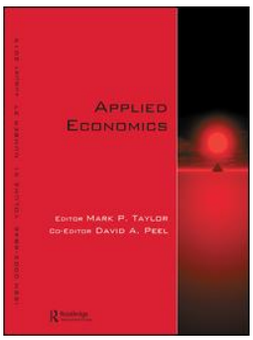

Applied Economics

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/raec20

In search of China's income-health gradient: a biomarker-based analysis

Peng Nie, Qing Li, Alan A. Cohen & Alfonso Sousa-Poza

To cite this article: Peng Nie, Qing Li, Alan A. Cohen & Alfonso Sousa-Poza (2021): In search of China's income-health gradient: a biomarker-based analysis, Applied Economics, DOI: 10.1080/00036846.2021.1927962

To link to this article: https://doi.org/10.1080/00036846.2021.1927962

APPLIED ECONOMICS
https://doi.org/10.1080/00036846.2021.1927962

Routledge
Taylor & Francis Group

In search of China's income-health gradient: a biomarker-based analysis

Peng Nie [a,b,c], Qing Li [d], Alan A. Cohen [e] and Alfonso Sousa-Poza [b,c]

[a]School of Economics and Finance, Xi'an Jiaotong University, Xi'an, China; [b]Institute for Health Care & Public Management, University of Hohenheim, Stuttgart, Germany; [c]IZA, Bonn, Germany; [d]School of Economics and Management, Xinjiang University, Urumqi, China; [e]Department of Family Medicine, University of Sherbrooke, Sherbrooke, Canada

ABSTRACT

Using data from the 1991–2009 China Health and Nutrition Survey, this study investigates China's income-health gradient by analysing the effect of both current and long-term household income on 22 blood-based biomarkers, 4 used as individual variables and all 22 assessed as a composite. We employ estimation approaches that allow for analysis 'beyond the mean' and accounting for individual-specific unobserved heterogeneity. After applying a two-step residual inclusion estimator, we find limited evidence of an income-health gradient irrespective of whether the income measure is current or long term. Because risky behaviour may attenuate income's positive effects on health, we also analyse the associations between income and such health-influencing factors as alcohol consumption, smoking, diet, physical activity, and dietary knowledge. Although we find that higher incomes go hand-in-hand with some of these factors (in particular, a higher number of cigarettes smoked per day), they also promote poorer diets. However, the fact that these effects are small, dependent on income measure, and susceptible to reporting biases makes it unlikely that they are attenuating income's potentially positive effects on health. Our findings highlight the importance of considering more accurate measures of health in assessing income-health gradients in future research.

KEYWORDS

China; income-health gradient; biomarkers; measuring health

I. Introduction

The existence of an income-health gradient is well established in the health economics literature, with the common finding that a higher income is linked to better health in Europe, the US, and elsewhere in the developed world (Adda, Banks, and Von Gaudecker 2009; Adeline & Delattre, 2017; Apouey and Clark 2014; Benzeval and Judge 2001; Benzeval, Taylor, and Judge 2000; Carrieri and Jones 2017; Davillas, Jones, and Benzeval 2017; Deaton and Paxson 1998; Ettner 1996; Frijters, Haisken-denew, and Shields 2005; Jones and Wildman 2008; Lindahl 2005). The underlying logic is that more income provides more avenues to better health, including better nutrition, improved access to health care, more opportunities for physical activity, more public safety, and lower environmental risk (Evans, Wolfe, and Adler 2012). However, the relation between income and health is not as clear-cut as one may assume. Quoting Cutler, Lleras-Muney, and Vogl (2008): 'Perhaps as individuals age, their health is best thought of as a stock that is relatively impervious to small changes in circumstance, . . . However, the preponderance of the evidence from developed countries today suggests that income does not have a large causal effect on adult health, whereas adult health has a large effect on adult income.' (p. 21). An ambiguous socioeconomic status (SES)-health gradient can also be illustrated in a special case of the Grossman model (Grossman 1972a, 1972b) when health is treated as a pure investment good whose demand depends solely on the returns to health capital which, in turn, remain unaffected by income.

China presents a particularly interesting case for assessing the income-health gradient because few other countries have experienced such a rapid and dramatic economic, social, and demographic transition. For example, following implementation of the Reform and Opening-Up Policy, Chinese per capita GDP increased from 385 yuan in 1978 to 59,660 yuan in 2017, while the average life expectancy increased from 67.77 in 1981 to 76.34 in 2015

CONTACT Alfonso Sousa-Poza alfonso.sousa-poza@uni-hohenheim.de Institute for Health Care & Public Management, University of Hohenheim, 70599, Stuttgart, Germany

(National Bureau of Statistics 2018). Yet rising standards of living coupled with a rapidly ageing population have also increased the prevalence of noncommunicable diseases, with substantial increases over recent decades in the incidence of hypertension, smoking, dyslipidemia, type 2 diabetes (T2DM), and obesity (Papagianni and Tziomalos 2018).

Nonetheless, identifying the income-health gradient presents several challenges, not least that the most common general health measure, self-reported health (SRH), is subject to inherent self-reporting bias with comparability problems on both the individual and national level (Carrieri and Jones 2017; Johnston, Propper, and Shields 2009). In particular, as Bago d'Uva, O'Donnell, and Van Doorslaer (2008) emphasize, SRH may differ in both conceptions of what leads to better health and expectations for one's own health status. If such beliefs and expectations vary with SES, then SRH differences provide a biased measure of SES inequality in health, one whose tendency to vary systematically with income and other SES components raises questions about reliability (Bago d'Uva et al. 2011; Bago d'Uva, O'Donnell, and Van Doorslaer 2008; Johnston, Propper, and Shields 2009; Rossouw 2018). This problem of reporting bias also extends to other common SRH measures such as functional limitations and chronic conditions (see, for instance, Johnston, Propper, and Shields 2009). More recent studies thus explore the income-health gradient using more objective health measures; most especially, blood-based biomarkers (Banks et al. 2006; Carrieri and Jones 2017; Davillas, Jones, and Benzeval 2017; Muennig, Sohler, and Mahato 2007).

A second major challenge is the appropriate definition of income, which according to the permanent income hypothesis, requires clear differentiation of permanent versus temporary or current income. Not only does the former have a far greater impact on health than the latter (Benzeval and Judge 2001), but in 13 longitudinal studies, long-term income was more strongly correlated with SRH than current earnings (Gunasekara, Carter, and Blakely 2011). Benzeval and Judge (2001) thus suggest that from a life-course perspective, long-term income may be especially relevant to health outcomes because it can capture cumulative disadvantages.

A third challenge is the common practice in income-health gradient studies of using regression models of the conditional mean of health outcomes (see, for instance, Banks et al. 2006; Benzeval, Taylor, and Judge 2000; Ettner 1996; Johnston, Propper, and Shields 2009; Muennig, Sohler, and Mahato 2007), which omit important information from other parts of the health outcome distribution. For example, both Carrieri and Jones (2017) and Davillas, Jones, and Benzeval (2017), using UK data, find a strong income gradient at the upper distributions of biomarkers. Hence, given that it is often the distribution tails that are of most concern to clinicians (Carrieri and Jones 2017), it is important to explore the income gradient across the entire distribution of health outcomes.

An additional methodological concern is that research seldom explores the mechanisms underlying the income-health gradient, including such health-enhancing behaviours as good diet and physical exercise versus health-compromising activities like smoking or drinking alcohol ((Brasher et al. 2017). For instance, whereas income shocks are detrimental to individual lifestyles that include smoking and social drinking (Apouey and Clark 2014; Van Kippersluis and Galama 2014), income can positively affect individual health status via better health knowledge and health service access.

Lastly, most of the literature on the income-health gradient, although it reports a clear positive association between income and good health, is based on developed nations, prompting considerable debate on the gradient's consistency and direction in developing countries (Mceniry 2013; Monteiro et al. 2004; Rosero-Bixby and Dow 2009; Smith and Goldman 2007; Zimmer et al. 2004). In the context of China, the evidence is even sparser and nowhere near as conclusive. For example, in two studies using 2006 China Health and Nutrition Survey (CHNS) data, whereas Yang and Kanavos (2012) report a positive correlation between higher income and better SRH with fewer physical activity limitation, Qi (2006) finds no linkage at all between income and SRH. This latter result is reinforced by both Rarick et al. (2017) Shanghai-based study showing no income-SRH association and Bakkeli (2016) evidence that individual income has little impact on objective health measures like obesity and blood pressure. In fact,

Deaton (2006) even reports a negative relation between rates of economic growth and reduced infant or child mortality in China.

This paper thus uses a combination of cross-sectional and longitudinal CHNS household income, health-related behaviour, and blood-based biomarker data to examine China's income-health gradient with an eye to all the above concerns. In doing so, we contribute to the income-health gradient literature in several ways: First, we employ both individual measures and a composite measure of the blood-based biomarkers, which have the distinct advantage of being objective and free from reporting bias. Second, we employ unconditional quantile regressions (UQR) to identify possible gradients at different points along the health distribution. Third, we introduce both current income and longitudinal mean income measures to assess the importance of permanent versus temporary income. Fourth, we explore the mechanisms underlying income's impacts on health by including individual health-affecting behaviours and dietary knowledge.

II. Data and methods

Data

The data are taken from the 2009 China Health and Nutrition Survey (CHNS), whose multistage randomly clustered sample covers nine provinces (Liaoning, Heilongjiang, Jiangsu, Shandong, Henan, Hubei, Hunan, Guangxi and Guizhou) (Zhang et al. 2014). We choose the 2009 wave because it was the first to contain blood biomarker data. The blood samples are taken through venipuncture in the morning on an empty stomach and tested immediately for glucose and haemoglobin A1C (HbA1C) (Yan et al. 2012). Plasma and serum samples are then frozen and stored at −86° C for laboratory analysis (Yan et al. 2012). All samples are further analysed in a national central laboratory in Beijing under strict quality control. In our study, we restrict our final sample to adults aged 18 and older for whom the 2009 data set provides detailed demographic, socioeconomic, and biomarker information. We also leverage the panel structure of the 1991–2009 CHNS to estimate longitudinal household income.

Health measures

Following Carrieri and Jones (2017) and Davillas, Jones, and Benzeval (2017), we use four biomarkers as dependent variables – glycated haemoglobin (HbA1c), cholesterol ratio, C-reactive protein (CRP), and white blood cell count (WBC). HbA1C, measured as the 3-month average plasma glucose concentration (in mmol/l), is found in high levels in individuals with elevated blood sugar (e.g. diabetes), whereas cholesterol ratio, calculated as the ratio of total cholesterol to high-density lipoprotein cholesterol, is associated with a high risk of cardiovascular disease and mortality risks (Prospective Studies Collaboration 2007). We also introduce two biomarkers of (systemic) inflammation: CRP, an acute-phase protein in blood that is synthesized in the liver in response to inflammation (Brenner et al. 2014), and WBC, a measure of total white blood cells, generally indicative of infection and associated with lung cancer risk (Brenner et al. 2014).

To expand upon the analyses of Carrieri and Jones (2017) and Davillas, Jones, and Benzeval (2017), we additionally develop a composite measure for all 22 biomarkers covered by the CHNS (abumin, alanine aminotransferase, apolipoprotein A-1, creatinine, ferritin, glucose, high-density lipoprotein cholesterol, insulin, low-density lipoprotein cholesterol, serum magnesium, total cholesterol, triglyceride, total protein, transferrin, soluble transferrin receptor, uric acid, urea, haemoglobin, WBC, red blood cell, platelet count, and HbA1c). This measure, first proposed by Cohen et al. (2013) in their study of physiological dysregulation, captures objective health as the Mahalanobis distance (DM) to a point of ideal health in the biomarker space. A greater distance from this centroid (i.e. larger values of the health indicator) implies worse health.

We calculate DM as follows:

$$D_M(x) = \sqrt{(x-\mu)^T S^{-1}(x-\mu)} \qquad (1)$$

where x is a vector of biomarker values for a given individual, μ is the equivalent-length vector of the means for each variable representing the ideal health level, and S is the variance-covariance matrix for the variables. We standardize all

variables by subtracting their mean and dividing by their standard deviation. Extensive validation analyses (see, e.g. Cohen et al. 2015, 2014; Li et al. 2015) suggest that the mean value is nearly optimal as a reference point, although a subset of younger healthier individuals may sometimes provide a better reference. Liu (2020) alsovalidates the DM measure in CHNS and shows that this measure captures a general signal of health in China similar to that of Western countries.We report the descriptive statistics for the composite biomarker measure sample in Table A1.

Income measures

Our current income measure is household per capita income (i.e. total household income divided by household size), which in the CHNS questionnaire comprises nine sources: farming, gardening, livestock/poultry, fishing, business, subsidies, retirement income, nonretirement earnings, and other. In this analysis, we use logged household per capita income to allow for income-health nexus concavity and to capture income distribution skewness (Contoyannis, Jones, and Rice 2004; Davillas, Jones, and Benzeval 2017). Similar to Davillas, Jones, and Benzeval (2017), we also introduce a measure for permanent (i.e. long-term) income, defined by calculating the within-individual mean of the natural logarithm of household per capita income over the available time period (maximum of 7 CHNS rounds).

Behavioural measures

To assess how changes in income may affect health, we examine the association between income and four aspects: risky behaviours (smoking and drinking alcohol), diet (macronutrients), physical activity, and knowledge of dietary guidelines. We define smoking by number of smoked cigarettes per day, and alcohol consumption by frequency: 1 = no more than once a month, 2 = once or twice a month, 3 = once or twice a week, 4 = 3–4 times a week, and 5 = almost every day. Because the CHNS monitors individual dietary intake for three consecutive days by asking all respondents about all foods consumed inside and outside the home on a 24-hour recall basis, we define calorie intake as the 3-day average intake in kilocalories; and fat, carbohydrate, and protein intake as 3-day average values in grams. We generate a variable for time spent on physical activity by summing up individual expenditure on specific sports: martial arts, gymnastics/dancing/acrobatics, track and field/swimming, ball sports (e.g. soccer/basketball/tennis, badminton/volleyball), and other (e.g. tai chi). Lastly, we define knowledge of dietary guidelines as a binary variable equal to 1 if the respondent is familiar with the five-level Chinese Pagoda or similar dietary guidelines and 0 otherwise. In addition to recommending portions from different food groups (i.e. grains, fresh vegetables, poultry and meat, and edible oil), the Chinese Pagoda also recommends drinking plenty of water and engaging in physical activity.

Controls

Controls for longitudinal income estimation

In line with previous studies estimating household income (see, Davillas, Jones, and Benzeval 2017), we introduce five individual and household characteristics: age (10 age group dummies for 5-year intervals with the 18–25 age group as the reference group); education (0 = illiterate, 1 = primary school, 2 = middle school, 3 = high school, 4 = technical/vocational school, and 5 = university or higher, with the illiterate group omitted); marital status (1 = never married, 2 = married, 3 = divorced/widowed/separated, with never married as the reference group); employment status (1 = employed, 0 = otherwise), and household size. To capture the aggregate income shocks associated with time-variant reporting changes (Davillas, Jones, and Benzeval 2017), we also employ wave dummies, with 1991 as the reference wave.

Controls for cross-sectional health estimation

Because health is age and gender dependent, in addition to controls for education level, marital and employment status, and household size, we introduce 10 age dummies for each gender to capture a flexible link between age, gender, and health. In addition, when analysing HbA1c and cholesterol ratio, we introduce an antidiabetic medications dummy to account for possible mediation effects

(Rahkovsky and Gregory 2013). Lastly, we include provincial dummies to capture geographic/regional variations.

Estimation strategy

We first investigate the income-health gradient using the following OLS model:

$$BIO_i = \beta_0 + \beta_1 PCHI_i + \beta_2 X_i + \beta_3 P_i + \varepsilon_i \quad (2)$$

where BIO_i denotes individual i's biomarker variable or biomarker-based composite measure, $PCHI_i$ is the per capita total household income (either current or long term as defined below), X_i is a vector of individual and household characteristics, P_i is a vector of provincial dummies, and ε_i is the error term. To detect possible heterogeneous effects of household income across the full distribution of health outcomes, we then estimate an unconditional quantile regression (UQR) model, which in its simplest form is estimable as an OLS regression on a transformed dependent variable using the recentered influence function (RIF) (Firpo, Fortin, and Lemieux 2009). Unlike the conditional quantile regression (which identifies covariate impacts on the conditional quantiles of the dependent variable), the UQR explores its *unconditional* quantile partial effects (Firpo, Fortin, and Lemieux 2009).

We estimate our UQR model as

$$RIF(BIO_i; Q_\tau, F_{BIO}) = \delta_i PCHI_i + \varphi_i X_i + \sigma_i P_i + \omega_i \quad (3)$$

where Q_τ denotes the τth quantile of the outcome cumulative distribution F_{BIO}. $PCHI_i$ and X_i follow the same logic as in equation (2), δ_i, φ_i and σ_i are the parameters to be estimated, and ω_i is an error term. RIF in equation (3) is thus

$$RIF(BIO_i; Q_\tau, F_{BIO}) = Q_\tau + (\tau - I[BIO_i \leq Q_\tau])/f_Y(Q_\tau) \quad (4)$$

where the probability distribution function of variable BIO_i is f_{BIO}, and $I[BIO_i \leq Q_\tau]$ represents the indicator function for whether a biomarker indicator is small or equal to the τth quantile. Like Jolliffe (2011), we use bootstrapping with 500 replications to obtain unbiased results for the variance-covariance matrix of the parameter estimates.

Next, to rule out any endogeneity-producing correlation between the individual-specific selection effects from our first-stage fixed effects income estimator and health outcomes, we employ a variant of the two-step residual inclusion estimator that allows for such time-invariant unobserved heterogeneity (cf. Davillas, Jones, and Benzeval (2017). More specifically, using 1991–2009 CHNS panel data, we first disentangle the time-invariant unobserved individual heterogeneity by estimating the following fixed effects model for household income:

$$ln(PCHI_{it}) = \theta' X_{it} + \nu_i + \varepsilon_{it} \quad (5)$$

where $PCHI_{it}$ represents individual i's per capita household total income at time t, X_{it} is a vector of the time-variant explanatory variable, ν_i denotes the time-invariant individual specific effects, and ε_{it} is a randomly distributed idiosyncratic error term. ν_i, which captures the time-invariant unobserved individual heterogeneity, is obtained as follows:

$$\hat{\nu}_i = ln(\widehat{PCHI_{it}}) - \widehat{\theta}' \bar{X}_{it} \quad (6)$$

We then introduce $\hat{\nu}_i$ into the health outcomes estimation as an additional regressor, allowing us to estimate the individual-specific selection effects (from the time-invariant unobserved heterogeneity associated with both long-term income and health outcomes) with a common factor structure:

$$u_i = \delta \hat{\nu}_i + \pi_i \quad (7)$$

where π_i is the new idiosyncratic error term in the health outcome estimation.

Finally, to analyse the association between income and health behaviours, we estimate several regressions of the following general form:

$$B_i = \beta_0 + \beta_1 PCHI_i + \beta_2 X_i + \beta_3 P_i + \varepsilon_i \quad (8)$$

where B_i denotes a specific health behaviour such as smoking, alcohol consumption, macronutrient intake, physical activity, or knowledge of dietary guidelines. Because the dependent variables have different forms, we use different models for each: tobit for smoking, ordered probit for alcohol consumption, quantile regression for micronutrients, OLS for physical activity, and probit for dietary guidelines. Here,

$PCHI_i$ is per capita total household income (either current or long term), X_i is a vector of individual and household characteristics, P_i is a vector of provincial dummies, and ε_i is the error term.

III. Results

Income-health gradient using cross-sectional and longitudinal income measures

As Tables 1 and 2 clearly show, the estimated translog per capita household income coefficients of the OLS and RIF regressions are insignificant, regardless of whether current or longitudinal mean income is used. Hence, in sharp contrast to the findings for England (Davillas, Jones, and Benzeval 2017) of a strong negative income gradient for all four biomarkers selected (HbA1c, cholesterol ratio, CRP and fibrinogen), our results point to the absence of any income-health gradient. In the UQR estimates, although using long-run mean income yields the expected increase to larger coefficients, none of the biomarkers are significant at conventional levels.

The outcomes for our composite measure based on the 22 biomarkers are also clearly insignificant (Table 3), indicating no income gradient in the composite health measure either. The only exception is a positive coefficient at the 50th percentile when using permanent income; however, this coefficient is also very small (0.01 elasticity).[1]

Table 1. OLS and RIF estimates for HbA1c and cholesterol ratio among adults aged 18 + .

Panel A: HbA1c	OLS	25th	50th	75th	90th	95th
Ln(current income)	0.002	0.001	0.003	0.007	0.003	−0.004
	(0.010)	(0.007)	(0.007)	(0.009)	(0.023)	(0.052)
N	6730	6730	6730	6730	6730	6730
Adj. R^2	0.195	0.124	0.184	0.180	0.146	0.168
Ln(permanent income)	0.004	0.002	0.004	0.010	0.015	0.027
	(0.009)	(0.006)	(0.006)	(0.008)	(0.023)	(0.050)
N	6730	6730	6730	6730	6730	6730
Adj. R^2	0.195	0.124	0.184	0.180	0.146	0.168
Ln(permanent income)	0.004	0.001	0.005	0.011	0.020	0.056
	(0.012)	(0.008)	(0.008)	(0.011)	(0.028)	(0.058)
Individual-specific effects	−0.000	0.003	−0.003	−0.003	−0.015	−0.081
	(0.020)	(0.014)	(0.013)	(0.018)	(0.049)	(0.104)
N	6730	6730	6730	6730	6730	6730
Adj. R^2	0.195	0.124	0.184	0.180	0.146	0.168
Panel B: Cholesterol ratio	OLS	25th	50th	75th	90th	95th
Ln(current income)	0.013	0.015	0.007	0.027	−0.003	−0.043
	(0.014)	(0.013)	(0.017)	(0.023)	(0.034)	(0.053)
N	6760	6760	6760	6760	6760	6760
Adj. R^2	0.073	0.050	0.064	0.049	0.028	0.019
Ln(permanent income)	0.020	0.017	0.015	0.036	0.011	−0.017
	(0.013)	(0.013)	(0.015)	(0.020)	(0.030)	(0.052)
N	6760	6760	6760	6760	6760	6760
Adj. R^2	0.069	0.047	0.061	0.046	0.027	0.018
Ln(permanent income)	−0.018	−0.001	−0.018	−0.026	−0.033	−0.067
	(0.017)	(0.016)	(0.019)	(0.024)	(0.038)	(0.065)
Individual-specific effects	0.108**	0.053	0.093**	0.178**	0.126*	0.142
	(0.029)	(0.027)	(0.033)	(0.043)	(0.065)	(0.110)
N	6760	6760	6760	6760	6760	6760
Adj. R^2	0.075	0.050	0.065	0.052	0.029	0.019

Notes: The dependent variables are HbA1c and cholesterol ratio. The controls include individual characteristics (age-gender dummies, education level, marital status, employment status, and antidiabetes medication), translog household equivalized income, household size, and provincial dummies (with Liaoning as the reference). Robust standard errors for the OLS estimates are in parentheses; standard errors for the UQR estimates are bootstrapped with 500 replications. * $p < 0.05$, ** $p < 0.01$.

[1]We conducted a number of robustness checks for healthier reference groups, including (i) doctor-diagnosed high blood pressure, diabetes, myocardinal infarction, apolexy, bone fracture and asthma information; (ii) high-density lipoprotein cholesterol, low-density lipoprotein cholesterol, total cholesterol, triglyceride, general obesity, abdominal obesity and high blood pressure; (iii) high-density lipoprotein cholesterol, low-density lipoprotein cholesterol, total cholesterol and triglyceride. We use the healthier reference groups to recalculate the composite measure of health. We generally obtain results, available from the authors upon request, that are qualitatively similar to those in Table 3.

Table 2. OLS and RIF estimates for CRP and WBC among adults aged 18 + .

Panel A: CRP	OLS	25th[a]	50th	75th	90th	95th
Ln(current income)	0.002		0.007	0.050	−0.088	0.055
	(0.041)		(0.023)	(0.041)	(0.113)	(0.186)
N	5089		5089	5089	5089	5089
Adj. R^2	0.013		0.010	0.013	0.009	0.004
Ln(permanent income)	−0.002		0.020	0.065	−0.099	−0.029
	(0.037)		(0.020)	(0.038)	(0.105)	(0.160)
N	5089		5089	5089	5089	5089
Adj. R^2	0.013		0.010	0.014	0.009	0.004
Ln(permanent income)	0.024		0.008	0.075	0.001	0.119
	(0.048)		(0.025)	(0.047)	(0.136)	(0.192)
Individual-specific effects	−0.071		0.031	−0.027	−0.273	−0.406
	(0.086)		(0.045)	(0.085)	(0.229)	(0.364)
N	5089		5089	5089	5089	5089
Adj. R^2	0.013		0.010	0.014	0.009	0.004
Panel B: WBC	OLS	25th	50th	75th	90th	95th
Ln(current income)	−0.006	0.014	−0.015	0.001	−0.012	−0.021
	(0.022)	(0.023)	(0.026)	(0.033)	(0.052)	(0.080)
N	6804	6804	6804	6804	6804	6804
Adj. R^2	0.049	0.042	0.044	0.032	0.016	0.009
Ln(permanent income)	−0.011	0.024	−0.007	−0.009	−0.042	−0.044
	(0.021)	(0.020)	(0.023)	(0.029)	(0.050)	(0.072)
N	6804	6804	6804	6804	6804	6804
Adj. R^2	0.049	0.042	0.044	0.032	0.016	0.009
Ln(permanent income)	−0.017	0.023	−0.012	−0.045	−0.089	−0.071
	(0.027)	(0.025)	(0.028)	(0.035)	(0.062)	(0.096)
Individual-specific effects	0.016	0.003	0.014	0.101	0.132	0.077
	(0.046)	(0.046)	(0.050)	(0.063)	(0.102)	(0.157)
N	6804	6804	6804	6804	6804	6804
Adj. R^2	0.049	0.042	0.043	0.032	0.016	0.009

Notes: The dependent variables are C-reactive protein (CRP) and white blood-cell count (WBC). The controls include individual characteristics (age-gender dummies, education level, marital status, and employment status), translog household equivalisized income, household size, and provincial dummies (with Liaoning as the reference). Robust standard errors for the OLS estimates are in parentheses; standard errors for the UQR estimates are bootstrapped with 500 replications. * $p < 0.1$, ** $p < 0.05$.

[a]The 25th percentile cannot be estimated for CRP because just under 50% of the observations have the minimum value of 1.

Table 3. OLS and RIF estimates for composite measure of health among adults aged 18 + .

	OLS	25th	50th	75th	90th	95th
Ln(current income)	0.001	−0.005	0.002	0.004	0.009	0.011
	(0.003)	(0.004)	(0.004)	(0.004)	(0.007)	(0.012)
N	6651	6651	6651	6651	6651	6651
Adj. R^2	0.038	0.036	0.037	0.024	0.008	0.004
Ln(permanent income)	0.003	−0.005	0.005	0.006	0.011	0.011
	(0.003)	(0.003)	(0.003)	(0.004)	(0.007)	(0.011)
N	6651	6651	6651	6651	6651	6651
Adj. R^2	0.038	0.036	0.037	0.024	0.008	0.004
Ln(permanent income)	0.007	0.00004	0.010*	0.008	0.016	0.006
	(0.004)	(0.004)	(0.004)	(0.005)	(0.009)	(0.013)
Individual-specific effects	−0.011	−0.015*	−0.014	−0.007	−0.015	0.014
	(0.007)	(0.007)	(0.007)	(0.009)	(0.016)	(0.022)
N	6651	6651	6651	6651	6651	6651
Adj. R^2	0.038	0.037	0.037	0.024	0.007	0.004

Notes: The dependent variable is the translog composite measure of health, constructed with 22 biomarkers. The controls are individual characteristics (age-gender dummies, education level, marital status, and employment status), translog household equivalized income, household size, and provincial dummies (with Liaoning as the reference). Robust standard errors for the OLS estimates are in parentheses; standard errors for the UQR estimates are bootstrapped with 500 replications. * $p < 0.05$, ** $p < 0.01$.

Income-health gradient accounting for unobserved individual heterogeneity and underlying mechanisms

For the two-step estimator used to rule out individual selection effects from time-invariant unobserved heterogeneity, except for the cholesterol ratio, the individual-specific selection effects are insignificant. These results, like the UK findings (Davillas, Jones, and Benzeval 2017), suggest no systematic selection effects for our objectively measured health outcomes. Nonetheless, this clearly insignificant relation between income and health could be the result of potentially attenuating effects of income-dependent poor health behaviours. For example, in Table 4, higher levels of both longitudinal average income

Table 4. Tobit/ordered probit estimates for smoking and alcohol consumption among adults aged 18 + .

	Number of cigarettes smoked per day	Frequency of alcohol consumption				
	Tobit	Ordered probit (marginal effects)				
		1	2	3	4	5
Ln(current income)	0.651*	−0.006	−0.005	−0.001	0.001	0.012
	(0.270)	(0.004)	(0.004)	(0.001)	(0.001)	(0.008)
N	1931	2265	2265	2265	2265	2265
Pseudo R^2	0.034	0.050	0.050	0.050	0.050	0.050
Ln(permanent income)	0.623*	−0.003	−0.003	−0.001	0.001	0.006
	(0.250)	(0.004)	(0.003)	(0.001)	(0.001)	(0.007)
N	1931	2265	2265	2265	2265	2265
Pseudo R^2	0.034	0.050	0.050	0.050	0.050	0.050
Ln(permanent income)	0.618	−0.002	−0.002	−0.0004	0.0003	0.003
	(0.326)	(0.005)	(0.004)	(0.001)	(0.001)	(0.010)
Individual-specific effects	0.015	−0.005	−0.004	−0.001	0.001	0.009
	(0.585)	(0.009)	(0.008)	(0.002)	(0.002)	(0.017)
N	1931	2265	2265	2265	2265	2265
Pseudo R^2	0.034	0.050	0.050	0.050	0.050	0.050

Notes: The dependent variables are number of smoked cigarettes per day and frequency of alcohol consumption (1 = no more than once a month, 2 = once or twice a month, 3 = once or twice a week, 4 = 3–4 times a week, and 5 = almost every day). The controls include individual characteristics (age-gender dummies, education level, marital status, and employment status), translog household equivalized income, household size, and provincial dummies (with Liaoning as the reference). * $p < 0.05$, ** $p < 0.01$.

and current income are associated with an increase in the number of cigarettes smoked per day, echoing both Li and Zhu (2006) for China and Apouey and Clark (2014) for the UK. Neither of these observations, however, hold for alcohol consumption. On the other hand, when we include our two-stage estimator for heterogeneity control, longitudinal mean income is uncorrelated with the level of either smoking or alcohol consumption.

As regards income's effect on individual diet (Table 5), the OLS estimates indicate that current income is positively associated with higher intake of calories, fat, and protein, whereas the longitudinal mean income is positively and significantly correlated with fat and protein (Panels C and D) but negatively correlated with carbohydrates (Panel B). In the RIFR estimates, once we account for the individual-specific selection effects – which initially produce heterogeneities in the current income gradient for calories, fat, and protein and the average income gradient for carbohydrates – all the macronutrients are impervious to longitudinal average income with the exception of carbohydrates at the 25th percentile.[2] We also note that the individual-specific effects are often significant, once again highlighting that measures based on self-reports are susceptible to reporting bias. This finding is important given the great reliance in much of the micronutrient research on such self-reported consumption measures.

The results reveal no associations, however, between time spent on physical activity and either current or longitudinal mean income (Table 6), even when we adjust the latter for individual-specific selection effects, which once again are highly significant. Nor do the results show any correlation between these two income variables and knowledge of dietary guidelines either before or after we account for unobserved individual heterogeneity.

Robustness checks

Addressing endogeneity using the Lewbel (2012) approach

Although research on the income-health gradient typically focuses on association rather than causality, it is highly probable that poor health impairs individual productivity, and thereby income (Evans, Wolfe, and Adler 2012). It is also likely that certain common factors – for example, individual motivation or genetics – influence both income and health outcomes (Apouey and Clark 2014; Evans, Wolfe, and Adler 2012). Identifying possible causal routes is thus greatly hindered by the potential risk for reverse causality or omitted factors (Evans, Wolfe, and Adler 2012). Hence, although several lottery-based studies identify significant positive income effects on health outcome (Apouey and

[2]As a robustness test, we also run probit estimates using dummies for four macronutrients (calories (1 ≥ 2000 kcal), carbohydrates (1 ≥ 150 g), fat (1 ≥ 78 g), and protein (1 ≥ 56 g)), which yields quantitatively similar results to those in Table 5 (see Table A2).

Table 5. OLS and RIFR estimates for macronutrients among adults aged 18 + .

Panel A: calories	OLS	25th	50th	75th	90th	95th
Ln(current income)	0.010**	0.007	0.010*	0.009*	0.007	0.013*
	(0.004)	(0.006)	(0.004)	(0.005)	(0.006)	(0.007)
N	6827	6827	6827	6827	6827	6827
Adj. R^2	0.168	0.101	0.122	0.105	0.075	0.045
Ln(permanent income)	0.002	−0.001	0.003	−0.001	−0.001	0.007
	(0.003)	(0.005)	(0.004)	(0.004)	(0.005)	(0.006)
N	6827	6827	6827	6827	6827	6827
Adj. R^2	0.167	0.101	0.121	0.104	0.075	0.044
Ln(permanent income)	0.001	−0.003	0.001	0.004	0.005	0.013
	(0.005)	(0.006)	(0.005)	(0.005)	(0.007)	(0.009)
Individual-specific effects	0.002	0.004	0.005	−0.016	−0.017	−0.017
	(0.008)	(0.011)	(0.009)	(0.010)	(0.012)	(0.015)
N	6827	6827	6827	6827	6827	6827
Adj. R^2	0.167	0.101	0.121	0.104	0.075	0.044
Panel B: carbohydrates	OLS	25th	50th	75th	90th	95th
Ln(current income)	−0.005	−0.015*	−0.005	−0.006	−0.001	0.001
	(0.004)	(0.007)	(0.005)	(0.005)	(0.006)	(0.008)
N	6827	6827	6827	6827	6827	6827
Adj. R^2	0.166	0.087	0.128	0.107	0.068	0.048
Ln(permanent income)	−0.014**	−0.024**	−0.013**	−0.015**	−0.013*	−0.008
	(0.004)	(0.006)	(0.005)	(0.005)	(0.005)	(0.007)
N	6827	6827	6827	6827	6827	6827
Adj. R^2	0.167	0.088	0.128	0.108	0.069	0.048
Ln(permanent income)	0.003	−0.015*	0.007	0.009	0.001	0.011
	(0.005)	(0.008)	(0.006)	(0.006)	(0.007)	(0.009)
Individual-specific effects	−0.050**	−0.025	−0.058**	−0.069**	−0.042**	−0.055**
	(0.009)	(0.013)	(0.011)	(0.011)	(0.012)	(0.015)
N	6827	6827	6827	6827	6827	6827
Adj. R^2	0.172	0.089	0.132	0.113	0.070	0.050
Panel C: fat	OLS	25th	50th	75th	90th	95th
Ln(current income)	0.027**	0.025**	0.030***	0.026***	0.013	0.013
	(0.007)	(0.010)	(0.008)	(0.007)	(0.008)	(0.009)
N	6827	6827	6827	6827	6827	6827
Adj. R^2	0.104	0.066	0.073	0.062	0.047	0.034
Ln(permanent income)	0.021**	0.024***	0.024***	0.017**	0.010	0.009
	(0.006)	(0.009)	(0.007)	(0.007)	(0.007)	(0.008)
N	6827	6827	6827	6827	6827	6827
Adj. R^2	0.103	0.066	0.072	0.061	0.047	0.034
Ln(permanent income)	−0.009	−0.009	−0.004	−0.006	−0.008	−0.003
	(0.008)	(0.012)	(0.009)	(0.009)	(0.009)	(0.012)
Individual-specific effects	0.086**	0.095***	0.078***	0.067***	0.051***	0.034
	(0.014)	(0.021)	(0.015)	(0.015)	(0.017)	(0.022)
N	6827	6827	6827	6827	6827	6827
Adj. R^2	0.109	0.069	0.076	0.063	0.048	0.034
Panel D: protein	OLS	25th	50th	75th	90th	95th
Ln(current income)	0.025**	0.022**	0.019**	0.016**	0.015**	0.016*
	(0.004)	(0.006)	(0.005)	(0.005)	(0.007)	(0.007)
N	6827	6827	6827	6827	6827	6827
Adj. R^2	0.181	0.094	0.123	0.121	0.080	0.047
Ln(permanent income)	0.020**	0.019**	0.014**	0.009*	0.010	0.011
	(0.004)	(0.005)	(0.005)	(0.005)	(0.006)	(0.007)
N	6827	6827	6827	6827	6827	6827
Adj. R^2	0.180	0.094	0.122	0.120	0.080	0.047
Ln(permanent income)	0.004	−0.002	−0.007	−0.001	0.009	0.004
	(0.007)	(0.006)	(0.006)	(0.008)	(0.009)	(0.007)
Individual-specific effects	0.043**	0.047**	0.047**	0.031*	0.006	0.043**
	(0.012)	(0.010)	(0.011)	(0.014)	(0.017)	(0.012)
N	6827	6827	6827	6827	6827	6827
Adj. R^2	0.096	0.124	0.122	0.080	0.047	0.096

Notes: The dependent variables are the macronutrients. The controls include individual characteristics (age-gender dummies, education level, marital status, and employment status), translog household equivalized income, household size, and provincial dummies (with Liaoning as the reference). For the OLS estimates, robust standard errors are in parentheses. * $p < 0.05$, ** $p < 0.01$.

Clark 2014; Gardner and Oswald 2007; Lindahl 2005), Gunasekara, Carter, and Blakely (2011) criticize the practice on the grounds that, even in the same individual, 'income windfalls from [such] natural experiments … may not be associated with the same health behaviours' as occur with the expectation of increased permanent income (p. 200). In a similar vein, they criticize Frijters, Haisken-denew, and Shields' (2005) assumption that income increases related

Table 6. OLS/probit estimates for physical activity and knowledge of dietary guidelines among adults aged 18+ (marginal effects).

	Time spent on physical activity (hours/week)			Knowledge of dietary guidelines		
	OLS			Probit (marginal effects)		
Ln(current income)	−0.072			0.019		
	(0.093)			(0.015)		
Ln(permanent income)		−0.131	−0.189		0.025	0.023
		(0.085)	(0.110)		(0.015)	(0.018)
Individual-specific effects			0.154			0.002
			(0.175)			(0.019)
N	562	562	562	883	883	883
Pseudo R^2	0.046	0.049	0.049	0.186	0.188	0.188

Notes: The dependent variables are time spent on physical activity (measured in hours/week) and knowledge of the Chinese Pagoda or similar dietary guidelines. The former encompasses martial arts, gymnastics/dancing/acrobatics, track and field, soccer/basketball/tennis, badminton/volleyball and other activities. The controls include individual characteristics (age-gender dummies, education level, marital status, and employment status), translog household equivalized income, household size, and provincial dummies (with Liaoning as the reference). Marginal errors are reported for the probit estimates, with robust standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$.

to East-West German reunification are exogenous for East Germans, claiming that such exogeneity is likely to diminish over time, leading to bias (Gunasekara, Carter, and Blakely 2011, 199). To add weight to their argument, they further point out that natural experiments, although advantageous for exogeneity, have relatively limited generalizability (Gunasekara, Carter, and Blakely 2011).

Although the two-step estimator used here does account for certain endogeneity, in the case of long-term income gradients in health, it is restricted to that from time-invariant unobserved heterogeneity (i.e. individual-specific selection effects). Hence, to detect any additional endogeneity issues in the income-health relation, we apply Lewbel's (2012) heteroscedasticity-based 2SLS, which requires the presence of heteroscedasticity as a precondition for identification (confirmed here by a Breusch and Pagan (1979) test) but enables IV estimation in the absence of any obvious instruments.[3] As before, the 2SLS results show no income gradient for either current or (longitudinal) average income (see Table A3).

Urban versus rural

As a final check for the presence of an income-health gradient in China, we consider geographic differences by performing a split analysis for urban versus rural residents. Yet again, we observe no income-health gradient for either rural or urban areas (see Appendix Tables A4-A6).[4]

Non-monotonic effects of income

Biologically speaking, our quantile regression approach allowed us to test a wide range of hypotheses (some more plausible, some less). For example, CRP is an inflammatory marker that shows low-grade increases with chronic metabolic conditions such as diabetes and heart disease, but that spikes to much higher levels during acute infections. It was therefore plausible that individuals in the lowest income quartile, who might be more exposed to acute infections (Ticinesi et al. 2017), would have a higher probability of being in the 95th quantile, whereas higher income individuals might have a higher probability of being at the 75th quantile, with little distinction for the lower quantiles. This was not what we found; indeed, globally and for the 50th through 90th quantiles, income Q2 (slightly higher income) had lower CRP levels than income Q1 (lowest income), with no differences between income Q1 and income Q3 or Q4 (higher or highest income) (Appendix Table A7). Broadly speaking, there was little if any support for the various such biologically plausible scenarios. In fact, the only pattern of relationships between biomarkers and income appears when considering income by quartiles (Appendix Table A7). Income Q2 has somewhat better health than income Q1 across HbA1C, cholesterol ratios, and CRP, though not for WBC or the composite measure, and not for all quantiles. This result is not replicated for income Q3 or Q4, implying that if there is a relationship between health and income in this dataset, it is non-monotonic, with intermediate incomes being healthiest. Puzzlingly, this finding was consistent across the three markers of cardio-metabolic health

[3] See Appendix B for a detailed description of this approach.

[4] Several other robustness checks for income variables – including adding a squared term, income quantiles and and household income divided by the square root of the total number of household members – also yield no significant results for income gradients in health. Even introducing a community-level urbanicity score, which reflects population size and density, community health infrastructure, sanitation, and socioeconomic characteristics, does not change the outcomes: we find no income gradient for biomarkers. All these additional results are available from the authors upon request.

(HbA1C, cholesterol ratios, and CRP), which we might expect to group together, and which we might expect to be related to access to modern diets and lifestyles in the opposite direction from our finding. However, we urge caution in over-interpreting this finding. While consistent across several markers, these markers are themselves correlated, and given the large number of tests conducted in this analysis, it could be a spurious finding.

Income-body mass index gradient

As a developing country, China has witnessed a rapid transition from historical undernutrition to a sharp increase in overweight and obesity (Xi et al. 2012; Nie, Ding, and Sousa-Poza 2019), and also more rapid than recorded in any other country is a dramatic shift in diet (mostly notably, the increased intake of edible oils, fried foods, animal-sourced foods and snacks) accompanied by a sharp decline in occupational and domestic PA (Nie, Ding, and Sousa-Poza 2019). Considering this, we also take a look at how income affects the body mass index (BMI) (Appendix Table 8). Our OLS estimates show that current income is positively correlated with higher BMI (Panel A), and this is also the case for the longitudinal mean income (Panel B). This finding here recalls our observation that a higher level of income is associated with poor diet (Table 5). When we account for the individual-specific selection effects, our RIFR estimates indicate that there is no association between the longitudinal mean income and BMI, though estimated coefficients are generally positive (Panel C). Furthermore, the individual-specific effects are significant except for 95th quantile of the BMI distribution.

IV. Conclusions

Although a large body of literature examines the income-health gradient, relatively few of these studies consider non-Western countries, particularly China. Those that do address the Chinese context paint an inconclusive picture, with some revealing a positive income-health gradient and others no relation. The research that does provide evidence of a positive income-health gradient not only tends to be based on SRH, known to be susceptible to reporting bias, but often addresses only current income when long-term (permanent) income is probably more relevant for capturing income-related changes in health behaviours (and thus health). This present study of China's income-health gradient overcomes these shortcomings by basing its analysis on both individual and composite measures of blood biomarker data from the China Health and Nutrition Survey and including both current and long-term income in the regressions. It is also innovative in assessing the income-health gradient across the entire health-measure distribution while using a two-step residual inclusion estimator and heteroscedasticity-based 2SLS Lewbel (2012) to control for endogeneity, and in employing not only individual health-affecting behaviours but also health knowledge to explore the underlying mechanisms through which income may impact health.

Our major finding is that, contrary to the results for England (Carrieri and Jones 2017; Davillas, Jones, and Benzeval 2017), there appears to be no biomarker-based income-health gradient in China, possibly because income increases are fostering such health-damaging behaviours as poor diet, smoking, and alcohol consumption. However, although current income is positively linked not only to fat and protein consumption (as is permanent income) but also to increased calorie intake, its association with increased cigarette smoking fails to hold when we adjust for individual-specific selection effects. We also find no linkage between either current or longitudinal mean income and time spent on physical activity or individual knowledge of dietary guidelines. Hence, although some of these behaviours may be attenuating income's positive effects on health, the small magnitude of their effects, the differences between permanent and current income, and the potential for some to even improve health outcomes (e.g. through dietary intake) leads us to doubt that they are responsible for the insignificance of the income-health gradient. Rather, it is worth noting that the effects of these health behaviours become insignificant once we control for endogeneity from time-invariant unobserved heterogeneity, implying that the behavioural measures themselves, being based on self-reported survey data, may be biased.

Another possible explanation given the tendency for income-health gradients to be less steep when health systems are more developed and coverage more comprehensive, is that the recent expansion of China's public health system in terms of both accessibility and affordability may be muting the association between SES and health (Lowry and Xie 2009). In fact, as a result of several health reforms since 2003, by 2011, 96% of China's households had health insurance and enjoyed significantly reduced out-of-pocket payments for total health expenditures, as well as substantially less inequality in insurance coverage and access to care (Papagianni and Tziomalos 2018).

One final explanation for the apparent absence of an income-health gradient in China could be the family support that is an integral part of Chinese society and on which about 85% of rural elderly depend (Gong et al. 2012). This support, which acts as a type of insurance, may not be captured in a household's income yet could positively affect health outcomes. Indeed, this failure to capture a link between income and health may even extend to our dataset of blood-based biomarkers, which, although numerous, may not fully account for all aspects of health. Even with such considerations, however, our results are extremely puzzling in their refutation of the positive income-health gradient so widely reported in the literature. They also differ greatly from the few studies that use biomarkers to assess the income-health gradient in Western countries. Unfortunately, the absence of long-term biomarker data makes pinpointing the exact reason for these differences extremely difficult, although it does present an interesting challenge for future research. Our findings, however, highlight the importance of considering more accurate measures of health (such as biomarker) when assessing the income-health gradient, especially in developing countries, in future research.

Acknowledgments

This research uses data from China Health and Nutrition Survey (CHNS). We thank the National Institute for Nutrition and Health, China Center for Disease Control and Prevention, Carolina Population Center (P2C HD050924, T32 HD007168), the University of North Carolina at Chapel Hill, the NIH (R01-HD30880, DK056350, R24 HD050924, and R01-HD38700) and the NIH Fogarty International Center (D43 TW009077, D43 TW007709) for financial support for the CHNS data collection and analysis files from 1989 to 2015 and future surveys, and the China-Japan Friendship Hospital, Ministry of Health for support for CHNS 2009, Chinese National Human Genome Center at Shanghai since 2009, and Beijing Municipal Center for Disease Prevention and Control since 2011. This work was supported by the National Natural Science Foundation of China (grant numbers 71804142; 72074178) and the Start-up Fund for Young Talent Support Plan (grant number 71211201010701). We would like to thank the participants in the Global Health Institute seminar in Xi'an, Sino-UK Health Forum in Ningbo and Second China Health Economics Forum in Chengdu. We also thank the editor, Mark Taylor, and one anonymous referee for valuable comments on an earlier version of this paper. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. All errors are our own.

Declarations of interest

The authors declare that they have no competing interests.

Disclosure of potential conflicts of interest

No potential conflict of interest was reported by the authors.

ORCID

Peng Nie http://orcid.org/0000-0002-5322-6324
Qing Li http://orcid.org/0000-0002-4191-9596
Alan A. Cohen http://orcid.org/0000-0003-4113-3988
Alfonso Sousa-Poza http://orcid.org/0000-0002-6524-9654

References

Adda, J., J. Banks, and H. M. Von Gaudecker. 2009. "The Impact of Income Shocks on Health: Evidence from Cohort Data." *Journal of the European Economic Association* 7 (6): 1361–1399. doi:10.1162/JEEA.2009.7.6.1361.

Adeline, A., and E. Delattre. 2017. "Some Microeconometric Evidence on the Relationship between Health and Income." *Health Economics Review* 7 (1): 27. doi:10.1186/s13561-017-0163-5.

Apouey, B., and A. E. Clark. 2014. "Winning Big but Feeling No Better? The Effect of Lottery Prizes on Physical and Mental Health." *Health Economics* 24 (5): 516–538. doi:10.1002/hec.3035.

Bago d'Uva, T., M. Lindeboom, O. O'Donnell, and E. Van Doorslaer. 2011. "Slipping Anchor?: Testing the Vignettes Approach to Identification and Correction of Reporting Heterogeneity." *Journal of Human Resources* 46 (4): 875–906.

Bago d'Uva, T., O. O'Donnell, and E. Van Doorslaer. 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. doi:10.1093/ije/dyn146.

Bakkeli, N. Z. 2016. "Income Inequality and Health in China: A Panel Data Analysis." *Social Science & Medicine* 157: 39–47. doi:10.1016/j.socscimed.2016.03.041.

Banks, J., M. Marmot, Z. Oldfield, and J. P. Smith. 2006. "Disease and Disadvantage in the United States and in England." *JAMA* 295 (17): 2037–2045. doi:10.1001/jama.295.17.2037.

Benzeval, M., J. Taylor, and K. Judge. 2000. "Evidence on the Relationship between Low Income and Poor Health: Is the Government Doing Enough?" *Fiscal Studies* 21 (3): 375–399. doi:10.1111/j.1475-5890.2000.tb00029.x.

Benzeval, M., and K. Judge. 2001. "Income and Health: The Time Dimension." *Social Science & Medicine* 52 (9): 1371–1390. doi:10.1016/S0277-9536(00)00244-6.

Brasher, M. S., L. K. George, X. Shi, Z. Yin, and Y. Zeng. 2017. "Incorporating Biomarkers into the Study of Socio-economic Status and Health among Older Adults in China." *SSM - Population Health* 3: 577–585. doi:10.1016/j.ssmph.2017.07.003.

Brenner, D. R., D. Scherer, K. Muir, J. Schildkraut, P. Boffetta, M. R. Spitz, … R. J. Hung. 2014. "A Review of the Application of Inflammatory Biomarkers in Epidemiologic Cancer Research." *Cancer Epidemiology, Biomarkers & Prevention* 23 (9): 1729–1751. doi:10.1158/1055-9965.EPI-14-0064.

Breusch, T. S., and A. R. Pagan. 1979. "A Simple Test for Heteroscedasticity and Random Coefficient Variation." *Econometrica* 47 (5): 1287–1294. doi:10.2307/1911963.

Carrieri, V., and A. M. Jones. 2017. "The Income–health Relationship 'Beyond the Mean': New Evidence from Biomarkers." *Health Economics* 26 (7): 937–956. doi:10.1002/hec.3372.

Cohen, A. A., E. Milot, J. Yong, C. L. Seplaki, T. Fülöp, K. Bandeen-Roche, and L. P. Fried. 2013. "A Novel Statistical Approach Shows Evidence for Multi-system Physiological Dysregulation during Aging." *Mechanisms of Ageing and Development* 134 (3): 110–117. doi:10.1016/j.mad.2013.01.004.

Cohen, A. A., E. Milot, Q. Li, V. Legault, L. P. Fried, and L. Ferrucci. 2014. "Cross-population Validation of Statistical Distance as a Measure of Physiological Dysregulation during Aging." *Experimental Gerontology* 57: 203–210. doi:10.1016/j.exger.2014.04.016.

Cohen, A. A., Q. Li, E. Milot, M. Leroux, S. Faucher, V. Morissette-Thomas, L. Ferrucci, L. P. Fried, and L. Ferrucci. 2015. "Statistical Distance as a Measure of Physiological Dysregulation Is Largely Robust to Variation in Its Biomarker Composition." *PLOS ONE* 10 (4): e0122541. doi:10.1371/journal.pone.0122541.

Contoyannis, P., A. M. Jones, and N. Rice. 2004. "The Dynamics of Health in the British Household Panel Survey." *Journal of Applied Econometrics* 19 (4): 473–503. doi:10.1002/jae.755.

Cutler, D. M., A. Lleras-Muney, and T. Vogl. 2008. *Socioeconomic Status and Health: Dimensions and Mechanisms. NBER Working Paper No. 14333*. Cambridge, MA: National Bureau of Economic Research.

Davillas, A., A. M. Jones, and M. Benzeval (2017). *The income-health gradient: Evidence from self-reported health and biomarkers using longitudinal data on income*. "Health, Econometrics and Data Group (HEDG) Working Papers 17/04, HEDG, C/o Department of Economics, University of York." University of York.

Deaton, A. S. 2006. *Global Patterns of Income and Health: Facts, Interpretations, and Policies*. Cambridge: National Bureau of Economic Research.

Deaton, A. S., and C. H. Paxson. 1998. "Aging and Inequality in Income and Health." *American Economic Review* 88 (2): 248–253.

Ettner, S. L. 1996. "New Evidence on the Relationship between Income and Health." *Journal of Health Economics* 15 (1): 67–85. doi:10.1016/0167-6296(95)00032-1.

Evans, W., B. Wolfe, and N. Adler. 2012. "The SES and Health Gradient: A Brief Review of the Literature." In *The Biological Consequences of Socioeconomic Inequalities*, edited by W. Evans, B. Wolfe, and T. E. Seeman, 1–37. New York: Russell Sage Foundation.

Firpo, S., N. M. Fortin, and T. Lemieux. 2009. "Unconditional Quantile Regressions." *Econometrica* 77 (3): 953–973. doi:10.3982/ECTA6822.

Frijters, P., J. P. Haisken-denew, and M. A. Shields. 2005. "The Causal Effect of Income on Health: Evidence from German Reunification." *Journal of Health Economics* 24 (5): 997–1017. doi:10.1016/j.jhealeco.2005.01.004.

Gardner, J., and A. J. Oswald. 2007. "Money and Mental Wellbeing: A Longitudinal Study of Medium-sized Lottery Wins." *Journal of Health Economics* 26 (1): 49–60. doi:10.1016/j.jhealeco.2006.08.004.

Gong, P., S. Liang, E. J. Carlton, Q. Jiang, J. Wu, L. Wang, and J. V. Remais. 2012. "Urbanisation and Health in China." *The Lancet* 379 (9818): 843–852. doi:10.1016/S0140-6736(11)61878-3.

Grossman, M. 1972a. The Demand for Health: A Theoretical and Empirical Investigation. New York: Columbia University Press

Grossman, M. 1972b. "On the Concept of Health Capital and the Demand for Health." *Journal of Political Economy* 80 (2): 223–255. doi:10.1086/259880.

Gunasekara, F. I., K. Carter, and T. Blakely. 2011. "Change in Income and Change in Self-rated Health: Systematic Review of Studies Using Repeated Measures to Control for Confounding Bias." *Social Science & Medicine* 72 (2): 193–201. doi:10.1016/j.socscimed.2010.10.029.

Johnston, D. W., C. Propper, and M. A. Shields. 2009. "Comparing Subjective and Objective Measures of Health: Evidence from Hypertension for the Income/health Gradient." *Journal of Health Economics* 28 (3): 540–552. doi:10.1016/j.jhealeco.2009.02.010.

Jolliffe, D. 2011. "Overweight and Poor? On the Relationship between Income and the Body Mass Index." *Economics and Human Biology* 9 (4): 342–355. doi:10.1016/j.ehb.2011.07.004.

Jones, A. M., and J. Wildman. 2008. "Health, Income and Relative Deprivation: Evidence from the BHPS." *Journal of Health Economics* 27 (2): 308–324. doi:10.1016/j.jhealeco.2007.05.007.

Lewbel, A. 2012. "Using Heteroscedasticity to Identify and Estimate Mismeasured and Endogenous Regressor Models." *Journal of Business and Economic Statistics* 30 (1): 67–80. doi:10.1080/07350015.2012.643126.

Li, H.B., & Zhu, Y. (2006). Income, income inequality, and health: Evidence from China. Journal of Comparative Economics, 34: 668–693. doi:10.1016/j.jce.2006.08.005

Li, Q., S. Wang, E. Milot, P. Bergeron, L. Ferrucci, L. P. Fried, and A. A. Cohen. 2015. "Homeostatic Dysregulation Proceeds in Parallel in Multiple Physiological Systems." *Aging Cell* 14 (6): 1103–1112. doi:10.1111/acel.12402.

Lindahl, M. 2005. "Estimating the Effect of Income on Health and Mortality Using Lottery Prizes as an Exogenous Source of Variation in Income." *The Journal of Human Resources* 40 (1): 144–168. doi:10.3368/jhr.XL.1.144.

Liu, Z.Y. (2020). Development and validation of 2 composite aging measures using routine clinical biomarkers in the Chinese population: Analyses from 2 prospective cohort studies, The Journals of Gerontology: Series A, glaa238. doi:10.1093/gerona/glaa238

Lowry, D., and Y. Xie. 2009. *Socioeconomic Status and Health Differentials in China: Convergence or Divergence at Older Ages (Pp. 1–23)*. Ann Arbor: Population Studies Center, University of Michigan.

Mceniry, M. 2013. *Cross-national Comparisons of Health Disparities among Aging Populations in Latin America, the Caribbean, Asia and Africa (Pp. 1–35)*. Ann Arbor. Michigan: Population Studies Center University of Michigan.

Monteiro, C. A., E. C. Moura, W. L. Conde, and B. M. Popkin. 2004. "Socioeconomic Status and Obesity in Adult Populations of Developing Countries: A Review." *Bulletin of the World Health Organization* 82 (12): 940–946.

Muennig, P., N. Sohler, and B. Mahato. 2007. "Socioeconomic Status as an Independent Predictor of Physiological Biomarkers of Cardiovascular Disease: Evidence from NHANES." *Preventive Medicine* 45 (1): 35–40. doi:10.1016/j.ypmed.2007.04.005.

National Bureau of Statistics. 2018. *China Statistical Yearbook 2018*. Beijing: China Statistics Press.

Nie, P., L. L. Ding, and A. Sousa-Poza. 2019. "Decomposing Adult Obesity Trends in China (1991–2011)." *Economics and Human Biology* 34: 5–15. doi:10.1016/j.ehb.2019.02.001.

Papagianni, M., and K. Tziomalos. 2018. "Healthcare Reform in China: Challenges and Opportunities." *Current Medical Research and Opinion* 34 (5): 821–823. doi:10.1080/03007995.2017.1401988.

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. doi:10.1016/S0140-6736(07)61778-4.

Qi, L. S. 2006. "Income, Income Inequality and Health: The Impacts of Rural-urban Gap and Occupational Status." *Economic Research Journal* 11: 16–26. (in Chinese)..

Rahkovsky, I., and C. A. Gregory. 2013. "Food Prices and Blood Cholesterol." *Economics and Human Biology* 11 (1): 95–107. doi:10.1016/j.ehb.2012.01.004.

Rarick, J. R. D., C. T. Dolan, W. J. Han, and J. Wen. 2017. "Relations between Socioeconomic Status, Subjective Social Status, and Health in Shanghai, China." *Social Science Quarterly* 99 (1): 390–405. doi:10.1111/ssqu.12360.

Rosero-Bixby, L., and W. H. Dow. 2009. "Surprising SES Gradients in Mortality, Health, and Biomarkers in a Latin American Population of Adults." *The Journals of Gerontology: Series B* 64B (1): 105–117. doi:10.1093/geronb/gbn004.

Rossouw, L., 2018. Bago d'Uva, T., & van Doorslaer, E. (2018). "Poor Health Reporting? Using Anchoring Vignettes to Uncover Health Disparities by Wealth and Race." *Demography* 55 (5). doi:10.1007/s13524-018-0709-0.

Smith, K. V., and N. Goldman. 2007. "Socioeconomic Differences in Health among Older Adults in Mexico." *Social Science & Medicine* 65 (7): 1372–1385. doi:10.1016/j.socscimed.2007.05.023.

Ticinesi, A., F. Lauretani, A. Nouvenne, E. Porro, G. Fanelli, M. Maggio, and T. Meschi. 2017. "C-reactive Protein (CRP) Measurement in Geriatric Patients Hospitalized for Acute Infection." *European Journal of Internal Medicine* 37: 7–12. doi:10.1016/j.ejim.2016.08.026.

Van Kippersluis, H., and T. J. Galama. 2014. "Wealth and Health Behavior: Testing the Concept of a Health Cost." *European Economic Review* 72: 197–220. doi:10.1016/j.euroecorev.2014.10.003.

Xi, B., Y. Liang, T. He, K. H. Reilly, Y. Hu, Q. Wang, Y. Yan, and J. Mi. 2012. "Secular Trends in the Prevalence of General and Abdominal Obesity among Chinese Adults, 1993–2009." *Obesity Review* 13 (3): 287–296. doi:10.1111/j.1467-789X.2011.00944.x.

Yan, S., J. Li, S. Li, B. Zhang, S. Du, P. Gordon-Larsen, L. Adair, et al. 2012. "The Expanding Burden of Cardiometabolic Risk in China: The China Health and Nutrition Survey." *Obesity Reviews* 13 (9): 810–821. doi:10.1111/j.1467-789X.2012.01016.x.

Yang, W., and P. Kanavos. 2012. "The Less Healthy Urban Population: Income-related Health Inequality in China." *BMC Public Health* 12 (1): 804. doi:10.1186/1471-2458-12-804.

Zhang, B., F. Y. Zhai, S. F. Du, and B. M. Popkin. 2014. "The China Health and Nutrition Survey, 1989–2011." *Obesity Reviews* 15 (S1): 2–7. doi:10.1111/obr.12119.

Zimmer, Z., N. Chayovan, H. S. Lin, and J. Natividad. 2004. "How Indicators of Socioeconomic Status Relate to Physical Functioning of Older Adults in Three Asian Societies." *Research on Aging* 26 (2): 224–258. doi:10.1177/0164027503260624.

Appendix A :

Table A1. Descriptive statistics

Variable	Obs.	Mean	Std. Dev.
Health			
HbA1C (mmol/l)	6,730	5.645	0.919
Cholesterol ratio	6,760	3.614	1.157
CRP	5,089	2.716	2.926
WBC	6,804	6.292	1.922
Log (current income)	6,730	9.060	1.064
Log (permanent income)	6,730	8.503	1.186
Male: 18–24	6,730	0.004	0.067
Male: 25–29	6,730	0.011	0.106
Male: 30–34	6,730	0.025	0.157
Male: 35–39	6,730	0.050	0.218
Male: 40–44	6,730	0.063	0.242
Male: 45–49	6,730	0.065	0.247
Male: 50–54	6,730	0.069	0.253
Male: 55–59	6,730	0.072	0.258
Male: 60–64	6,730	0.056	0.231
Male: 65+	6,730	0.106	0.307
Female: 18–24	6,730	0.009	0.096
Female: 25–29	6,730	0.016	0.125
Female: 30–34	6,730	0.027	0.163
Female: 35–39	6,730	0.046	0.210
Female: 40–44	6,730	0.052	0.222
Female: 45–49	6,730	0.058	0.233
Female: 50–54	6,730	0.061	0.239
Female: 55–59	6,730	0.064	0.245
Female: 60–64	6,730	0.050	0.218
Female: 65+	6,730	0.095	0.293
Education: illiterate	6,730	0.266	0.442
Education: primary school	6,730	0.207	0.405
Education: middle school	6,730	0.319	0.466
Education: high school	6,730	0.105	0.307
Education: vocational school	6,730	0.061	0.240
Education: university or higher	6,730	0.042	0.200
Marital status: never married	6,730	0.030	0.170
Marital status: married	6,730	0.867	0.339
Marital status: divorced/widowed/separated	6,730	0.103	0.304
Household size	6,730	3.678	1.667
Province: Liaoning	6,730	0.093	0.290
Province: Heilongjiang	6,730	0.109	0.311
Province: Jiangsu	6,730	0.129	0.335
Province: Shandong	6,730	0.111	0.314
Province: Henan	6,730	0.108	0.310
Province: Hubei	6,730	0.115	0.319
Province: Hunan	6,730	0.120	0.325
Province: Guangxi	6,730	0.118	0.322
Province: Guizhou	6,730	0.099	0.299

Notes: Based on data from the 2009 wave of the CHNS.
[†]The number of observations differs dependent on the availability of CHNS biomarker data.

Table A2. Probit estimates for macronutrients among adults aged 18 +

	Calories (1 ≥ 2000 kcal)	Carbohydrates (1 ≥ 150 g)	Fat (1 ≥ 78 g)	Protein (1 ≥ 56 g)
Ln(current income)	0.011	−0.00004	0.028**	0.023**
	(0.006)	(0.003)	(0.006)	(0.006)
N	6827	6827	6827	6827
Pseudo R^2	0.092	0.098	0.056	0.088
Ln(permanent income)	0.001	−0.003	0.021**	0.017**
	(0.005)	(0.003)	(0.006)	(0.005)
N	6827	6827	6827	6827
Pseudo R^2	0.092	0.099	0.056	0.087
Ln(permanent income)	−0.003	−0.001	−0.002	−0.003
	(0.007)	(0.003)	(0.007)	(0.007)
Individual-specific effects	0.012	−0.006	0.066**	0.059**
	(0.012)	(0.005)	(0.013)	(0.012)
N	6827	6827	6827	6827
Pseudo R^2	0.092	0.099	0.059	0.090

Notes: The dependent variables are dummies for calories, carbohydrates, fat and protein. The controls include individual characteristics (age-gender dummies, education level, marital status, and employment status), translog household equivalized income, household size, and provincial dummies (with Liaoning as the reference). Standard errors are in parentheses. * $p < 0.05$, ** $p < 0.01$.

Table A3. Lewbel's heteroscedasticity-based 2SLS estimates for individual biomarkers and composite measure of health among adults aged 18 +

Panel A	HbA1c	Cholesterol ratio	CRP	WBC	Composite health
Ln(current income)	−0.037	−0.006	0.073	−0.033	0.005
	(0.026)	(0.034)	(0.103)	(0.059)	(0.008)
Controls	YES	YES	YES	YES	YES
First-stage estimate					
R^2	0.331	0.334	0.328	0.326	0.327
F-statistics	35.99	37.31	28.24	36.51	36.25
Hansen's J p-value	0.806	0.950	0.557	0.623	0.010
N	6730	6760	5089	6804	6651
Panel B	HbA1c	Cholesterol ratio	CRP	WBC	Composite health
Ln(permanent income)	−0.024	0.014	0.052	−0.030	0.003
	(0.025)	(0.033)	(0.100)	(0.056)	(0.008)
Controls	YES	YES	YES	YES	YES
First-stage estimate					
R^2	0.353	0.355	0.353	0.349	0.351
F-statistics	31.59	32.74	24.88	32.93	33.03
Hansen's J p-value	0.582	0.816	0.212	0.853	0.760
N	6730	6760	5089	6804	6651

Notes: The dependent variables are HbA1C, cholesterol ratio, CRP, white blood-cell count, and the translog composite measure of health constructed with 22 biomarkers (abumin, alanine aminotransferase, apolipoprotein A-1, creatinine, ferritin, glucose, high-density lipoprotein cholesterol, insulin, low-density lipoprotein cholesterol, serum magnesium, total cholesterol, triglyceride, total protein, transferrin, soluble transferrin receptor, uric acid, urea, haemoglobin, white blood cell, red blood cell, platelet count, and haemoglobin A1C). The controls include individual characteristics (age-gender dummies, education level, marital status, and employment status), translog household equivalized income, household size, and provincial dummies (with Liaoning as the reference). For the estimates of HbA1C and cholesterol ratio, we also control antidiabetes medication. Robust standard errors are in parentheses.

Table A4. OLS and RIF estimates for HbA1c and cholesterol ratio among adults aged 18+ in urban China

Panel A: HbA1c	OLS	25th	50th	75th	90th	95th
Ln(current income)	0.031	−0.008	0.010	0.026	0.076	0.030
	(0.017)	(0.011)	(0.012)	(0.017)	(0.059)	(0.102)
N	2098	2098	2098	2098	2098	2098
Adj. R^2	0.271	0.116	0.172	0.207	0.236	0.283
Ln(permanent income)	0.034*	−0.009	0.011	0.025	0.079	0.022
	(0.015)	(0.011)	(0.011)	(0.015)	(0.051)	(0.086)
N	2098	2098	2098	2098	2098	2098
Adj. R^2	0.271	0.116	0.172	0.207	0.236	0.283
Ln(permanent income)	0.047*	0.005	0.025	0.027	0.119	0.040
	(0.019)	(0.014)	(0.014)	(0.022)	(0.061)	(0.130)
Individual-specific effects	−0.033	−0.035	−0.036	−0.005	−0.103	−0.047
	(0.033)	(0.024)	(0.023)	(0.035)	(0.108)	(0.179)
N	2098	2098	2098	2098	2098	2098
Adj. R^2	0.271	0.117	0.173	0.207	0.236	0.282
Panel B: Cholesterol ratio	OLS	25th	50th	75th	90th	95th
Ln(current income)	0.024	0.021	0.016	0.073*	0.028	0.045
	(0.027)	(0.026)	(0.032)	(0.036)	(0.054)	(0.090)
N	2113	2113	2113	2113	2113	2113
Adj. R^2	0.080	0.068	0.072	0.052	0.029	0.026
Ln(permanent income)	0.035	0.025	0.029	0.083*	0.052	0.077
	(0.025)	(0.024)	(0.028)	(0.034)	(0.051)	(0.084)
N	2113	2113	2113	2113	2113	2113

(*Continued*)

(Continued).

Panel A: HbA1c	OLS	25th	50th	75th	90th	95th
Adj. R^2	0.080	0.068	0.072	0.053	0.030	0.026
Ln(permanent income)	0.016	0.010	0.034	0.072	0.021	−0.048
	(0.032)	(0.032)	(0.036)	(0.045)	(0.065)	(0.104)
Individual-specific effects	0.052	0.040	−0.013	0.030	0.083	0.334*
	(0.052)	(0.055)	(0.059)	(0.076)	(0.124)	(0.162)
N	2113	2113	2113	2113	2113	2113
Adj. R^2	0.080	0.068	0.072	0.053	0.029	0.028

Notes: The dependent variables are HbA1c and cholesterol ratio. The controls include individual characteristics (age-gender dummies, education level, marital status, employment status, and antidiabetes medication), translog household equivalized income, household size, and provincial dummies (with Liaoning as the reference). Robust standard errors for the OLS estimates are in parentheses; standard errors for the UQR estimates are bootstrapped with 500 replications. * $p < 0.05$, ** $p < 0.01$.

Table A5. OLS and RIF estimates for HbA1c and cholesterol ratio among adults aged 18+ in rural China

Panel A: HbA1c	OLS	25th	50th	75th	90th	95th
Ln(current income)	−0.011	0.003	0.002	−0.002	−0.029	−0.070
	(0.012)	(0.009)	(0.008)	(0.012)	(0.026)	(0.064)
N	4632	4632	4632	4632	4632	4632
Adj. R^2	0.168	0.129	0.196	0.180	0.114	0.121
Ln(permanent income)	−0.007	0.005	0.005	0.008	−0.007	−0.014
	(0.012)	(0.008)	(0.008)	(0.011)	(0.025)	(0.061)
N	4632	4632	4632	4632	4632	4632
Adj. R^2	0.168	0.129	0.196	0.180	0.113	0.121
Ln(permanent income)	−0.012	0.001	0.001	0.001	−0.012	0.024
	(0.015)	(0.010)	(0.010)	(0.013)	(0.029)	(0.078)
Individual-specific effects	0.016	0.014	0.013	0.023	0.015	−0.118
	(0.026)	(0.019)	(0.017)	(0.023)	(0.051)	(0.139)
N	4632	4632	4632	4632	4632	4632
Adj. R^2	0.168	0.129	0.196	0.180	0.113	0.121
Panel B: Cholesterol ratio	OLS	25th	50th	75th	90th	95th
Ln(current income)	0.003	0.013	0.001	−0.025	−0.008	−0.077
	(0.017)	(0.016)	(0.019)	(0.027)	(0.042)	(0.073)
N	4647	4647	4647	4647	4647	4647
Adj. R^2	0.070	0.045	0.061	0.051	0.029	0.014
Ln(permanent income)	0.009	0.014	0.007	−0.013	−0.002	−0.057
	(0.016)	(0.015)	(0.018)	(0.027)	(0.038)	(0.063)
N	4647	4647	4647	4647	4647	4647
Adj. R^2	0.070	0.045	0.061	0.051	0.029	0.014
Ln(permanent income)	−0.026	−0.007	−0.029	−0.079*	−0.058	−0.079
	(0.020)	(0.019)	(0.021)	(0.032)	(0.047)	(0.083)
Individual-specific effects	0.110**	0.065	0.113**	0.204**	0.175	0.068
	(0.037)	(0.036)	(0.040)	(0.059)	(0.090)	(0.149)
N	4647	4647	4647	4647	4647	4647
Adj. R^2	0.072	0.046	0.063	0.054	0.029	0.014

Notes: The dependent variables are HbA1c and cholesterol ratio. The controls include individual characteristics (age-gender dummies, education level, marital status, employment status, and antidiabetes medication), translog household equivalized income, household size, and provincial dummies (with Liaoning as the reference). Robust standard errors for the OLS estimates are in parentheses; standard errors for the UQR estimates are bootstrapped with 500 replications. * $p < 0.05$, ** $p < 0.01$.

Table A6. OLS and RIF estimates for composite measure of health among adults aged 18+ (urban versus rural)

Panel A: Urban	OLS	25th	50th	75th	90th	95th
Ln(current income)	0.001	−0.004	0.001	0.004	0.013	0.032
	(0.006)	(0.006)	(0.007)	(0.008)	(0.014)	(0.019)
N	2071	2071	2071	2071	2071	2071
Adj. R^2	0.067	0.059	0.059	0.036	0.021	0.014
Ln(permanent income)	0.004	−0.004	0.004	0.007	0.015	0.032
	(0.005)	(0.006)	(0.006)	(0.008)	(0.014)	(0.021)
N	2071	2071	2071	2071	2071	2071
Adj. R^2	0.068	0.059	0.059	0.036	0.021	0.015
Ln(permanent income)	0.007	−0.002	0.007	0.011	0.019	0.035
	(0.007)	(0.008)	(0.008)	(0.010)	(0.018)	(0.029)
Individual-specific effects	−0.008	−0.005	−0.009	−0.009	−0.009	−0.008
	(0.012)	(0.013)	(0.014)	(0.018)	(0.032)	(0.051)
N	2071	2071	2071	2071	2071	2071
Adj. R^2	0.067	0.059	0.059	0.036	0.021	0.014
Panel B: Rural	OLS	25th	50th	75th	90th	95th
Ln(current income)	0.002	−0.001	0.003	0.004	0.005	−0.001
	(0.004)	(0.005)	(0.004)	(0.005)	(0.008)	(0.013)
N	4580	4580	4580	4580	4580	4580
Adj. R^2	0.033	0.031	0.036	0.026	0.006	0.002
Ln(permanent income)	0.002	−0.002	0.005	0.004	0.006	0.002
	(0.004)	(0.004)	(0.004)	(0.005)	(0.007)	(0.012)
N	4580	4580	4580	4580	4580	4580
Adj. R^2	0.033	0.031	0.036	0.026	0.006	0.002
Ln(permanent income)	0.004	−0.001	0.007	0.004	0.008	0.003
	(0.004)	(0.005)	(0.005)	(0.006)	(0.010)	(0.015)
Individual-specific effects	−0.004	−0.006	−0.007	−0.001	−0.006	−0.004
	(0.008)	(0.010)	(0.009)	(0.010)	(0.019)	(0.027)
N	4580	4580	4580	4580	4580	4580
Adj. R^2	0.033	0.031	0.036	0.026	0.006	0.002

Notes: The dependent variable is the translog composite measure of health, constructed with 22 biomarkers (abumin, alanine aminotransferase, apolipoprotein A-1, creatinine, ferritin, glucose, high-density lipoprotein cholesterol, insulin, low-density lipoprotein cholesterol, serum magnesium, total cholesterol, triglyceride, total protein, transferrin, soluble transferrin receptor, uric acid, urea, haemoglobin, white blood cell, red blood cell, platelet count, and haemoglobin A1C). The controls include individual characteristics (age-gender dummies, education level, marital status, and employment status), translog household equivalized income, household size, and provincial dummies (with Liaoning as the reference). Robust standard errors for the OLS estimates are in parentheses; standard errors for the UQR estimates are bootstrapped with 500 replications. * $p < 0.05$, ** $p < 0.01$.

Table A7. OLS and RIF estimates for individual biomarkers and composite health among adults aged 18 +

Panel A: HbA1C	OLS	25th	50th	75th	90th	95th
Income quartile 2	−0.059	−0.023	0.001	−0.033	−0.116	−0.313*
	(0.030)	(0.020)	(0.017)	(0.029)	(0.059)	(0.144)
Income quartile 3	−0.001	0.014	0.014	0.013	0.022	0.021
	(0.030)	(0.019)	(0.020)	(0.029)	(0.064)	(0.149)
Income quartile 4	0.010	0.001	0.006	0.056	0.044	0.049
	(0.030)	(0.020)	(0.020)	(0.029)	(0.069)	(0.158)
N	6730	6730	6730	6730	6730	6730
Adj. R^2	0.195	0.124	0.184	0.181	0.147	0.169
Panel B: Cholesterol ratio	OLS	25th	50th	75th	90th	95th
Income quartile 2	−0.097*	−0.038	−0.068	−0.080	−0.232*	−0.511**
	(0.040)	(0.037)	(0.045)	(0.061)	(0.094)	(0.146)

(*Continued*)

(Continued).

Panel A: HbA1C	OLS	25th	50th	75th	90th	95th
Income quartile 3	−0.020	−0.017	−0.050	0.024	−0.087	−0.060
	(0.040)	(0.038)	(0.046)	(0.063)	(0.092)	(0.165)
Income quartile 4	0.057	0.064	0.046	0.098	0.031	−0.047
	(0.043)	(0.041)	(0.048)	(0.069)	(0.101)	(0.165)
N	6760	6760	6760	6760	6760	6760
Adj. R^2	0.075	0.050	0.065	0.050	0.029	0.021
Panel C: CRP	OLS	25th[a]	50th	75th	90th	95th
Income quartile 2	−0.241*		−0.126*	−0.145	−0.715*	−0.530
	(0.119)		(0.061)	(0.108)	(0.298)	(0.502)
Income quartile 3	−0.027		−0.032	0.145	−0.273	−0.090
	(0.122)		(0.057)	(0.113)	(0.332)	(0.515)
Income quartile 4	−0.057		−0.019	0.089	−0.401	−0.149
	(0.129)		(0.066)	(0.123)	(0.345)	(0.525)
N	5089		5089	5089	5089	5089
Adj. R^2	0.014		0.011	0.014	0.010	0.004
Panel D: WBC	OLS	25th	50th	75th	90th	95th
Income quartile 2	−0.064	−0.107	−0.087	0.034	0.061	−0.137
	(0.067)	(0.061)	(0.065)	(0.084)	(0.163)	(0.211)
Income quartile 3	−0.127*	−0.097	−0.108	−0.094	−0.264	−0.287
	(0.064)	(0.062)	(0.071)	(0.092)	(0.150)	(0.215)
Income quartile 4	−0.048	0.023	−0.033	0.021	−0.030	−0.043
	(0.067)	(0.070)	(0.072)	(0.094)	(0.157)	(0.235)
N	6804	6804	6804	6804	6804	6804
Adj. R^2	0.049	0.043	0.044	0.032	0.017	0.009
Panel E: Composite health	OLS	25th	50th	75th	90th	95th
Income quartile 2	−0.004	−0.020*	0.002	−0.010	−0.000	0.000
	(0.009)	(0.009)	(0.010)	(0.012)	(0.020)	(0.032)
Income quartile 3	0.001	−0.005	0.012	−0.005	−0.008	0.004
	(0.009)	(0.010)	(0.010)	(0.013)	(0.022)	(0.033)
Income quartile 4	0.001	−0.019	0.007	0.003	0.012	0.028
	(0.010)	(0.011)	(0.011)	(0.013)	(0.022)	(0.035)
N	6651	6651	6651	6651	6651	6651
Adj. R^2	0.038	0.037	0.037	0.024	0.007	0.003

Notes: The dependent variables are HbA1C, cholesterol ratio, CRP, WBC and composite health. Controls include individual characteristics (including age-gender dummies, education level, marital status and employment status), translog household income, household size and provincial dummies (Liaoning as the reference province). For HbA1C and cholesterol ratio, antidiabetes medication is also controlled. For OLS estimates, robust standard errors are in parentheses, and UQR standard errors are bootstrap estimates with 500 replications. * $p < 0.05$, ** $p < 0.01$.

a The 25th percentile cannot be estimated for CRP because just under 50% of the observations have the minimum value of 1.

Table A8. OLS and RIF estimates for BMI among adults aged 18+ in urban China

	OLS	25th	50th	75th	90th	95th
Ln(current income)	0.204***	0.172***	0.206***	0.241***	0.198***	0.393***
	(0.043)	(0.053)	(0.056)	(0.066)	(0.089)	(0.134)
N	6685	6685	6685	6685	6685	6685
Adj. R^2	0.101	0.071	0.072	0.045	0.024	0.013
Ln(permanent income)	0.194***	0.185***	0.218***	0.209***	0.194***	0.417***
	(0.039)	(0.047)	(0.052)	(0.060)	(0.081)	(0.124)
N	6685	6685	6685	6685	6685	6685
Adj. R^2	0.102	0.078	0.073	0.045	0.024	0.014
Ln(permanent income)	0.036	0.068	0.075	−0.009	0.017	0.253
	(0.049)	(0.060)	(0.063)	(0.077)	(0.099)	(0.149)
Individual-specific effects	0.451***	0.335***	0.408***	0.620***	0.502***	0.467
	(0.688)	(0.104)	(0.109)	(0.138)	(0.174)	(0.285)
N	6685	6685	6685	6685	6685	6685
Adj. R^2	0.105	0.079	0.075	0.048	0.025	0.014

Notes: The dependent variable is BMI. The controls include individual characteristics (age-gender dummies, education level, marital status and employment status), translog household equivalized income, household size, and provincial dummies (with Liaoning as the reference). Robust standard errors for the OLS estimates are in parentheses; standard errors for the UQR estimates are bootstrapped with 500 replications. * p < 0.05, ** p < 0.01.

Appendix B

Lewbel's (2012) heteroscedasticity-based two-stage least squares (2SLS) estimation

As a robustness check, we also adopt Lewbel's (2012) heteroscedasticity-based 2SLS to explore the causal relation between income and health. We first consider a structural model of the form below:

$$Y_1 = X'\alpha_1 + Y_2\gamma_1 + \varepsilon_1 \quad (1)$$

$$Y_2 = X'\alpha_2 + \varepsilon_2, where \varepsilon_2 = \rho_2 U + \omega_2 \quad (2)$$

In our case, Y_1 is the health outcome and Y_2 is per capita household total income, U represents unobserved factors such as individual motivation or genetics, and ε_1 and ε_2 are idiosyncratic error terms. As Lewbel (2012) suggests, we can take a vector Z of observed exogenous variables and employ $[Z\text{-}E(Z)]\varepsilon_2$ as an instrument if

$$E(X\varepsilon_1) = 0, E(X\varepsilon_2) = 0, cov(Z, \varepsilon_1, \varepsilon_2) = 0 \quad (3)$$

The rationale for using $[Z\text{-}E(Z)]\varepsilon_2$ as an instrument is that identification can be achieved by obtaining regressors that are uncorrelated with the product of the heteroscedastic errors (Lewbel 2012). In practice, Z could either be a subset of X or equal to X. We use the latter case for our IV estimation. Drawing on this instrument, we can use 2SLS to run the IV estimation even without the existence of conventional IVs.