). Recent studies have explored the association between measures of socioeconomic status and more objective and proximal health measures such as blood-based or nurse-administered biomarkers[1]. Using blood-based biomarker data for inflammation, diabetes and blood pressure a number of studies found a negative association with higher socioeconomic position such as higher income or educational attainment (e.g., Banks et al., 2006; Johnston et al., 2009; Jürges et al., 2013; Muennig et al., 2007; Murasko, 2008; Powdthavee, 2010).

[1] Biomarkers are objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention (Biomarkers Definition Working Group, 2001).

A third challenge in income-health studies is how to measure income. Dating back to Friedman (1957), a prolonged discussion about the importance of permanent versus short-term income levels, which can be found in the general economics literature, might be also relevant for health (Fuchs, 2004). According to the "permanent income hypotheses", it might be anticipated that permanent income (as opposed to transitory) may be more relevant as a determinant of the demand for health services (Feldstein, 1966). From a life-course perspective, long-term socioeconomic position may be more relevant to health since it may better reflect cumulative disadvantage (see e.g., Benzeval and Judge, 2001; Singh-Manoux et al., 2004). Moreover, employing measures of income based on long income histories are better indicators of an individuals' economic status since they are less sensitive to temporal income variations, such as a brief spell of unemployment or a short period of hard times, and they help to reduce concerns regarding the role of any potential effects of health shocks on income (Menchik, 1993).

Fourthly, studies of the link between income and health measures typically explore the effect of the former on the conditional mean of the health outcome (for instance, Johnston et al., 2009; Jürges et al., 2013; Muennig et al., 2007; Powdthavee, 2010). However, analyses based solely on the mean may mask important information in other parts of the distribution (Bitler et al., 2006). This is particularly important in the case of the health-income association, where clinical concern is typically focused on the tails of the health distribution (Carrieri and Jones, 2016). For instance, individuals with higher income who experience ill-health may be more likely to initiate behavioural adjustments. In this context, Carrieri and Jones (2016) and Jolliffe (2011) explore the association of income with selected blood-based biomarkers and adiposity measures using quantile regression techniques in order to estimate how the income gradients may vary at different points of the distribution of biomarkers. Hence, evaluating the income gradients at different points of the health distribution may be beneficial.

In this paper we seek to address all of these concerns by using both self-reported health outcomes and nurse measured and blood-based biomarkers in an analysis that compares short-run and long-run measures of income and evaluates the health-income gradient at the mean and across the full distribution of the outcomes. The absorption of the British Household Panel Survey (BHPS) into the Understanding Society (the UK Household Longitudinal Survey, UKHLS) gives us the rare opportunity of combining cross-sectional (UKHLS wave 3 data) and long-running (up to a maximum of 18 BHPS waves) longitudinal

household income data with a large set of self-reported and objective health measures. We estimate short- and long-run income gradients in health in order to explore the relative importance of "permanent" versus current measures of income. In subsequent analysis, a two-step approach is used to account for selection effects due to the time-invariant unobserved heterogeneity when estimating the long-run income gradients in health.

Our paper contributes to the literature in a number of ways. Firstly, we believe that using biomarkers in the analysis of the income gradient in health has its virtues: a) compared to the conventional self-reported health measures, biomarkers are objective measures of health; b) they provide direct information on pre-disease mechanisms that are below the individual's threshold of perception or clinical diagnosis thresholds and, thus, allow for a better understanding of the income-health gradient when diseases have not yet become explicit; and c) they give useful insights about the causal physiological pathways in the complex relationship between socioeconomic status and health, since they are more proximal outcomes compared to SAH.

The most popular biomarkers from the health economics literature are used in this study: adiposity measures, blood pressure, resting heart rate, inflammatory biomarkers, blood glucose (HbA1C) and cholesterol ratio (Carrieri and Jones, 2016; Frankenberg et al., 2016; Johnston et al., 2009). These biomarkers capture different health dimensions and are considered as "secondary" physiological responses to stress and, thus, they are more proximal outcomes in the process through which economic status get "under the skin" (Glei et al., 2013; Turner et al., 2016)[2]. Self-reported health measures may be subject to reporting bias that is correlated with important determinants of health. However, they have been shown to be predictors of future mortality, even after accounting for more objective health measures (Idler and Benyamini, 1997; Jürges et al., 2013; Jylhä, 2009). Complementary to the objective measures, therefore, we also consider three self-reported health measures: SAH, functional disabilities and physical-health functioning (PCS-12). Identifying differences between self-reported and objective health measures that may be driven by reporting heterogeneity in the self-reported health could be of particular importance.

[2] These biomarkers are used by the medical literature to construct allostatic load, i.e. a measure of the wear and tear on the body reflecting the physiological consequences of exposure to stress (Turner et al., 2016). However they are used separately in this study to capture different dimensions of health.

Secondly, capitalizing on the richness of the data we explore associations of health with contemporaneous income as well as relatively long (a maximum of 18 waves) longitudinal income histories; this facilitates identification of the relative importance of the short-run versus long-run measures of income for health. Long-run income is defined by calculating the within-individual mean of the household income over the available time period. In subsequent analysis, a two-step estimator is used to account for selection effects due to the time-invariant unobserved heterogeneity that may be associated with both current health and long-run income. This approach involves the estimation of a longitudinal fixed-effects model for household income in the first stage followed by using the *predicted* individual-specific fixed effects in the health regression models on our measure of long-run income.

In this context, this is the first study, to our knowledge, that estimates the association of biomarkers with both cross-sectional and long-run income measures after accounting for individual-specific effects as well as employing econometric techniques that facilitate "beyond the mean" analysis. Over and above the conventional ordinary least square (OLS) models, unconditional quantile regression (UQR) techniques are used for the case of our continuous health measures. Building on recent work from Firpo et al. (2009), UQR models - based on the recentered influence function (RIF) approach - are used to estimate the income gradient at different points of the *unconditional* distribution of each biomarker of interest. Exploring the health-income gradients "beyond the mean" is particularly important given the heterogeneity of the health risks across the distribution of biomarkers. This allows us to assess whether the association is more evident at the higher quantiles of the biomarker distribution where elevated risks are prominent, implying greater illnesses for individuals and possibly high costs for the health care system.

Results from our cross-sectional regressions of health on household income indicate the presence of clear income gradients across all the self-reported health measures and most of the nurse-administered and blood-based biomarkers. Analysis "beyond the mean" shows that the cross-sectional income gradient is substantially larger in the upper tail of the distribution of our continuous health measures, corresponding to higher health risks. This is particularly true for self-reported physical health functioning measures (PCS-12) and for biomarkers of adiposity (BMI, WC), heart rate, inflammation (CRP), diabetes and cholesterol. However, we find that the long-run income gradients are much greater in magnitude and more statistically significant than those based on the cross-sectional income measure. Moreover,

the corresponding UQR results reveal the presence of greater heterogeneity in the income gradients, being larger in magnitude and following a more clearly increasing pattern towards the tails of the distribution that correspond to higher health risks, when long-term average income is used as opposed to cross-sectional income. Further analysis allows us to disentangle the role of long-run income from individual-specific selection effects, using a two-step estimator that includes the estimated individual-specific selection effects from panel data regressions of household income within the health outcome regressions. Our results show a clear distinction between the self-reported and the objectively measured biomarkers. We find that selection effects due to the time-invariant unobserved heterogeneity are highly significant in the case of the self-reported health outcomes but not in the case of the objectively measured health indicators. This may suggest that reporting heterogeneity in self-reported health, assumed to be driven by individual-specific characteristics that are correlated with socio-economic status and income, may bias the income-health gradients in the case of the pertinent health measures.

The rest of the paper is organized as follows. Section 2 presents our empirical methodology, Section 3 introduces the data and Section 4 presents the results of the study. The final section summarizes and concludes.

2. Methods

In this section we present the empirical strategy that we employ in this study. We first give a brief illustration of the regression models that we use to explore the association between income and our different health measures. This is followed by a presentation of the health specifications for the case of short-run and long-run income measures and the method employed to account for selection-effects due to the time-invariant unobserved heterogeneity that may be associated with both income and health measures.

2.1 Health outcome regression models

Ordered probit models and probit models are used to test the association of short-run and long-run household income with SAH and functional difficulties, respectively. The continuous health measures (PCS-12, nurse-measured and blood-based biomarkers) are

initially modelled using the conventional linear regression model (OLS); in subsequent analysis, quantile regression techniques are employed to explore the whole distribution of the health measures. In this context, a general model specification can be written as:

$$H_i^* = \gamma' I_i + \delta' z_i + u_i \quad (1)$$

where, H_i^* stands for the health outcome of interest, I_i represents the household income variable, z_i stands for the covariates and γ and δ are the regression coefficients to be estimated. In the case of the continuous health outcomes (OLS models), H_i^* coincides with the observed health measure (H_i). Regarding the probit models for functional difficulties and ordered probit models for SAH, H_i^* stands for latent variable.

We also apply quantile regression techniques that allow us to consider the entire distribution of the continuous health outcomes and to investigate the potentially differential effect of household income across different points of their distribution. UQR models are employed in this study (Firpo et al., 2009). Unlike the conventional conditional quantile regression models, which explore the effect of covariates on the conditional quantiles of the outcome variable (Koenker and Bassett, 1978), the UQR technique estimates *unconditional* quantile partial effects.

The estimation of the UQR is based on the RIF. This can be estimated directly from the data by computing sample quantiles of the health measure (q_τ) and then estimating the density of the distribution of health measures at that quantiles using kernel density methods. Specifically, for an observed quantile (q_τ), a RIF is generated which can take one of two values depending upon whether or not the observation's value of the health measure is less than or equal to the observed quantile (q_τ):

$$RIF(H_i; q_\tau) = q_\tau + \frac{\tau - 1[H_i \leq q_\tau]}{f_H(q_\tau)} \quad (2)$$

where, q_τ is the observed sample quantile, $1[H_i \leq q_\tau]$ is an indicator that equals to one if the observation value of the health measure of interest is less than or equal to the observed quantile q_τ and zero otherwise. $f_H(q_\tau)$ is the estimated kernel density of the particular health measure at the τ^{th} quantile. The RIF is then regressed on a set of covariates z_i using OLS; this

constitutes a rescaled linear probability models. We use the bootstrap method with 500 replications to obtain unbiased estimates of the variance-covariance matrix of the parameter estimates (Buchinsky, 1998; Jolliffe, 2011).

2.2 Model specifications

Cross-sectional regressions of health on income (*specification 1*) are initially estimated using current household income (collected at UKHLS wave 3). We then enhance this cross-sectional approach by using a long-run average measure of household income (within-individual mean of the natural logarithm of the household income) derived from the longitudinal income histories covering a long period (maximum of 18 BHPS waves) prior to the health outcomes (*specification 2*). In subsequent analysis, a two-step approach is used to account for selection effects due to the time-invariant unobserved heterogeneity (*specification 3*). This analysis allows us to separate the role of individual-specific selection effects from that of our measure of "permanent" income[3].

2.2.1 Two-step approach and longitudinal household income model

To account for individual-specific selection effects (due to the time-invariant unobserved heterogeneity) in the case of the long-run income gradients in health, we adopt a two-step estimation approach, with a fixed effects income model used in the first-stage.[4] Specifically, the availability of long-running (up to a maximum of 18 BHPS waves) longitudinal household income data facilitates the estimation of a fixed effect model for household income; this model allow for disentangling the time-invariant unobserved individual heterogeneity whose correlation with health may lead to endogeneity concerns. For each individual i at time period (wave) t, the household income equation can be specified as follows:

$$\ln(Y_{it}) = \beta' x_{it} + v_i + \varepsilon_{it} \qquad (i = 1, \dots, N; t = 1, \dots, T_i) \tag{3}$$

[3] A number of studies have used comparable approaches to proxy "permanent income" in the context of other economic research fields (Bhalla, 1980; Mincer, 1962).

[4] Originally, two-step residual inclusion estimators have been devolved to address endogeneity in the context of instrumental variable models (Hausman, 1978). The two-step residual inclusion and the two-stage least squares estimator are identical for the case of linear models and, thus, both consistent (Terza et al., 2008). For the needs of our paper, we use a variant of the two-step residual inclusion estimator technique that limited to allow for the time-invariant unobserved heterogeneity (i.e., the individual-specific selection effects from our first-stage fixed effects income estimator, whose correlation with health outcomes may lead to endogeneity) to be taken into account in the estimation of long-run income gradients in health.

where, Y_{it} is the equivalised deflated household income, x_{it} is a vector of explanatory variables (mainly describing individual i and their household) and β is a vector of regression coefficients to be estimated. The error term has two components: v_i, which represents time-invariant individual-specific effects, and ε_{it}, which is a randomly distributed idiosyncratic error term. v_i can be correlated with the covariates, while ε_{it} is assumed to be uncorrelated across individuals and waves as well as strictly exogenous. The fixed effects (v_i) capture time-invariant unobserved characteristics that affect household income and can be obtained as $\hat{v}_i = \overline{\ln(Y_i)} - \hat{\beta}'\bar{x}_{it}$.

At the second stage of our two-step approach, $\hat{v}_i$, is included in our health outcome regressions (eq. 1) as an additional regressor. In the context of our general health outcomes model specification (eq. 1), the role of selection due to unobserved heterogeneity that is correlated with both long-run income and the health outcomes can be modelled through a common factor structure, where:

$$u_i = \theta\hat{v}_i + \kappa_i \quad (4)$$

and κ_i is the new idiosyncratic error term of the health outcome models.

3. The UKHLS and BHPS datasets

The data come from the BHPS sub-sample of UKHLS. UKHLS is a large, national representative longitudinal study of the members of about 32,000 households (at wave 1) in the UK (Knies, 2015). At wave 2 (2010-2011), the sample of around 8,000 households from the BHPS was absorbed into the UKHLS. The BHPS is a widely used representative longitudinal UK study that covered the period between 1991 and 2009 (18 waves) up to the time it was incorporated in the UKHLS.

For the BHPS respondents followed up in the UKHLS, a set of objective health measures as well as a non-fasted blood sample were collected by trained nurses, as part of the UKHLS wave 3 main survey (Benzeval et al., 2014; McFall et al., 2014). All the other contemporaneous information (such as self-reported health measures, socioeconomic

characteristics etc.) was collected as part of the UKHLS wave 3 main survey. Longitudinal household income histories (and other covariates that are used for the panel household income models) are extracted from BHPS waves 1-18. In order to exploit all the available observations, an unbalanced sample of the BHPS waves 1-18 is employed for the estimation of the longitudinal household income model.

For our study we merge BHPS waves 1-18 with UKHLS wave 3 data for the BHPS sample members who were followed up and were interviewed at least once during the BHPS waves 1-18. When available, the relevant biomedical data from the nurse visits that followed UKHLS wave 3 are also included. Respondents were eligible for the nurse visits if they took part in the main survey, were aged 16+, lived in Great Britain (not Northern Ireland), and were not pregnant (McFall et al., 2014). Blood sample collections were further restricted to those who had no clotting or bleeding disorders and had never had a fit (Benzeval et al., 2014). The resulting potential sample for the longitudinal household income analysis has 195,176 observations across all the eighteenth BHPS waves (25,804 unique individuals). Of those, 8,086 also participated in the UKHLS wave 3, while 4,512 took part in the nurse visits. The blood-based biomarker data are available for 3,054 respondents[5].

3.1 Health Measures

In addition to the conventional SAH and functional disability measures, we also use a set of continuous health indicators. The latter are constituted by a self-reported physical health functioning measure (PCS-12) and a number of objectively measured health indicators: adiposity measures and biomarkers that derived from nurse-administered measurements or analysis of blood samples. Continuous scale health measures allow for exploring the income gradients across different points of the distribution of the pertinent health measure.

Self-reported Health

Three self-reported health measures are used. SAH categorizes respondents on a five-category scale, ranging from "excellent" (value of 1) to "poor" (value of 5) health. We also consider a self-reported functional disability measure. A dichotomous variable is constructed taking the value of one if the respondent reported any long-standing functional difficulty with

[5] Comparison of the summary statistics across different samples reveals similar results (Table A1, appendix), indicating that the implications of the reduction in the sample size (in the case of the nurse visits and the blood data) are limited in our analysis.

any domain of life and zero otherwise. The SF-12 is a self-administered measure of health-related quality of life. For this study, we use the physical component sub-measure (PCS-12). By definition, PCS-12 scores have values between zero and 100 and are standardized to have a mean of 50 and a standard deviation of 10. To facilitate consistency with the interpretation of our results since we intent to measure ill-health, PCS-12 is inverted such as higher values indicate worse physical health functioning.

Adiposity Measures

Anthropometrics were measured during the nurse visits (McFall et al., 2014). We employ waist circumference (WC), to capture central adiposity, in addition to the conventional BMI. The mean of the WC measurements (the two closest, if there were three) is used for the purpose of our study (Davillas and Benzeval, 2016). Body weight and height are used to calculate BMI as the weight (in kilograms) over the square of height (in meters). It has been shown that there is a J-shaped association of BMI and WC with mortality risks; mortality risk is elevated for the underweight and it gradually increases with higher levels of BMI and WC (Pischon et al, 2008; Prospective Studies Collaboration, 2009).

Blood pressure and resting health rate

Systolic blood pressure (SBP) is the maximum pressure in an artery at the moment when the heart is pumping blood; diastolic blood pressure (DBP) is the lowest pressure in an artery in the moments between beats when the heart is resting. A large body of medical studies have demonstrated that the cardiovascular morbidity and mortality risks gradually increase with higher levels of SBP and DBP (Sesso et al., 2000). Heart rate is an overall measure of heart function and cardiovascular fitness. Heart rate values above 90 heart beats per minute (bpm) are indicative for excess health risks (Seccareccia et al., 2001).

Blood-based biomarkers

Two biomarkers of inflammation are examined: CRP and fibrinogen. CRP (in mg/ L) is an acute phase protein that mainly reflects general chronic or systemic inflammation. It has been shown that the risk of ischaemic vascular disease, metabolic syndrome and mortality are gradually increasing in CRP (Emerging Risk Factors Collaboration, 2010). A number of cut-points are used in the medical literature with CRP values over 5 mg/L considered as elevated, while CRP over 3 mg/L as a high risk for cardiovascular diseases (Emerging Risk Factors Collaboration, 2010; Ferrari et al., 2015); values over 10 mg/L are regarded as suggestive of

acute infections (Ishii et al., 2012). Fibrinogen (in g/L) is a glycoprotein that stops bleeding by helping blood clots to form. As such, fibrinogen is directly related to coronary artery thrombosis; however, it is also regarded as an inflammatory biomarker. There is an approximately log-linear association of fibrinogen levels with cardiovascular conditions and mortality (Fibrinogen Studies Collaboration, 2005).

Glycated haemoglobin (HbA1c) is a validated diagnostic test for diabetes (WHO, 2011). Implications for health are not homogenous across the distribution of HbA1c since different levels suggest distinct conditions and severity. HbA1c levels between 42 mmol/mol and 48 mmol/mol indicate pre-diabetes risk, HbA1c $\geq$ 48 mmol/mol indicates diagnosis of diabetes, and higher HbA1c levels suggestive of more severe conditions (WHO, 2011).

Cholesterol concentrations measure the "fat in the blood". For this study, the cholesterol ratio is calculated as the ratio of total cholesterol over high-density lipoprotein cholesterol. This is a stronger predictor of cardiovascular morbidity and mortality risks, than each of the individual cholesterol concentrations, with a dose-response association (Prospective Studies Collaboration, 2007).

3.2 Household income variables

The monthly gross household income is used as the dependent variable in the panel household income model for BHPS waves 1-18. Current household income (i.e. UKHLS wave 3) is available as a derived variable in UKHLS. The household income variables are transformed to natural logarithms in order to allow for the concavity of the health-income associations (e.g. Contoyannis et al., 2004a) and because of the skewness of the income distribution. To facilitate comparisons over time and between households, household income is deflated, using the Retail Price Index, to express income in January 2010 prices and equivalised (using the modified OECD scale).

3.3 Other covariates

Longitudinal income regression models

In keeping with previous studies that model household income, a number of household and individual level covariates are also included (Cappellari and Jenkins, 2002 and 2004).[6] Gender, ethnicity and age (age group dummies for five years intervals between 15 and 84 and a dummy for those over 84) of the HoH are included. We also account for education and job status of the HoH in order to capture important demographic and socioeconomic characteristics that may affect the welfare of the household members (Devicienti, 2011). Measures of the composition and the labour market attachment of the household (number of family members working) are also included since they are either directly associated with earned income or linked to the demographic composition of the household (Cappellari and Jenkins, 2004). Standard regional dummies are also included.

In order to account for any potential effect of individual's health on household income (Michaud and Van Soest, 2008) we control for SAH and dummy indicators for having any long-lasting health problem/disability and health-related limitations on daily activities[7]. Respondent age (dummies defined analogously to the case of HoH) is also included. A vector of wave dummies is added to account for aggregate income shocks that are not captured by deflated income as well as for time-varying reporting changes. Descriptive statistics for these variables are presented in Table A2 (Appendix).

Health outcome regression models

The covariates (collected during the UKHLS wave 3) that are used to model our health outcomes are presented in Table A1 (Appendix), along with summary statistics. The estimation models include fourteen age dummies for each gender (as already described, although the two youngest categories are grouped due to sample size), to allow for a flexible association between health, age and gender. Ethnicity dummies are also included in the health models. We include marital status since it may affect household production of health and demand for health (Fuchs, 2004). Education is also included given evidence on the positive

[6] Although these studies focus on poverty, a similar set of covariates can be used in our analysis since poverty is a function of the continuous household income measures.

[7] Since the wording of the SAH question is different for BHPS wave 9, data from the two adjacent waves are used to impute SAH for BHPS wave 9. Data from adjacent waves are used to impute the missing "health-related limitations" variable at BHPS waves 9 and 14.

association between schooling and health (Contoyannis et al., 2004a). Regional dummies are also added to capture regional variations.

Medications may affect the level of the biomarkers. Following previous literature (Godoy et al., 2007; Powdthavee, 2010; Rahkovsky and Gregory, 2013), we adjust for taking relevant medications. This allows exploration of the health-income gradients on the whole population, controlling for the role of medications. A dummy for anti-hypertensive medications is included in the blood pressure models, while dummies for statins and anti-inflammatory medications are added to the CRP models. Anti-inflammatory medications are also accounted for in fibrinogen models. The cholesterol ratio and the HbA1c regression models include indicators for statins and anti-diabetic medications, respectively.[8]

4. Empirical Results

We present results on income gradients in health using cross-sectional and long-run income measures and the self-reported and objective health indicators. We then present the findings from our two-step approach. To save space, the results from the longitudinal fixed-effects model for household income, serving as the first-stage estimation to obtain the *predicted* individual-specific fixed effects (sub-section 2.2.1), are presented in the Appendix (Section B).

4.1 Income gradients in health using cross-sectional versus long-run income measures

Income gradients for our different health indicators using cross-sectional (*specification 1*) and long-run (within-individual mean of the natural logarithm of the household income over up to 18 BHPS waves) household income measures (*specification 2*) are presented in Tables 1-6. Our results are broadly in accordance with previous studies that have found a stronger association between long-run income measures and self-reported measures of health and disability compared to short-run income measures (e.g., Benzeval et al., 2000, Contoyannis et al., 2004a,b) and extends them to a set of objective health measures, employing "beyond the mean" analytical techniques.

[8] Since an unbalanced BHPS sample is used, we also account for the number of BHPS waves that each individual is observed in the case of the regressions of health on long-run income (Verbeek and Nijman, 1992).

Tables 1 and 2 present the results for our self-reported health measures. As expected, we find a strong cross-sectional income gradient in both SAH and functional disability; higher income is related to a better SAH (i.e., lower SAH values since SAH is coded from excellent [1] to poor health [5]) and to a lower probability of functional disability (*specification 1*, Table 1). There is also a negative association between higher cross-sectional income and poor physical health functioning (inverted PCS-12; *specification* 1, Table 2). The UQR estimates show that the gradient is more evident beyond the median of the PCS-12 distribution (corresponding to lower physical functioning). For instance, the income gradient at the 10th percentile is about five times higher than that at the 75th percentile (-1.922 vs. -0.391).

Table 1. Income gradients in self-assessed health and functional disabilities using cross-sectional and long-run income measures.

Panel A: Self-assessed health						
	Ordered Probit					
	Coeff. (s.e.)†	APE (s.e.)†				
		Excellent	Very good	Good	Fair	Poor
Specification 1						
Ln(current income)	-0.233***	0.052***	0.033***	-0.021***	-0.037***	-0.027***
	(0.024)	(0.005)	(0.003)	(0.002)	(0.004)	(0.003)
Specification 2						
Long-run mean ln(income)	-0.406***	0.090***	0.057***	-0.035***	-0.063***	-0.048***
	(0.029)	(0.006)	(0.004)	(0.003)	(0.005)	(0.004)
Specification 3						
Long-run mean ln(income)	-0.687***	0.152***	0.095***	-0.059***	-0.107***	-0.080***
	(0.061)	(0.013)	(0.008)	(0.005)	(0.009)	(0.008)
Individual-specific effects	0.387***	-0.086***	-0.054***	0.033***	0.061***	0.045***
	(0.071)	(0.016)	(0.010)	(0.006)	(0.011)	(0.009)
Sample size	7,978					
Panel B: Functional disabilities						
	Probit					
	Coeff. (s.e.)†	APE (s.e.)†				
Specification 1						
Ln(current income)	-0.176***	-0.048***				
	(0.032)	(0.009)				
Specification 2						
Long-run mean ln(income)	-0.388***	-0.105***				
	(0.041)	(0.011)				
Specification 3						
Long-run mean ln(income)	-0.954***	-0.256***				
	(0.086)	(0.022)				
Individual-specific effects	0.782***	0.210***				
	(0.100)	(0.027)				
Sample size	7,724					

Abbreviations: APE, average partial effects; Coeff., coefficients; s.e., standard errors.
† Robust standard errors in parenthesis.
***P<0.01; **P<0.05; *P<0.10

Employing our long-run income measure (proxy of "permanent income") we find much higher income gradients (*specification 2*) than those based on cross-sectional income (*specification 1*). Specifically, the long-run income gradients in SAH and the functional disability (average partial effects) are 1.6 to 2 times higher than those based on cross-

sectional income. Moreover, long-run income, compared to current income, exhibits greater heterogeneity in the income gradients in (inverted) PCS-12 (Table 2, *specification 2*). Specifically, the long-run income gradient at the 90th percentile is about 8 times higher than at the 10th (inverted) PCS-12 percentile.

Table 2. Income gradients in (inverted) PCS-12 using cross-sectional and long-run income measures.

	OLS	**Unconditional quantile regressions**					
		Q10	Q25	Q50	Q75	Q90	Q95
	Coeff. (s.e.)†	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡
Specification 1							
Ln(current income)	-1.310***	-0.391**	-0.916***	-1.523***	-1.922***	-1.696**	-1.145
	(0.231)	(0.166)	(0.157)	(0.249)	(0.516)	(0.725)	(0.735)
Specification 2							
Long-run mean ln(income)	-3.058***	-0.799***	-0.885***	-2.292***	-5.409***	-6.510***	-6.100***
	(0.294)	(0.196)	(0.182)	(0.299)	(0.710)	(1.039)	(1.037)
Specification 3							
Long-run mean ln(income)	-6.305***	-0.771*	-1.032***	-3.947***	-11.62***	-17.17***	-13.45***
	(0.641)	(0.402)	(0.361)	(0.631)	(1.377)	(2.257)	(2.223)
Individual-specific effects	4.465***	-0.0387	0.202	2.276***	8.545***	14.66***	10.11***
	(0.733)	(0.465)	(0.438)	(0.758)	(1.595)	(2.523)	(2.323)
Sample size				7,048			

Abbreviations: Coeff., coefficients; OLS, ordinary least squares; s.e., standard errors.
† Robust standard errors in parenthesis.
‡ UQR standard errors are bootstrap estimates with 500 replications.
***P<0.01; **P<0.05; *P<0.10

The income gradients in adiposity are presented in Table 3. Again, in the BMI models the income coefficients are larger in magnitude for long-run versus cross-sectional income measures. Although there is no systematic association at the mean (OLS), the long-run income coefficient at the 95th BMI percentile (corresponds to BMI values within the range of severe obesity, i.e. ≥35 kg/m^2; Prospective Studies Collaboration, 2009) is more than five times higher than the corresponding OLS coefficient. The income gradients in WC are more pronounced. This is broadly in accordance with previous studies that found stronger socioeconomic gradients for central adiposity measures rather than BMI, reflecting the fact that BMI is a noisy adiposity measure that cannot distinguish fat from lean body mass (Davillas and Benzeval, 2016; Ljungvall et al., 2015). Income gradients in WC are also higher in magnitude in the case of long-run versus current income measures, notably at the right tails of the WC distribution. A closer look at the long-run income gradients (*specification 2*) reveals that the OLS estimator averages out notable differences across the WC distribution. Specifically, the UQR models suggest no systematic income gradients at the lower percentiles of WC, while there are statistically significant and gradually increasing income gradients at higher percentiles (up to three times larger than the OLS estimates).

Table 3. Income gradients in adiposity using cross-sectional and long-run income measures.

	OLS	Unconditional quantile regressions¶					
Panel A: Body mass index							
		Q10 (22.0 kg/m^2)	Q25 (24.4 kg/m^2)	Q50 (27.4 kg/m^2)	Q75 (31.1 kg/m^2)	Q90 (35.2 kg/m^2)	Q95 (38.3 kg/m^2)
	Coeff. (s.e.)†	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡
Specification 1							
Ln(current income)	-0.091	0.099	-0.026	-0.078	-0.059	-0.282	-0.757*
	(0.154)	(0.197)	(0.187)	(0.199)	(0.256)	(0.367)	(0.430)
Specification 2							
Long-run mean ln(income)	-0.254	0.228	-0.115	-0.284	-0.476	-0.743	-1.414**
	(0.194)	(0.239)	(0.230)	(0.230)	(0.315)	(0.457)	(0.579)
Specification 3							
Long-run mean ln(income)	-0.120	0.591	0.148	-0.132	-0.160	-1.591	-2.755**
	(0.422)	(0.502)	(0.465)	(0.444)	(0.634)	(1.040)	(1.381)
Individual-specific effects	-0.180	-0.487	-0.353	-0.205	-0.423	1.138	1.799
	(0.483)	(0.560)	(0.537)	(0.535)	(0.727)	(1.179)	(1.540)
Sample size				4,224			
Panel B: Waist circumference							
	OLS	**Unconditional quantile regressions¶**					
		Q10 (76.2 cm)	Q25 (84.2 cm)	Q50 (93.7 cm)	Q75 (103.7 cm)	Q90 (113.6 cm)	Q95 (120 cm)
	Coeff. (s.e.)†	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡
Specification 1							
Ln(current income)	-0.820**	-0.780	0.0300	-0.237	-1.310**	-1.821**	-2.110**
	(0.377)	(0.563)	(0.546)	(0.491)	(0.590)	(0.815)	(0.955)
Specification 2							
Long-run mean ln(income)	-1.377***	-0.543	-0.322	-0.682	-1.630**	-3.020***	-4.302***
	(0.493)	(0.770)	(0.670)	(0.614)	(0.689)	(1.036)	(1.251)
Specification 3							
Long-run mean ln(income)	-1.528	-0.847	0.758	-0.475	-0.716	-4.024	-8.187***
	(1.068)	(1.706)	(1.454)	(1.355)	(1.556)	(2.468)	(3.109)
Individual-specific effects	0.203	0.409	-1.451	-0.279	-1.228	1.350	5.222
	(1.202)	(1.920)	(1.673)	(1.585)	(1.851)	(2.825)	(3.375)
Sample size				4,372			

Abbreviations: Coeff., coefficients; OLS, ordinary least squares; s.e., standard errors.
¶ Body mass index and waist circumference values that correspond to each percentile of the distribution are also presented.
† Robust standard errors in parenthesis.
‡ UQR standard errors are bootstrap estimates with 500 replications.
***P<0.01; **P<0.05; *P<0.10

Table 4 presents the results for blood pressure and heart rate. There is some evidence of considerably higher income gradients at the right tails of the distribution for both systolic and diastolic blood pressure, especially in the case of long-run income measures, albeit only statistically significant at the 10% level. More pronounced income gradients are evident for our measure of overall cardiovascular fitness (heart rate). Long-run income gradients in heart rate are larger in magnitude and increase towards the right tail of the heart rate distribution than those for current income. No systematic gradients are found at the lowest percentiles of the heart rate distribution (up to 60 bmp; most likely reflecting athletic lifestyles), whereas gradually increasing negative income gradients are evident towards the higher percentiles, with a peak at the 95th percentile that is close to the clinical threshold for elevated health risks (>90 bmp; Seccareccia et al., 2001).

Table 4. Income gradients in blood pressure and heart rate measurements using cross-sectional and long-run income measures.

	Panel A: Systolic blood pressure						
	OLS	**Unconditional quantile regressions¶**					
		Q10 (106 mmHg)	**Q25** (114.5 mmHg)	**Q50** (124.5 mm Hg)	**Q75** (136.5 mm Hg)	**Q90** (148.5 mm Hg)	**Q95** (156.5 mmHg)
	Coeff. (s.e.)†	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡
Specification 1							
Ln(current income)	-0.687	-0.508	-0.575	-0.211	-0.587	-0.613	-2.061
	(0.470)	(0.628)	(0.576)	(0.633)	(0.839)	(0.978)	(1.558)
Specification 2							
Long-run mean ln(income)	-0.050	1.023	-0.144	1.008	0.547	-0.610	-4.238*
	(0.608)	(0.812)	(0.773)	(0.765)	(0.975)	(1.396)	(2.390)
Specification 3							
Long-run mean ln(income)	1.025	0.642	-0.824	2.300	1.548	0.917	-2.729
	(1.291)	(1.904)	(1.630)	(1.686)	(2.215)	(2.552)	(4.164)
Individual-specific effects	-1.423	0.506	0.902	-1.711	-1.326	-2.023	-2.000
	(1.508)	(2.277)	(1.924)	(1.979)	(2.683)	(3.029)	(5.165)
Sample size				3,632			
	Panel B: Diastolic Blood pressure						
	OLS	**Unconditional quantile regressions¶**					
		Q10 (59.5 mmHg)	**Q25** (65.5 mmHg)	**Q50** (73 mm Hg)	**Q75** (80 mmHg)	**Q90** (87 mmHg)	**Q95** (91.5 mmHg)
	Coeff. (s.e.)†	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡
Specification 1							
Ln(current income)	-0.809**	-0.573	-0.971**	-0.707*	-1.002*	-1.049	-0.861
	(0.326)	(0.448)	(0.456)	(0.430)	(0.520)	(0.693)	(0.978)
Specification 2							
Long-run mean ln(income)	-0.478	0.258	-0.199	0.0385	-1.131*	-1.619*	-1.520
	(0.422)	(0.680)	(0.607)	(0.545)	(0.665)	(0.886)	(1.114)
Specification 3							
Long-run mean ln(income)	-0.744	0.546	-1.269	0.090	-1.217	-0.647	-3.170
	(0.903)	(1.501)	(1.317)	(1.159)	(1.341)	(1.811)	(2.380)
Individual-specific effects	0.352	-0.382	1.418	-0.0683	0.114	-1.288	2.186
	(1.053)	(1.816)	(1.570)	(1.402)	(1.568)	(2.015)	(2.567)
Sample size				3,632			
	Panel C: Resting Heart rate						
	OLS	**Unconditional quantile regressions¶**					
		Q10 (56 bmp)	**Q25** (61.5 bmp)	**Q50** (68.5 bmp)	**Q75** (75.5 bmp)	**Q90** (84 bmp)	**Q95** (89 bmp)
	Coeff. (s.e.)†	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡
Specification 1							
Ln(current income)	-1.425***	-0.673	-0.743*	-1.481***	-1.463***	-3.476***	-2.934***
	(0.348)	(0.535)	(0.434)	(0.467)	(0.534)	(0.848)	(0.869)
Specification 2							
Long-run mean ln(income)	-1.742***	-0.391	-0.992*	-1.712***	-2.118***	-4.078***	-5.061***
	(0.455)	(0.652)	(0.544)	(0.587)	(0.634)	(1.081)	(1.226)
Specification 3							
Long-run mean ln(income)	-1.755*	0.937	0.0821	-2.533**	-2.603*	-5.018**	-6.567***
	(1.015)	(1.352)	(1.113)	(1.204)	(1.433)	(2.489)	(3.001)
Individual-specific effects	0.0172	-1.760	-1.423	1.088	0.643	1.246	2.447
	(1.174)	(1.625)	(1.288)	(1.333)	(1.690)	(2.741)	(3.365)
Sample size				3,636			

Abbreviations: Coeff., coefficients; OLS, ordinary least squares; s.e., standard errors.

¶ Blood pressure and heart rate values that correspond to each percentile of the distribution are also presented.

† Robust standard errors in parenthesis.

‡ UQR standard errors are bootstrap estimates with 500 replications.

***P<0.01; **P<0.05; *P<0.10

Income gradients in inflammatory biomarkers are also higher in magnitude for long-run versus cross-sectional income (Table 5). Analysis "beyond the mean" reveals that the differences between them are more evident at the tails of the CRP distribution, where both higher health risks and steeper income gradients are observed. For example, the long-run income gradient at the 95th percentile (reflecting acute inflammation; Ishii et al., 2012) is about 7 times higher than the OLS coefficient. However, we find limited variation in the magnitude of the negative income gradient in fibrinogen across its distribution; this result is in accordance with previous evidence (Carrieri and Jones, 2016).

Table 5. Income gradients in inflammatory biomarkers using cross-sectional and long-run income measures.

	Panel A: C-reactive protein						
	OLS	**Unconditional quantile regressions¶**					
		Q10 (0.3 mg/L)	**Q25 (0.6 mg/L)**	**Q50 (1.4 mg/L)**	**Q75 (3.1 mg/L)**	**Q90 (6.7 mg/L)**	**Q95 (11 mg/L)**
	Coeff. (s.e.)†	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡
Specification 1							
Ln(current income)	-0.176	0.002	0.001	-0.048	-0.249	-1.195**	-2.151*
	(0.225)	(0.033)	(0.038)	(0.071)	(0.184)	(0.566)	(1.251)
Specification 2							
Long-run mean ln(income)	-0.815***	-0.031	-0.082*	-0.222**	-0.943***	-2.037***	-5.733**
	(0.280)	(0.048)	(0.049)	(0.089)	(0.234)	(0.730)	(2.378)
Specification 3							
Long-run mean ln(income)	-0.764	0.047	-0.008	-0.105	-0.563	-1.343	-12.78**
	(0.636)	(0.077)	(0.104)	(0.192)	(0.515)	(1.531)	(5.473)
Individual-specific effects	-0.0667	-0.103	-0.0972	-0.153	-0.498	-0.911	9.252
	(0.677)	(0.093)	(0.119)	(0.222)	(0.605)	(1.693)	(5.694)
Sample size				2,932			
	Panel B: Fibrinogen						
	OLS	**Unconditional quantile regressions¶**					
		Q10 (2.1 g/L)	**Q25 (2.4 g/L)**	**Q50 (2.8 g/L)**	**Q75 (3.2 g/L)**	**Q90 (3.6 g/L)**	**Q95 (3.8 g/L)**
	Coeff. (s.e.)†	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡
Specification 1							
Ln(current income)	-0.075***	-0.057**	-0.073***	-0.085***	-0.099***	-0.061	-0.046
	(0.020)	(0.025)	(0.025)	(0.025)	(0.031)	(0.045)	(0.059)
Specification 2							
Long-run mean ln(income)	-0.148***	-0.112***	-0.143***	-0.133***	-0.172***	-0.121**	-0.166**
	(0.025)	(0.033)	(0.034)	(0.031)	(0.038)	(0.060)	(0.069)
Specification 3							
Long-run mean ln(income)	-0.200***	-0.061	-0.187***	-0.202***	-0.292***	-0.288**	-0.351**
	(0.055)	(0.071)	(0.070)	(0.066)	(0.081)	(0.135)	(0.175)
Individual-specific effects	0.068	-0.067	0.057	0.090	0.157	0.220	0.244
	(0.063)	(0.082)	(0.084)	(0.080)	(0.098)	(0.154)	(0.203)
Sample size				2,894			

Abbreviations: Coeff., coefficients; OLS, ordinary least squares; s.e., standard errors
¶ Biomarker values that correspond to each percentile of the biomarker distribution are also presented.
† Robust standard errors in parenthesis.
‡ UQR standard errors are bootstrap estimates with 500 replications.
***P<0.01; **P<0.05; *P<0.10

The long-run income measure shows larger income gradients and a sharper increase in the income gradients towards the right tail of the distribution of our "blood sugar" (HbA1c) and

"fat in the blood" (cholesterol ratio) biomarkers than current income measures (Table 6). For HbA1c we find a steeper pattern across the HbA1c distribution for the long-run measure; with the gradient increasing almost linearly up to 90th percentile and much sharper after this point (95th HbA1c percentile, i.e., values close to the diabetes threshold). The long-run income gradient in cholesterol ratio (*specification 2*) gradually increases towards the right tails of its distribution with two "peak" points: the first at around the 75th percentile (close to the clinical threshold of 4; Millán et al., 2009) and another peak afterwards (95th percentile). The corresponding income gradient at the right tail of the cholesterol ratio distribution is 37 times higher compared to the bottom of the distribution (-0.475 vs -0.013).

Table 6. Income gradients in HbA1c and Cholesterol ratio using cross-sectional and long-run income measures.

	Panel A: HbA1c						
	OLS	**Unconditional quantile regressions¶**					
		Q10 (31 mmol/mol)	**Q25 (33 mmol/mol)**	**Q50 (36 mmol/mol)**	**Q75 (39 mmol/mol)**	**Q90 (43 mmol/mol)**	**Q95 (50 mmol/mol)**
	Coeff. (s.e.)†	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡
Specification 1							
Ln(current income)	-0.673***	-0.342*	-0.455**	-0.347**	-0.450**	-0.308	-3.486*
	(0.227)	(0.197)	(0.181)	(0.166)	(0.228)	(0.602)	(1.865)
Specification 2							
Long-run mean ln(income)	-1.248***	-0.274	-0.488**	-0.774***	-0.938***	-1.123*	-5.549**
	(0.366)	(0.278)	(0.247)	(0.211)	(0.297)	(0.667)	(2.850)
Specification 3							
Long-run mean ln(income)	-2.121***	0.157	-1.075	-0.708	-1.347**	-3.698**	-13.390**
	(0.807)	(0.629)	(0.670)	(0.458)	(0.596)	(1.790)	(6.519)
Individual-specific effects	1.145	-0.566	0.771	-0.0867	0.536	3.208	12.222
	(0.855)	(0.700)	(0.699)	(0.534)	(0.736)	(2.234)	(7.333)
Sample size				2,779			
	Panel B: Cholesterol ratio						
	OLS	**Unconditional quantile regressions¶**					
		Q10 (2.35 units)	**Q25 (2.81 units)**	**Q50 (3.5 units)**	**Q75 (4.45 units)**	**Q90 (5.55 units)**	**Q95 (6.3 units)**
	Coeff. (s.e.)†	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡
Specification 1							
Ln(current income)	-0.104**	-0.044	-0.096*	-0.138***	-0.231***	0.070	-0.085
	(0.046)	(0.039)	(0.050)	(0.052)	(0.077)	(0.119)	(0.157)
Specification 2							
Long-run mean ln(income)	-0.229***	-0.013	-0.122**	-0.221***	-0.379***	-0.305**	-0.475**
	(0.056)	(0.050)	(0.059)	(0.065)	(0.098)	(0.143)	(0.197)
Specification 3							
Long-run mean ln(income)	-0.294**	-0.047	-0.128	-0.236*	-0.490**	-0.577*	-1.402***
	(0.131)	(0.124)	(0.119)	(0.140)	(0.209)	(0.328)	(0.536)
Individual-specific effects	0.086	0.045	0.008	0.020	0.147	0.358	1.129*
	(0.156)	(0.142)	(0.141)	(0.159)	(0.242)	(0.388)	(0.637)
Sample size				2,932			

Abbreviations: Coeff., coefficients; OLS, ordinary least squares; s.e., standard errors

¶ Biomarker values that correspond to each percentile of the biomarker distribution are also presented.

† Robust standard errors in parenthesis.

‡ UQR standard errors are bootstrap estimates with 500 replications.

***P<0.01; **P<0.05; *P<0.10

4.2 Accounting for unobserved heterogeneity in the long-run health-income gradients

The analysis so far does not account for the role of selection due to unobserved heterogeneity that may be associated with both health and long-run income. In subsequent analysis (*specification 3,* Tables 1-6), a two-step estimator is implemented by adding selection effects obtained using the fixed effects panel model (results available in Table A3 in the Appendix), as an additional regressor in the health outcome regressions for long-run income.

Overall, we find that accounting for selection effects due to the time-invariant unobserved heterogeneity (*specification 3*) results in a clear distinction between the self-reported and the objectively measured health indicators. Although the individual-specific selection effects are highly statistically significant in the case of our self-reported health models (i.e., SAH, functional disabilities and PCS-12), no such systematic effects are observed for the biomarker models (*specification 3*, Tables 1-6).[9] This suggests endogeneity of long-run income in the health outcome regressions based on self-reported health indicators, while this is not the case for the objectively measured health indicators (at least as far as the time-invariant individual-specific heterogeneity is concerned). This difference might be explained by the fact that the individual-specific selection effects in the self-report measures capture subjective reporting error.

Reporting heterogeneity in self-reported health can occur for a number of reasons, and has been shown to be associated with the respondent's socio-economic status and income (Campolieti, 2002; Johnston et al., 2009; Ziebarth, 2010). In this context, models accounting for unobserved heterogeneity (selection) effects (*specification 3*) are preferred for the self-reported health outcomes. Finally, the fact that we observe larger income gradients in self-reported health (SAH, functional disabilities and PCS-12) in the two-stage estimation models (*specification 3*) compared to those that do not account for selection-effects (*specification 2*) may be primarily because of attenuation bias associated with measurement error in self-reported health measures[10].

[9] To explore whether our results are driven by differences in the sample size, Tables A4 and A5 (Appendix) present income gradients in the self-reported health measures for the nurse visits sample. We find that our results are robust to the choice of alternative samples.

[10] In accordance with our conclusions, Johnston et al. (2009) found that reporting error in self-reported health measures may be result in an underestimation of the income-gradient. Our results indicate that those with higher incomes may have a tendency to be more harsh in evaluations of their own health. Those with higher income may have higher expectations for their own health or perhaps a better understanding of health conditions. It has

5. Conclusions

This paper uses data from the BHPS subsample of the UKHLS which allows for a large set of self-reported and objective measures of health as well as for both short- and long-run measures of income. We use a range of self-reported health measures (SAH, disability and physical-health functioning), nurse-administered (adiposity, blood pressure and heart rate) and blood-based biomarkers (inflammatory, blood sugar and "fat in the blood" biomarkers). The availability of this large set of health measures, in combination with longitudinal income histories, give us the rare opportunity for a detailed econometric analysis that is not limited to cross-sectional income measures but also explores long-run income gradients in health accounting for the role of individual-specific selection effects (time-invariant unobserved heterogeneity). To our knowledge, this is the first study that explores income-health gradients using such a broad set of self-reported and objectively measured health indicators, both short-run and long-run income measures, and estimation techniques that account for individual-specific selection effects and facilitate "beyond the mean" analysis.

Our results show clear income gradients across all the self-reported health measures and most of the nurse-administered and blood-based biomarkers when we use cross-sectional income. We find that the cross-sectional association of current income with self-reported physical health functioning measures (PCS-12) and biomarkers of adiposity (BMI, WC), heart rate, inflammation (CRP), diabetes and cholesterol varies across their distribution and is considerably larger at the tails of the distribution, where the health care risks are more evident. We find greater (in magnitude and statistical significance) income gradients in health when we employ a long-run income measure. Heterogeneity in these income gradients is evident, especially for long-run income, with the gradients larger in magnitude and following a steeper increasing pattern towards the tails of the distribution.

Using a two-step estimator, further analysis allows us to disentangle the role of "permanent" income from that attributed to individual-specific selection effects. Although the individual-specific selection effects are highly statistically significant for our self-reported health models

been shown that similar objective clinical health conditions can be differentially taken into account in self-assessments of health subject to the individual's knowledge of these conditions or symptoms (Jylhä, 2009).

(i.e., SAH, functional disabilities and PCS-12), no systematic effects are observed for the objectively measured health indicators. This may suggest that reporting heterogeneity in self-reported health, assumed to be driven by individual-specific characteristics that are correlated with income, may bias the income-health gradients in the case of the self-reported measures. Our findings suggesting that the income-health gradients may be contaminated by the role of systematic reporting heterogeneity in self-reported health highlighting the importance of considering more accurate measures of health (such as biomarkers) in future research.

References

Bago d'Uva, T., O'Donnell, O., van Doorslaer, E. (2008) Differential health reporting by education level and its impact on the measurement of health inequalities among older Europeans. International Journal of Epidemiology 37(6), 1375–1383

Baker, M., Stabile, M., Deri, C. (2004). What do self-reported, objective, measures of health measure? Journal of Human Resources, 39(4), 1067-1093.

Banks, J., Marmot, M., Oldfield, Z., Smith, J. P. (2006). Disease and disadvantage in the United States and in England. JAMA: the Journal of the American Medical Association, 295(17), 2037-2045.

Benzeval, M., Davillas, A., Kumari, M., Lynn, P. (2014). Understanding Society - UK Household Longitudinal Study: Biomarker User Guide and Glossary. Colchester: University of Essex.

Benzeval, M., Judge, K. (2001). Income and health: the time dimension. Social Science and Medicine, 52 (9), 1371-90.

Bhalla, S. S. (1980). The measurement of permanent income and its application to savings behavior. The Journal of Political Economy, 722-744.

Biomarkers Definition Working Group. (2001). Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. Clinical Pharmacology & Therapeutics 69, 89–95.

Bitler, M.P., Gelbach, J.B., Hoynes, H.W., (2006). What mean impacts miss: Distributional effects of welfare reform experiments. The American Economic Review 96, 988-1012.

Buchinsky, M. (1998). Recent advances in quantile regression models: a practical guideline for empirical research. Journal of Human Resources, 88-126.

Campolieti, M. (2002). Disability and the labor force participation of older men in Canada. Labour economics, 9(3), 405-432.

Cappellari, L., Jenkins, S.P. (2002). Who stays poor? Who becomes poor? Evidence from the British household panel survey. The Economic Journal, 112(478), C60-C67.

Cappellari, L., Jenkins, S.P. (2004). Modelling low income transitions. Journal of Applied Econometrics, 19(5), 593-610.

Carrieri, V., Jones, A.M., (2016). The income-health relationship 'beyond the mean': new evidence from biomarkers, Health Economics, DOI: 10.1002/hec.3372 (Online 15 July 2016).

Chou, S. Y., Grossman, M., Saffer, H. (2004). An economic analysis of adult obesity: results from the Behavioral Risk Factor Surveillance System. Journal of Health Economics, 23(3), 565-587.

Contoyannis, P., Jones, A. M., Rice, N. (2004a). The dynamics of health in the British Household Panel Survey. Journal of Applied Econometrics, 19(4), 473-503.

Contoyannis, P., Jones, A. M., Rice, N. (2004b). Simulation-based inference in dynamic panel probit models: an application to health. Empirical Economics, 29(1), 49-77.

Crossley, T. F., Kennedy, S. (2002). The reliability of self-assessed health status. Journal of Health Economics, 21(4), 643-658.

Daltroy, L. H., Larson, M. G., Eaton, H. M., Phillips, C. B., Liang, M. H. (1999). Discrepancies between self-reported and observed physical function in the elderly: the influence of response shift and other factors. Social Science & Medicine, 48(11), 1549-1561.

Davillas, A., Benzeval, M. (2016). Alternative measures to BMI: Exploring income-related inequalities in adiposity in Great Britain. Social Science & Medicine, 166, 223-232.

Deaton, A. S., Paxson, C. H. (1998). Aging and inequality in income and health. The American Economic Review, 88(2), 248-253.

Devicienti, F. (2011). Estimating poverty persistence in Britain. Empirical Economics, 40(3), 657-686.

Dowd, J.B., Zajacova, A. (2010). Does self-rated health mean the same thing across socioeconomic groups? Evidence from biomarker data. Annals of Epidemiology, 20(10), 743-749.

Dowd, J. B., Simanek, A. M., Aiello, A. E. (2009). Socio-economic status, cortisol and allostatic load: a review of the literature. International Journal of Epidemiology, 38(5), 1297-1309.

Ettner, S. L. (1996). New evidence on the relationship between income and health. Journal of Health Economics, 15(1), 67-85.

Feldstein, P. J. (1966). Research on the demand for health services. The Milbank Memorial Fund Quarterly, 44(3), 128-165.

Ferrari, M., Cuenca-Garcia, M., Valtuena, J., Moreno, L. A., Censi, L., González-Gross, M., Androutsos, O., Gilbert, C.C., Huybrechts, I., Dallongeville, J., Sjöström, M., Molnar, D., De Henauw, S., Gómez-Martínez, S., de Moraes, A.C.F., Kafatos, A., Widhalm, K., Leclercq, C.

(2015). Inflammation profile in overweight/obese adolescents in Europe: an analysis in relation to iron status. European Journal of Clinical Nutrition, 69(2), 247-255.

Fibrinogen Studies Collaboration (2005). Plasma fibrinogen level and the risk of major cardiovascular diseases and nonvascular mortality: an individual participant meta-analysis. JAMA: the Journal of the American Medical Association, 294(14), 1799-1809.

Firpo S, Fortin NM, Lemieux T. 2009. Unconditional quantile regressions. Econometrica 77, 953–973.

Frankenberg, E., Ho, J. Y., Thomas, D. (2016). Biological Health Risks and Economic Development. The Oxford Handbook of Economics and Human Biology, 454.

Friedman, M. (1957). The permanent income hypothesis. In: A theory of the consumption function (pp. 20-37). Princeton University Press.

Fuchs, V.R. (2004). Reflections on the socio-economic correlates of health. Journal of Health Economics, 23(4), 653-661.

Glei, D. A., Goldman, N., Shkolnikov, V. M., Jdanov, D., Shalnova, S., Shkolnikova, M., Weinstein, M. (2013). To what extent do biomarkers account for the large social disparities in health in Moscow? Social Science and Medicine, 77, 164-172.

Godoy, R., Goodman, E., Gravlee, C., Levins, R., Seyfried, C., Caram, M., Jha, N. (2007). Blood pressure and hypertension in an American colony (Puerto Rico) and on the USA mainland compared, 1886–1930. Economics & Human Biology, 5(2), 255-279.

Hausman, J. A. (1978). Specification tests in econometrics. Econometrica: Journal of the Econometric Society, 1251-1271.

Idler, E.L., Benyamini, Y. (1997). Self-rated health and mortality: a review of 27 community studies. Journal of Health and Social Behavior, 38:21-37.

Ishii, S., Karlamangla, A. S., Bote, M., Irwin, M. R., Jacobs Jr, D. R., Cho, H. J., Seeman, T.E. (2012). Gender, obesity and repeated elevation of C-reactive protein: data from the CARDIA cohort. PLoS One, 7(4), e36062.

Johnston, D.W., Propper, C., Shields, M.A. (2009). Comparing subjective and objective measures of health: Evidence from hypertension for the income/health gradient. Journal of Health Economics, 28(3), 540-552.

Jolliffe, D. (2011). Overweight and poor? On the relationship between income and the body mass index. Economics & Human Biology, 9(4), 342-355.

Jones, A.M., Wildman, J. (2008). Health, income and relative deprivation: Evidence from the BHPS. Journal of Health Economics, 27(2), 308-324.

Jürges H. (2007) True health vs response styles: exploring cross-country differences in self-reported health. Health Economics 16(2), 163–178

Jürges H. (2008) Self-assessed health, reference levels, and mortality. Applied Economics 40(5), 569–582

Jürges, H., Kruk, E., Reinhold, S. (2013). The effect of compulsory schooling on health-evidence from biomarkers. Journal of Population Economics, 26(2), 645-672.

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

Knies, G. (ed.) (2015). Understanding Society –UK Household Longitudinal Study: Wave 1-5, 2009-2014, User Manual. Colchester: University of Essex.

Koenker, R., Bassett, G. (1978). Regression quantiles. Econometrica: journal of the Econometric Society, 33-50.

Ljungvall, Å., Gerdtham, U. G., Lindblad, U. (2015). Misreporting and misclassification: implications for socioeconomic disparities in body-mass index and obesity. The European Journal of Health Economics, 16(1), 5-20.

McFall, S.L., Petersen, J., Kaminska, O., Lynn, P. (2014). Understanding Society –UK Household Longitudinal Study: Waves 2 and 3 Nurse Health Assessment, 2010-2012, Guide to Nurse Health Assessment. Colchester: University of Essex.

Menchik, P.L. (1993). Economic status as a determinant of mortality among black and white older men: does poverty kill? Population Studies, 47(3), 427-436.

Michaud, P.C., Van Soest, A. (2008). Health and wealth of elderly couples: Causality tests using dynamic panel data models. Journal of Health Economics, 27(5), 1312-1325.

Millán, J., Pintó, X., Muñoz, A., Zúñiga, M., Rubiés-Prat, J., Pallardo, L. F., Masana L., Mangas, A., Hernández-Mijares, A., González-Santos, P., Ascaso, J.F., Pedro-Botet, J. (2009). Lipoprotein ratios: physiological significance and clinical usefulness in cardiovascular prevention. Vascular Health and Risk Management, 5, 757.

Mincer, J. (1962). Labor force participation of married women: A study of labor supply. In Aspects of labor economics (pp. 63-105). Princeton University Press.

Muennig, P., Sohler, N., Mahato, B. (2007). Socioeconomic status as an independent predictor of physiological biomarkers of cardiovascular disease: evidence from NHANES. Preventive medicine, 45(1), 35-40.

Murasko, J.E. (2008). Male–female differences in the association between socioeconomic status and atherosclerotic risk in adolescents. Social Science & Medicine, 67(11), 1889-1897.

Pischon, T., Boeing, H., Hoffmann, K., Bergmann, M., Schulze, M. B., Overvad, K.,... Halkjaer, J. (2008). General and abdominal adiposity and risk of death in Europe. New England Journal of Medicine, 359(20), 2105-2120.

Powdthavee, N. (2010). Does education reduce the risk of hypertension? Estimating the biomarker effect of compulsory schooling in England. Journal of Human Capital, 4(2), 173-202.

Prospective Studies Collaboration. (2007). Blood cholesterol and vascular mortality by age, sex, and blood pressure: a meta-analysis of individual data from 61 prospective studies with 55 000 vascular deaths. The Lancet, 370(9602), 1829-1839.

Prospective Studies Collaboration. (2009). Body-mass index and cause-specific mortality in 900 000 adults: collaborative analyses of 57 prospective studies. The Lancet, 373(9669), 1083-1096.

Rahkovsky, I., Gregory, C.A. (2013). Food prices and blood cholesterol. Economics & Human Biology, 11(1), 95-107.

Seccareccia, F., Pannozzo, F., Dima, F., Minoprio, A., Menditto, A., Lo Noce, C., Giampaoli, S. (2001). Heart rate as a predictor of mortality: the MATISS project. American. Journal of Public Health, 91, 1258-1263.

Sesso, H. D., Stampfer, M. J., Rosner, B., Hennekens, C. H., Gaziano, J. M., Manson, J. E., Glynn, R. J. (2000). Systolic and diastolic blood pressure, pulse pressure, and mean arterial pressure as predictors of cardiovascular disease risk in men. Hypertension, 36(5), 801-807.

Singh-Manoux, A., Ferrie, J. E., Chandola, T., Marmot, M. (2004). Socioeconomic trajectories across the life course and health outcomes in midlife: evidence for the accumulation hypothesis? International Journal of Epidemiology, 33(5), 1072-1079.

Terza, J. V., Basu, A., Rathouz, P. J. (2008). Two-stage residual inclusion estimation: addressing endogeneity in health econometric modeling. Journal of Health Economics, 27(3), 531-543.

Turner, R. J., Thomas, C. S., Brown, T. H. (2016). Childhood adversity and adult health: Evaluating intervening mechanisms. Social Science & Medicine, 156, 114-124.

Van Doorslaer, E., Jones, A. M. (2003). Inequalities in self-reported health: validation of a new approach to measurement. Journal of Health Economics, 22(1), 61-87.

Verbeek, M., Nijman, T. (1992). Testing for selectivity bias in panel data models. International Economic Review, 681-703.

WHO (2011). Use of glycated haemoglobin (HbA1c) in the diagnosis of diabetes mellitus. World Health Organization, Geneva.

Ziebarth, N. (2010). Measurement of health, health inequality, and reporting heterogeneity. Social Science & Medicine, 71(1), 116-124.

Appendix

A. Descriptive statistics

Table A1. Descriptive statistics for all the independent variables used in the health outcome regression models.

	Maximum possible sample †		Nurse visits sample††		Blood sample¶	
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
Ethnicity						
White	0.968	0.177	0.969	0.174	0.975	0.156
Non-white (reference category)	0.032	0.177	0.031	0.174	0.025	0.156
Age-sex dummies						
Male (Age 19-24) (reference category)	0.030	0.170	0.019	0.137	0.016	0.125
Male (Age 25-29)	0.027	0.163	0.023	0.151	0.019	0.136
Male (Age 30-34)	0.035	0.184	0.032	0.176	0.027	0.163
Male (Age 35-39)	0.040	0.197	0.037	0.190	0.037	0.188
Male (Age 40-44)	0.046	0.210	0.045	0.208	0.044	0.205
Male (Age 45-49)	0.047	0.211	0.046	0.211	0.048	0.213
Male (Age 50-54)	0.043	0.203	0.044	0.205	0.049	0.216
Male (Age 55-59)	0.040	0.197	0.040	0.195	0.043	0.204
Male (Age 60-64)	0.038	0.192	0.039	0.193	0.039	0.194
Male (Age 65-69)	0.034	0.180	0.036	0.187	0.042	0.200
Male (Age 70-74)	0.027	0.163	0.029	0.167	0.031	0.174
Male (Age 75-79)	0.020	0.141	0.026	0.160	0.027	0.162
Male (Age 80-84)	0.016	0.124	0.015	0.123	0.016	0.124
Male (Age 85+)	0.009	0.097	0.011	0.102	0.008	0.091
Female (Age 19-24)	0.038	0.192	0.029	0.169	0.019	0.138
Female (Age 25-29)	0.032	0.177	0.030	0.169	0.022	0.147
Female (Age 30-34)	0.043	0.202	0.039	0.194	0.036	0.187
Female (Age 35-39)	0.047	0.211	0.048	0.214	0.051	0.219
Female (Age 40-44)	0.051	0.220	0.053	0.224	0.056	0.230
Female (Age 45-49)	0.058	0.233	0.060	0.237	0.060	0.238
Female (Age 50-54)	0.052	0.223	0.051	0.219	0.055	0.227
Female (Age 55-59)	0.042	0.200	0.044	0.206	0.045	0.207
Female (Age 60-64)	0.049	0.215	0.056	0.230	0.060	0.237
Female (Age 65-69)	0.040	0.197	0.045	0.208	0.050	0.217
Female (Age 70-74)	0.030	0.170	0.033	0.179	0.034	0.182
Female (Age 75-79)	0.028	0.166	0.032	0.175	0.033	0.179
Female (Age 80-84)	0.021	0.142	0.022	0.148	0.022	0.146
Female (Age 85+)	0.015	0.122	0.015	0.121	0.012	0.109
Educational attainment						
Degree	0.298	0.458	0.309	0.462	0.321	0.467
A-level or equivalent	0.236	0.425	0.222	0.415	0.214	0.410
O-level or basic qualification	0.321	0.467	0.325	0.469	0.322	0.467
No qualification (reference category)	0.145	0.352	0.145	0.352	0.143	0.350
Marital status						
Single	0.154	0.361	0.128	0.334	0.107	0.309
Married (reference category)	0.695	0.460	0.702	0.457	0.713	0.453
Separated/divorced	0.077	0.267	0.087	0.282	0.092	0.289
Widowed	0.074	0.262	0.083	0.276	0.088	0.284
Household size	2.735	1.372	2.626	1.276	2.595	1.235
Number of kids in household	0.531	0.931	0.499	0.897	0.491	0.873
Region						
North East	0.029	0.167	0.031	0.174	0.030	0.171
North West	0.084	0.278	0.092	0.289	0.096	0.295
Yorkshire & Humber	0.064	0.245	0.065	0.247	0.067	0.251
East Midlands	0.063	0.244	0.067	0.249	0.061	0.240
West Midlands	0.056	0.229	0.056	0.229	0.052	0.221
East of England	0.073	0.261	0.073	0.260	0.068	0.252
London	0.048	0.215	0.046	0.210	0.049	0.216
South East	0.102	0.302	0.105	0.307	0.102	0.303
South West	0.069	0.253	0.072	0.258	0.071	0.257
Wales	0.210	0.407	0.207	0.405	0.200	0.400
Scotland (reference category)	0.202	0.401	0.187	0.390	0.202	0.401
Sample size	7,979		4,474		3,003	

† Sample size corresponds to the maximum possible sample size for the health regression models. Sample size varies in Tables 1-6 depending on the health outcome considered.
†† Sample size corresponds to the nurse visits sub-sample.
¶ Sample size corresponds to the blood data sub-sample.

Table A2. Descriptive statistics for selected independent variables used in the longitudinal model of household income for the BHPS data

	Mean	Standard deviation
Household head (HoH) characteristics		
Male	0.613	0.487
White	0.968	0.175
HoH educational attainment		
Degree (reference category)	0.211	0.409
A-level or equivalent	0.203	0.402
O-level or basic qualification	0.331	0.470
No qualification	0.254	0.435
HoH employment status		
Employed (reference category)	0.518	0.500
Self-employed	0.091	0.287
Unemployed	0.032	0.176
Retired	0.228	0.420
Sick/disabled	0.047	0.211
Other activity	0.084	0.278
Household composition		
Lone parent	0.046	0.210
Couple without children (reference category)	0.423	0.494
Couple with children	0.296	0.456
Single: non-elderly	0.097	0.297
Single: elderly	0.090	0.286
Other (group HHs)	0.032	0.175
Multiple family households	0.016	0.124
Number of workers in the household	1.408	1.104
Individual's health indicators		
Self-assessed health		
Excellent (reference category)	0.232	0.422
Very good	0.465	0.499
Good	0.209	0.407
Fair	0.073	0.260
Poor	0.022	0.146
Health problems/disabilities	0.600	0.490
Health-related limitations on daily activities	0.166	0.372
Region of residence		
North East	0.038	0.190
North West	0.098	0.297
Yorkshire & Humber	0.076	0.265
East Midlands	0.069	0.254
West Midlands	0.071	0.257
East of England	0.072	0.259
London	0.072	0.258
South East	0.113	0.317
South West	0.075	0.264
Wales	0.146	0.354
Scotland (reference category)	0.169	0.375
Sample size	195,176	

Note: Although age dummies for individuals and the household head as well as year dummies are included in the income models, these variable are not presented to save space.

B. Panel data regressions for household income

Table A3 presents the results from the fixed effects model for equivalised household income. A strong negative association is observed between lower education of the HoH and higher equivalised household income. In a comparable context, Cappellari and Jenkins (2004) found that having a HoH with low education is associated with a higher risk of poverty. As expected, the employment status of the HoH is a significant correlate of equivalised income. Lone parent households, followed by couples with children and multiple family households are the type of households that are associated with the lowest equivalised family income compared to the references category of couples without children. Analogously, Cappellari and Jenkins (2004) found that living in lone parent or multiple family households as well as the presence of children increase the probability of transition to poverty. Moreover, individuals who are living in households with a higher number of working members have higher equivalised household income on average. A systematic association is also found between individual's SAH and household income. However, this association is relatively small in magnitude. For example, relative to those with excellent health, individuals reported fair health have a lower household income by about 2% (=(exp(-0.018)-1)*100). Finally, age dummies reveal an inverted U-shaped association if one compares the equivalised household income of our younger age group (15-19 years old; reference group) with any other age group (detailed results are not presented in Table A3 but are available upon request).

Table A3. Fixed effects model of household income for BHPS waves 1-18.

	Coefficient	Standard error
Household head (HoH) characteristics		
Male	0.001	0.005
White	0.090***	0.033
HoH educational attainment		
A-level or equivalent	-0.120***	0.010
O-level or basic qualification	-0.098***	0.010
No qualification	-0.118***	0.012
HoH employment status		
Self-employed	-0.197***	0.010
Unemployed	-0.425***	0.010
Retired	-0.290***	0.008
Sick/disabled	-0.236***	0.009
Other activity	-0.281***	0.008
Household composition		
Lone parent	-0.305***	0.012
Couple with children	-0.167***	0.005
Single: non-elderly	-0.111***	0.010
Single: elderly	-0.061***	0.013
Other (group HHs)	-0.068***	0.015
Multiple family households	-0.163***	0.011
Number of workers in the household	0.171***	0.003
Individual's health indicators		
Self-assessed health		
Very good	-0.006*	0.003
Good	-0.011***	0.004
Fair	-0.018***	0.006
Poor	-0.011	0.009
Health problems/disabilities	0.001	0.004
Health-related limitations on daily activities	0.001	0.003
Controls for:		
Individual age dummies	Y	
HoH age dummies	Y	
Year dummies	Y	
Regional dummies	Y	
Sample size	195,176	

Notes: Standard errors are clustered at the individual level.
***P<0.01; **P<0.05; *P<0.10

C. Income gradients in self-reported health measures for the nurse visits sub-sample

Table A4. Income gradients in self-assessed health and functional disabilities for the nurse visits sample.

Panel A: Self-assessed health						
	Ordered Probit					
	Coeff. (s.e.)†	APE (s.e.)†				
		Excellent	Very good	Good	Fair	Poor
Specification 1						
Ln(current income)	-0.264***	0.056***	0.041***	-0.022***	-0.043***	-0.032***
	(0.032)	(0.007)	(0.005)	(0.003)	(0.005)	(0.004)
Specification 2						
Long-run mean ln(income)	-0.409***	0.086***	0.062***	-0.034***	-0.066***	-0.048***
	(0.039)	(0.008)	(0.006)	(0.003)	(0.006)	(0.005)
Specification 3						
Long-run mean ln(income)	-0.675***	0.141***	0.102***	-0.055***	-0.109***	-0.080***
	(0.086)	(0.018)	(0.013)	(0.007)	(0.014)	(0.011)
Individual-specific effects	0.358***	-0.075***	-0.054***	0.029***	0.058***	0.042***
	(0.100)	(0.020)	(0.015)	(0.008)	(0.015)	(0.011)
Sample size	4,474					
Panel B: Functional disabilities						
	Probit					
	Coeff. (s.e.)†	APE (s.e.)†				
Specification 1						
Ln(current income)	-0.174***	-0.050***				
	(0.041)	(0.012)				
Specification 2						
Long-run mean ln(income)	-0.379***	-0.108***				
	(0.053)	(0.015)				
Specification 3						
Long-run mean ln(income)	-1.017***	-0.286***				
	(0.114)	(0.031)				
Individual-specific effects	0.857***	0.241***				
	(0.132)	(0.036)				
Sample size	4,474					

Abbreviations: APE, average partial effects; Coeff., coefficients; s.e., standard errors.
† Robust standard errors in parenthesis.
***P<0.01; **P<0.05; *P<0.10

Table A5. Income gradients in (inverted) PCS-12 for the nurse visits sample.

	OLS	**Unconditional quantile regressions**					
		Q10	Q25	Q50	Q75	Q90	Q95
	Coeff. (s.e.)†	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡	Coeff. (s.e.)‡
Specification 1							
Ln(current income)	-1.471***	-0.580***	-0.933***	-1.594***	-2.334***	-1.825**	-1.639**
	(0.299)	(0.203)	(0.188)	(0.325)	(0.701)	(0.906)	(0.810)
Specification 2							
Long-run mean ln(income)	-3.284***	-0.878***	-0.908***	-2.818***	-6.144***	-6.549***	-6.423***
	(0.397)	(0.263)	(0.243)	(0.421)	(1.007)	(1.308)	(1.154)
Specification 3							
Long-run mean ln(income)	-7.391***	-0.313	-1.162**	-5.387***	-14.58***	-19.60***	-15.53***
	(0.876)	(0.525)	(0.497)	(0.875)	(2.187)	(3.162)	(2.744)
Individual-specific effects	5.523***	-0.760	0.343	3.455***	11.34***	17.56***	12.25***
	(0.987)	(0.604)	(0.596)	(1.002)	(2.392)	(3.517)	(2.960)
Sample size	4,218						

Abbreviations: Coeff., coefficients; OLS, ordinary least squares; s.e., standard errors.
† Robust standard errors in parenthesis.
‡ UQR standard errors are bootstrap estimates with 500 replications.
***P<0.01; **P<0.05; *P<0.10