On the Measurement of Health and Its Effect on the Measurement of Health Inequality*

Erik T. Nesson†
Joshua J. Robinson‡

October 2, 2019

Abstract

We examine the extent to which self-reported health measures suffer from income-related reporting heterogeneity and then characterize how this reporting heterogeneity affects the estimation of income-related health inequality. We run a comprehensive set of tests of reporting heterogeneity using several self-reported health measures and several clinical measures of health from the National Health and Nutritional Examination Surveys. We propose the use of a multidimensional measure using clinical indicators of health in the context of measuring income-related health inequality, and we examine the extent of income-related health inequality, as measured by the concentration index, using both self-reported measures of health and the multidimensional clinical measure. Our results confirm the existence of significant, positive, income-related reporting heterogeneity and also suggest that higher income individuals react more strongly to a change in clinical health measures. Using self-assessed health suggests that income-related health inequality is about three times larger than when using more objective, self-reported health measures and ten times larger than when using the multidimensional clinical measure of health.

Keywords: Health Inequality, Concentration Index, Self-Reported Health, Self-Assessed Health, Allostatic Load
JEL Codes: C43, I12, I14, I18

*This work is supported by the Robert Wood Johnson Foundation (grant #72073). The views herein reflect the views of the authors and do not necessarily reflect those of the Robert Wood Johnson Foundation. We would like to thank Esfandiar Maasoumi, Kerry Anne McGeary, and participants at the 2014 Southern Economic Association Meetings and a seminar at Xavier University for excellent comments and Will Davis, Daniel Mowan, and August Clevenger for excellent research assistance.

†Department of Economics, Miller College of Business, Ball State University, Muncie, IN and NBER (E-Mail: etnesson@bsu.edu).

‡Corresponding author. Department of Marketing, Industrial Distribution, and Economics, Collat School of Business, University of Alabama at Birmingham, Birmingham, AL (E-Mail: jjr@uab.edu).

1 Introduction

The measurement of health, health inequality, and the relationship between health inequality and socioeconomic status (SES) have risen to the forefront of many researchers' and policy makers' minds in the past few years. While a large literature examines the statistical properties of health inequality measurements, one remaining issue is what measure of health should be used in these models. To measure income-related health inequality accurately requires a representative measure of health that does not suffer from reporting heterogeneity. Reporting heterogeneity is a problem that can occur for a self-reported measure if people are unable (or unwilling) to objectively report information. Health inequality measurements often use categorical self-assessed health status (SAH) as the measurement of health, despite its categorical nature and the possibility that SES may affect individuals' perception of their health status.

We examine how reporting heterogeneity affects the degree of income-related health inequality as measured by the concentration index. We use data from the National Health and Nutritional Examination Surveys (NHANES) which contain measures of self-reported health, including SAH, results from physical examinations, and blood, saliva, and urine samples.[1] We find evidence of income-related reporting heterogeneity in SAH—specifically that better responses about one's health are positively associated with income while holding clinical measures of health constant. We show this result is robust across three separate methods: (1) cluster analysis, (2) Lagrange multiplier tests developed by Lindeboom and van Doorslaer (2004), and (3) Chow tests from regressions of self-reported health on clinical health measures and demographic variables fully interacted with income. Each approach provides a unique perspective on the nature of reporting heterogeneity.

Consistent with this reporting heterogeneity, we also find that using SAH as a proxy for health suggests there is significantly more income-related health inequality than is sug-

[1]We use the term *self-reported health* to refer to any survey response about one's health, including responses to questions about physically healthy and mentally healthy days, while SAH refers to questions specific to rating one's overall health. Thus, SAH is one example of self-reported health.

gested by using indicators of health commonly used by physicians in a clinical setting. Our results together suggest that the use of SAH can overstate the degree of SES-related health inequality and that a more objective, multivariate measure is needed. We adapt a simple method for constructing such a measure of health based on increasingly available, clinical indicators of health, which we argue are more objective than self-reported measures. We combine these indicators based on the theory of allostatic load, a measure of the cumulative stresses on multiple physiological systems, which is a strong predictor of future mortality and morbidity (McEwen & Stellar, 1993; McEwen, 1998; Karlamangla et al., 2002; Seeman et al., 2001). This is a ratio scale measure that can be used with the concentration index to measure health inequality, and it addresses two common problems encountered by previous studies: the arbitrary choices of attributes and the arbitrary choice of healthy cut-points. Our adaptation of this measure additionally addresses a third concern also common in the literature: arbitrary weighting.

The rest of this paper proceeds as follows. Section 2 reviews the relevant literature measuring health inequality and our methods, Section 3 describes the NHANES data, Sections 4 and 5 summarize our results, and Section 6 concludes.

2 Background & Methods

2.1 The Concentration Index

The most common metric used to measure income-related inequality in health is the concentration index. This is a generalization of the Gini index applied to any positive and continuously differentiable function of income (Kakwani, 1977, 1980). For measuring SES-related health inequality, the health variable is assumed to be a continuous, positive function of income.

Formally, let $i = \{1, 2, \ldots, n\}$ be a ranking of n individuals according to their socioeconomic status, with 1 assigned to the least well off person and n assigned to most well off

person. The concentration index, C_h, for this population is

$$C_h = \frac{2}{n\mu_h} \sum_{i=1}^{n} h_i R_i - \frac{n+1}{n} \tag{1}$$

where h_i is the health status of an individual of socioeconomic rank i, μ_h is the mean level of health, and $R_i = {}^{i}/_{n}$ is the individual's rank relative to the highest rank. While Equation (1) is the most direct way to calculate the concentration index, its popularity is probably best attributed to the "convenient covariance" formula derived by Kakwani (1980), which shows that $C_h = {}^{2}/_{\mu_h}\text{cov}(h_i, R_i)$. Jenkins (1988) used this observation to show that when using individual-level data, a concentration index can be estimated with a simple linear regression, and Wagstaff et al. (1991) applied this technique to compute the health concentration index:

$$2\sigma_R^2 \left(\frac{h_i}{\mu_h} \right) = \beta_0 + \beta_1 R_i + \boldsymbol{x_i}\boldsymbol{\gamma} + u_i. \tag{2}$$

In this equation, h_i, R_i, and μ_h are defined as above. Additionally, σ_R^2 is the variance of the fractional rank and $\boldsymbol{x_i}$ is a vector of other explanatory variables (Kakwani et al., 1997; O'Donnell et al., 2008). β_1 is equivalent to the concentration index, and the standard error for β_1 allows researchers to conduct statistical inference.[2,3] Researchers have used these related measurements of the relationship between SES and health inequality in a wide range of settings (e.g. Zhang & Wang, 2004; Kennedy et al., 1998; Deaton & Paxson, 1998; Trannoy et al., 2010; Rosa Dias, 2009; Dolores Montoya Diaz, 2002).[4]

[2]Kakwani et al. (1997) note that the standard error associated with β_1 is not exactly correct, as it does not account for correlations in the error structure. This can be remedied by using cluster-robust standard errors (van Doorslaer & Jones, 2003).

[3]Numerous papers have made contributions and modifications to this original framework. Some notable examples are Wagstaff et al. (2003), which extends the framework of the concentration index to show how changes in health inequality can be decomposed into changes in the means and inequalities of the determinants of health inequality and changes in the size of the effects of the determinants on health inequality, and Wagstaff (2002), which proposes a generalization of the concentration index that allows for alternative normative choices concerning the weight given to various parts of the SES distribution.

[4]A number of methodological issues with the concentration index have been identified in the literature. We briefly review these issues in Appendix A.

2.2 Self-Assessed Health

Because health is a latent and multidimensional characteristic, many studies use a subjective health measure like SAH—often reported on a scale of one to five or one to ten—as an index for an individual's health (e.g. Zhang & Wang, 2004; Kennedy et al., 1998; Deaton & Paxson, 1998; Trannoy et al., 2010; Rosa Dias, 2009; Dolores Montoya Diaz, 2002). This presents two primary challenges in the context of measuring health inequality.

First, the concentration index can only be appropriately applied to ratio scale measures (a cardinal measure with an absolute zero), and SAH is an ordinal measure. There are three common methods of cardinalizing SAH for the purpose of measuring health inequality: dichotomization (Wagstaff et al., 1991), log-normal transformation (Wagstaff & van Doorslaer, 1994), and prediction from interval regression (van Doorslaer & Jones, 2003).[5] In this study, we primarily adopt the log-normalization method of cardinalization for SAH. Although any method of transformation has pros and cons, both the dichotomization and interval regression approaches have several methodological problems that are not yet resolved in the literature (Erreygers, 2009a; Ziebarth, 2010). We, therefore, believe that the log-normalization transformation is the least likely to unduly influence our results. However, we also run robustness checks with other transformations in the appendix.

Second, and a more serious set of concerns in our view, involves the reliability of SAH as an indicator for health status and whether its subjective assessments differ according to demographic or, more importantly, SES—a problem commonly referred to as *reporting heterogeneity* (or sometimes called *reporting bias*). The primary focus of this paper is to determine whether reporting heterogeneity exists, to measure the extent of its effects on the measurement of health inequality, and to evaluate an alternative measure of health which is

[5] A technique designed for discrete dependent variables, like ordered probit, can be used to examine the relationships between SAH and SES, but such a technique cannot be used to estimate a concentration index. It is worth noting that Makdissi and Yazbeck (2017) recently developed a stochastic dominance-based method for ranking SES health inequality using a categorical variable, like SAH, that is robust to any monotonic transformation. However, this method likely results in incomplete rankings due to the strict conditions of stochastic dominance.

less likely to suffer from reporting heterogeneity.

Research has shown that SAH is not necessarily a reliable indicator of health. For example, several studies find that there is considerable measurement error in SAH responses (Greene et al., 2015; Crossley & Kennedy, 2002; Zajacova & Dowd, 2011; Black et al., 2017). Groot (2000) finds evidence that adaptations to chronic conditions and pain can change an individual's reference points for SAH, and Frijters and Ulker (2008) find that controlling for individual fixed effects can dramatically change the statistical relationship between SAH and its determinants. Bound (1991) finds that self-reported measures of health suggest a stronger relationship between health and retirement decisions than more objective measures of self-reported health. Additionally, several studies look at the dynamics of health self-assessments and find them to be highly state dependent (e.g. Contoyannis, Jones, & Rice, 2004; Fernández-Val, Savchenko, & Vella, 2017; Harris & Kohn, 2017).

Reporting errors are more problematic if different groups of people, such as genders, ethnicities, ages, or SES groups, systematically self-report their health in different ways. If H_i^* is an individual's true health status, then the individual will choose $SAH_i = j$ if and only if $\alpha_{j-1} < H_i^* \leq \alpha_j$. The implicit assumption is that the reporting thresholds, α_j, are approximately the same for all groups of people. If this assumption does not hold, then SAH measures are not comparable between different groups of people because each group uses different criteria to choose the ordinal value based on their true health. With regard to using SAH to evaluate a health concentration index, SES-related reporting heterogeneity is particularly problematic because the concentration index is a measure of health inequality that is assumed to be driven by SES. If individuals evaluate and/or report their health differently due to differences in their SES, then the concentration index will yield a biased measure of health inequality.

There is a consensus in the literature that reporting heterogeneity exists across both age and gender, as well as the direction of the effects (age is positively correlated with SAH and women tend to report worse health than men). The evidence regarding SES-related

reporting heterogeneity, however, is more mixed, with some papers showing SES positively affects SAH, some papers showing SES negatively affects SAH, and a few papers showing no SES-related reporting heterogeneity. When using the Canadian HUI3, a generic measure, to control for objective health Lindeboom and van Doorslaer (2004) find no evidence of reporting heterogeneity by income or education, Humphries and van Doorslaer (2000) find lower income individuals understate their health, and Layes et al. (2012) find lower income individuals overstate their health. Shmueli (2003) finds that both SAH and SF-36 (another generic health measure) responses are positively biased by income.[6]

Another way to test for reporting heterogeneity is to examine the ability of SAH to predict mortality. Dowd and Zajacova (2007) show that the predictive ability of SAH varies significantly by income and education levels. However, van Doorslaer and Gerdtham (2003) find no evidence of income or education related reporting heterogeneity using a similar method. When conditioning on mortality, Jürges (2008) finds some evidence of income based heterogeneity, but only for women. Some researchers attempt to correct for reporting heterogeneity using health vignettes, in which respondents are asked to rate the health of a hypothetical person to serve as a reference point for their own self-assessment. Bago d'Uva, O'Donnell, and van Doorslaer (2008) show that better educated individuals are less prone to rate their health highly—leading to an underestimation of health inequality—but Bago d'Uva, van Doorslaer, et al. (2008) find that reporting better health is positively associated with income and negatively associated with education.

Yet another approach to identifying reporting heterogeneity is to use a clinical outcome to measure objective health. Johnston et al. (2009) find that individuals do not accurately report their clinical health—the false-negative report rate for hypertension is 85%—and that a person of low SES is more likely to give a false report, which implies a negative bias by income. Similarly, Cawley and Choi (2017) compare self-reports of health status to clinical measures of health and find that individuals with higher education self-report their health

[6]Shmueli (2003) uses a different type of self-assessment, the health related quality of life score (HRQL), for which respondents are asked to rate their health on a scale of 0 ("death") to 100 ("full health").

more accurately, and Dowd and Zajacova (2010) also find those with higher education more accurately report their health using biomarker data. However, Ziebarth (2010) finds that both SAH and the SF-12 (a generic health measure) are still positively correlated with income while holding grip strength, a clinical health measure, constant.[7]

From the variety of different results in the literature, it is clear that the issue of income-related reporting heterogeneity in SAH is far from settled. We extend the literature of assessing the reliability of SAH by conducting an extensive set of tests to measure the existence and extent of reporting heterogeneity in SAH. We introduce a new test for reporting heterogeneity, cluster analysis, which allows us to easily visualize the relative similarity of SAH and income relative to more objective health measures. We additionally use two tests used in previous research: ordered probit regression-based techniques similar to Lindeboom and van Doorslaer (2004), which allows for the distinction between cut-point and index shifts, and regression-based tests similar to Ziebarth (2010), which allows for us to intuitively characterize the direction of reporting heterogeneity. We briefly summarize these tests below.

2.2.1 Detecting reporting heterogeneity using cluster analysis

First, we introduce a new test of reporting heterogeneity: cluster analysis. The basic notion of clustering is to form groups of variables based on the "closeness" or similarity of their distributions. It has been widely used in psychology, marketing, management and the economic analysis of multidimensional well-being (e.g. Punj & Stewart, 1983; Borgen & Barnett, 1987; Henry, Tolan, & Gorman-Smith, 2005; Hirschberg, Maasoumi, & Slottje, 1991, 2001a, 2001b), but to the authors' knowledge, we are the first to use this method in an effort to detect reporting heterogeneity in SAH. The advantage of cluster analysis for examining reporting heterogeneity is that it offers an intuitively simple visualization of the relationships between key variables while making the relative similarity between those variables clear—something more traditional tests of reporting heterogeneity cannot do.

[7]See Dowd (2012) for a more comprehensive review of this literature.

We focus on the hierarchical agglomerative clustering technique proposed by Ward (1963). Ward's method begins with each variable forming its own group, known as a cluster. The number of clusters is iteratively reduced by combining the two clusters which are most distributionally similar to each other. Once a variable is combined into a cluster with at least one other variable it cannot be re-associated into another cluster. We quantify the similarity of distributions using the error sum of squares (ESS), which is the most relevant measure of distributional similarity for assessing the impact of reporting heterogeneity on a regression-based measure like the concentration index.[8] Thus, at each stage of amalgamation, the next two clusters are combined such that the sum of the ESS measures for all clusters is minimized. The process of combining clusters continues until there are only two clusters left, which are then trivially combined to obtain the final measure of dissimilarity. The result is a complete hierarchical map of the relative similarity between all variables considered.

To use cluster analysis to detect reporting heterogeneity in SAH, we start with a set of variables containing cardinalized SAH, two other self-reported health measures, nine clinical health measures, and an SES measure. Thus our cluster map will begin with 13 clusters, which will be iteratively combined into one cluster. When and how both SAH and the clinical health variables cluster with SES will be particularly revealing. If SAH is a good proxy for clinical health, then it should share a great deal of information with clinical measurements of health and quickly cluster with those variables. If SAH clusters with SES before the majority of the clinical variables, that would be an indication that the distribution of SAH shares more information with the distribution of SES than with clinical health measures. This would be indicative of SES-related reporting heterogeneity.

[8]Other approaches to cluster analysis use similarity measures that consider higher order moments of the distribution (see, for example the entropy-based approach of Hirschberg et al. (1991)). However, since the concentration index is a covariance-based measure, we believe there is minimal value to considering the relationship between the higher order moments of the distributions of these variables.

2.2.2 Detecting reporting heterogeneity using likelihood ratio tests

Lindeboom and van Doorslaer (2004) outline two different types of reporting heterogeneity. An *index shift* is where differential reporting behavior across subgroups leads to a parallel shift of the threshold parameters such that their relative position remains unchanged. Formally, an index shift means that an individual will choose $SAH_i = j$ if and only if $(\alpha_{j-1} + \alpha_m) < H_i^* \leq (\alpha_j + \alpha_m)$, where α_m is a group-specific constant added for individuals in group m to each of the reporting thresholds that apply to the whole population. A *cut-point shift*, on the other hand, occurs when differential reporting behavior affects the thresholds differently across groups such that the relative positions of the reporting thresholds are altered. This is equivalent to having a unique set of cut-points for each group, $\alpha_{j,m}$.

Following the method developed by Lindeboom and van Doorslaer (2004), we estimate ordered probit models, where the dependent variable is ordinal SAH, including controls and an objective health measure as independent variables. We implement various tests of the parallel regression assumption to establish the existence of reporting shifts. More formally, we first estimate a restricted model that imposes the parallel regression assumption for an objective health measure, SES, and other relevant covariates,

$$p_{ij} = p(SAH_i = j) = \Phi\left\{\alpha_{j-1} < \beta_H H_i + \sum_m (\beta_{I,m} I_{i,m}) + X_i'\boldsymbol{\beta}_x + \varepsilon_i < \alpha_j\right\} \tag{3}$$

where $\Phi(\cdot)$ is the cumulative density function of the normal distribution, $I_{i,m}$ is an indicator of individual i belonging to SES group m, and X_i is a vector of demographic and educational characteristics. It is critical to condition on a representative and objective health measure, H_i, to ensure that differences in reporting behavior are not purely due to differences in actual health (we will discuss this in greater detail below). The parallel regression assumption of the ordered probit model imposes that there is a common set of parameter estimates for all SES groups.

Next, we estimate Equation (3) separately for each SES group using the same set of covariates (minus the SES category dummies), and construct likelihood ratio tests for reporting heterogeneity, where $L(\lambda) = -2 \cdot (\lambda^R - \sum_{m=1}^{M} \lambda_m^U)$ is χ^2 distributed with $(k + J - 1)(M - 1)$ degrees of freedom. Here λ^R is the log-likelihood value of the restricted model and λ_m^U is the log-likelihood value of the unrestricted model for SES group m, M defines the number of SES categories, k is the number of parameters in each model, and J is the number of cut-points. The null hypothesis is that there is no reporting heterogeneity, so rejection of the null hypothesis indicates that reporting thresholds change based on an individual's SES.

We can test for cut-point shifts by running Equation (3) separately for each SES group, m and constraining the β parameters to be equal to the full, restricted model and only allowing the cut-point parameters, $\alpha_{j,m}$, to vary by SES group. We can then run a likelihood ratio test between the restricted model and the set of constrained models, $L(\lambda) = -2 \cdot (\lambda^R - \sum_{m=1}^{M} \lambda_m^C)$, which is χ^2 distributed with $J(M-1)$ degrees of freedom. The null hypothesis is that there is a common set of cut-points for all SES groups—that is, $\alpha_{j,m} = \alpha_j$ for all m and j. Rejecting the null would imply reporting thresholds change based on SES and using SAH to estimate the concentration index would result in over- or under-stating health inequality, depending on the nature of bias.

We can also test for index shifts using the likelihood ratio test $L(\lambda) = -2 \cdot (\sum_{m=1}^{M} \lambda_m^C - \sum_{m=1}^{M} \lambda_m^U)$, which is χ^2 distributed with $(M - 1)(K - 1)$ degrees of freedom. Index shifts are typically considered to be less problematic because they can be corrected by including group-specific dummy variables in a regression model. While this may be true for index shifts based on demographic difference or in some other contexts, it is certainly not the case for SES-based index shifts in SAH when estimating a concentration index. Since SES ranks are used to estimate the concentration index, there is no way to control for an SES-based index shift.

2.2.3 Detecting reporting heterogeneity using linear regression models

To gain greater intuition concerning how SES and clinical health are related to subjective health measures and how any reporting bias will ultimately affect the measurement of SES-related health inequality, we build off of the tests for reporting heterogeneity in Ziebarth (2010). We run linear regression models that are fully interacted with measures of SES on the cardinalized measure of SAH and other self-reported health measures,

$$SRH_i = \alpha + \beta_H H_i + \beta_{HI} H_i I_i + \beta_I I_i + \beta_{I2} I_i^2 + \sum_q (\beta_q D_{iq} + \beta_{Iq} D_{iq} I_i) + \sigma_s + \varepsilon_i, \qquad (4)$$

where SHR_i is a measure of self-reported health for individual i, which may be the cardinalized measure of SAH or other self-reported health measure; H_i is a measure of clinical health; I_i is a measure of SES; the D_{iq}'s represent demographic characteristics, and σ_s are survey wave fixed effects.

By specifically controlling for a clinical health measure in these models, any remaining correlations between self-reported health and SES (and its interactions) are driven by a mechanism other than differences in clinical health, which is consistent with SES-based reporting heterogeneity. Thus, statistically significant coefficients on the SES variable and the interactions between SES and the clinical health measure suggest that people of the same clinical health status but different levels of SES will report their health differently. The sign and magnitude of these coefficients will shed light on exactly how any reporting heterogeneity is related to SES and health, which can help us understand how other types of analysis done using subjective health measures might be affected by this bias.[9] A Chow test on the interacted coefficients provides an additional test of whether the determinants of self-reported health differ by SES.

[9] van Doorslaer and Gerdtham (2003) suggest that evidence of income-based reporting heterogeneity is only found when the coefficient on the health/income interaction term is significant. However, both Jürges (2008) and Ziebarth (2010) correctly point that a significant coefficient on the income coefficient is, by itself, evidence of income-based reporting heterogeneity and that the interaction term indicates whether the bias changes over the distribution of health.

2.3 Multivariate Health and Inequality

Despite the potential problems of using SAH to measure health inequality, it has the major advantage of being an index measure. That is, it represents the net effect of multiple dimensions of health in a single measure. Therefore, critiques of the use of SAH have little value unless a more suitable health index is proposed. A specific measure of clinical health may be too narrow to capture the broad scope of an individual's health status. This might explain why there are conflicting results in the literature on reporting heterogeneity, even when conditioning on a clinical measure of health (e.g. Johnston et al., 2009; Ziebarth, 2010). Generic health measures are also problematic because they are inherently based on self reports, which may also suffer from response bias. Indeed, there is evidence of reporting heterogeneity present in several generic measures (Ziebarth, 2010; Shmueli, 2003).

We propose using a variation of allostatic load to measure clinical health objectively and broadly for the purpose of assessing income-related health inequality. First developed by McEwen and Stellar (1993), allostatic load is a measure of cumulative physiological deterioration across a number of biological systems relevant to disease risk. A state in which these physiological systems, and thus the biomarkers, return to normal after experiencing a period of physiological stress is called *allostasis*. Allostatic load includes measures of four main systems: (1) the hypothalamic-pituitary-adrenal axis, a part of the neuroendocrine system that regulates digestion, the immune system, emotions, energy use and storage, and reactions to stress, among other things; (2) the sympathetic nervous system; (3) the cardiovascular system; and (4) the metabolic processes (McEwen & Stellar, 1993; McEwen, 1998; Seeman et al., 2001). In practice, measures of allostatic load include biological measures of obesity, such as body-mass index (BMI) or waist to hip ratio; cardiovascular health, such as blood pressure and albumin; measures of metabolism such as cholesterol levels, triglycerides and glycated hemoglobin; and measures of immunity and inflammation, such as white blood cell count or C-reactive protein levels (Seeman et al., 2001; Geronimus et al., 2006). When calculating a typical measure of allostatic load, individuals receive one point for each clinical

measure for which they fall over a high-risk threshold, and the cumulative number of points is the measure of allostatic load.

We modify this methodology to calculate allostasis, a measure of "good health" as described above, by assigning individuals one point for each measure for which they fall into the normal clinical range. We adjust the weighting of the clinical measures using *multiple correspondence analysis* (MCA). MCA is analogous to principal components analysis for categorical variables, and thus it can be used as a dimension reduction method for creating a single, continuous health index from multiple categorical variables.[10] For a detailed treatment of this method, we refer the reader to Greenacre and Blasius (2006), and for a rigorous discussion of how MCA can be used to create a health index, we refer the reader to Kohn (2012). For our purposes, MCA involves using the nine binary variables for health conditions (described above) to create an $18 \times n$ indicator matrix, $\mathbf{Z}$, (one column for each possible response). The cross-tabulation of this matrix, $\mathbf{Z}'\mathbf{Z}$, creates a symmetric positive definite matrix called the Burt's matrix. The first eigenvalues and corresponding eigenvectors of the Burt's matrix are used to create the weights used for the index. The resulting index can be interpreted as a weighted count of healthy conditions—that is, weighted allostasis measure.[11]

Allostasis is similar in some ways to measures of multidimensional poverty. Alkire and Foster (2011) define a measure of multidimensional poverty in which they aggregate multiple measures by counting the number of dimensions in which their indicators of poverty failed to exceed some predetermined cut-point, making the measure ratio-scale and compatible with concentration index measurement. This technique has been previously adapted for use in health inequality measurement by Makdissi et al. (2013) and Makdissi and Yazbeck (2014). There are three potential problems with this framework, however. First, the set of attributes used to measure health is possibly arbitrarily chosen and sample dependent; second, the cut-points used to define an outcome as "good" or "bad" are often arbitrarily defined; and third,

[10] We thank an anonymous referee for the suggestion to use MCA in calculating allostasis.

[11] The MCA procedure mean centers the data around zero; so the index must be rescaled to achieve this interpretation.

each dimension is given equal weight despite potentially measuring overlapping aspects of health or having different overall importance. Since allostasis is used for clinical assessment, using it as a basis for our health variable choice minimizes the risk of arbitrary inclusion. Moreover, since we are only using clinical measures, we can use clinically relevant cut-points to measure health achievement or failure, minimizing the risk of arbitrary cut-points. Finally, we adjust the weighting of the measures using *multiple correspondence analysis* (MCA).

It is important to note that our use of this measure to both determine the existence of reporting heterogeneity in SAH and as a measure of health in the concentration index is predicated on the notion that allostasis is a "good" measure of general health. If allostasis is missing an important dimension of health that is also correlated with income, then we may incorrectly conclude that SAH suffers from reporting heterogeneity or miss some important income-related health inequality. For example, some may doubt clinical markers' ability to capture mental health. However, SAH has not been definitively shown to be reliable measure of non-physical health either, as most tests of SAH reliability use physical health outcomes, like mortality (e.g. Mossey & Shapiro, 1982; Dowd & Zajacova, 2007; van Doorslaer & Gerdtham, 2003). Secondly, mental health problems often create physical symptoms, and allostatic load and allostasis are designed to capture the physiological consequences of mental stress (McEwen, 1998). We are able to test whether reporting heterogeneity results are affected by mental health, to a certain extent, which we discuss in Section 4.4.

3 Data

We use the 2005-2006, 2007-2008, and 2009-2010 waves of NHANES, a nationally representative cross-sectional survey of individuals in the United States conducted by the Centers for Disease Control and Prevention (CDC) that combines surveys, physical examinations by medical personnel, and laboratory tests. The NHANES surveys include multiple measures of self-reported health, which again may include include questions about physically healthy

and mentally healthy days or SAH. Most directly, NHANES asks individuals, "Would you say your health in general is...", where five represents the worst health and one represents excellent health. We transform this variable into a cardinal number, assuming it follows a log-normal distribution so that higher numbers indicate better health, in the manner of Wagstaff and van Doorslaer (1994). NHANES also asks individuals about the number of physically and mentally unhealthy days in the past month. Specifically, in regards to physical health, NHANES asks individuals, "Thinking about your physical health, which includes physical illness and injury, for how many days during the past 30 days was your physical health not good?" In regards to mental health, NHANES asks individuals, "Thinking about your mental health, which includes stress, depression, and problems with emotions, for how many days during the past 30 days was your mental health not good?" Since these values already represent ratio-scale numbers, we only transform them to represent numbers of physically and mentally healthy days.

The physical examination and laboratory components also collect numerous clinical indicators of health. We use nine clinical measures of health to construct our measure of allostasis. First, we use body mass index (BMI), calculated from physical measurements collected during the examination. Higher BMI has been linked to a host of adverse health conditions and is one of the leading causes of preventable death in the United States (Stommel & Schoenborn, 2010; Mokdad et al., 2004).[12] Second and third, we use resting diastolic and systolic blood pressure measurements collected during the examination. Higher blood pressure is linked to many negative cardiovascular events such as strokes, heart attacks, aneurysms, and heart failures (Collaboration, 2002). The fourth and fifth measures we use are overall cholesterol levels and triglyceride levels. Higher cholesterol and triglyceride levels are linked to many heart conditions (Stamler et al., 2000; Criqui et al., 1993). Sixth, we use individuals' glycated hemoglobin levels, a measure of diabetes which measures plasma glucose for the

[12]Very low BMI, usually defined as below 18.5, is also linked with adverse health conditions such as malnourishment, anorexia, and bulimia. However, only 185 out of over 12,600 individuals fall into this category in our data, so this is unlikely to be a problem.

previous four months (Trivelli et al., 1971). For our seventh measure, we use blood creatinine concentrations to calculate the estimated glomerular filtration rate (eGFR), a measure of kidney function, according to the CKD-EPI equation (Levey et al., 2009). Eighth, we use the white blood cell count, a measure of response to an infectious disease or other cause of tissue damage or inflammation. Our ninth measure is the level of c-reactive protein, a marker of inflammation, heart disease, and other conditions.

We assign individuals one point for each biomarker that falls in the normal clinical range. Our clinical thresholds are taken from a combination of published standards from the Mayo Clinic, the National Library of Medicine, the CDC, and other professional and governmental health organizations. Table 1 contains summary statistics for our allostasis biomarkers, clinical thresholds used, the fraction of our sample that falls in the normal clinical range for each biomarker (and thus considered low risk for the diseases associated with these tests), and the source(s) referenced for the relevant clinical cutoff. The original count measure of allostasis varies from zero (high-risk values for all health measures) to nine (low-risk values for all health measures). This measure is then re-weighted using the MCA procedure described in Section 2.3.[13]

We also use the survey portion of NHANES to collect other individual-level data. NHANES collects multiple measures of income, including the family income-to-poverty ratio. NHANES reports this ratio based on poverty guidelines from the Department of Health and Human Services specific to family size and each state.[14] Finally, NHANES collects numerous demographic characteristics, and we use these to control for a quadratic function of age and create indicator variables for individuals' gender, race, Hispanic origin, and education. We exclude

[13]There is some concern about how representative a single clinical assessment is for measuring general health. For example, blood pressure can fluctuate throughout the day or an individual may have had an acute infection when the labs were drawn. This risk of measurement error is somewhat mitigated by dichotomizing these measures, which implies that only individuals whose measures fluctuate around the unhealthy threshold will be measured with error. The aggregation of these measures into a single index further minimizes the effects of this kind of measurement error. To the extent that measurement error remains, it should be considered classical measurement error.

[14]This measure is ordinally equivalent to measure of actual income that has been adjusted for family size and regional cost of living differences. However, to reduce disclosure concerns, NHANES top-codes this value to 5. We discuss the robustness of our results to other income measures in the next section.

any individuals who do not report self-reported health information, are pregnant, or are in nursing homes. Table 2 contains summary statistics for all relevant variables not described in Table 1.

Figures 1, 2, and 3 show the distributions of SAH, clinical measures, and allostasis. The distribution of SAH appears to be roughly normal, although with slightly more density in the right tail. However, the distributions of physically and mentally healthy days appear much different, with the vast majority of respondents indicating the maximum number of healthy days. Most of the clinical measures show skewness in the direction of the unhealthy end of the distribution, with a few individuals having extreme levels of each clinical measure, but there is a lot of variation in the shapes of these density functions. When aggregated into the allostasis measure, the unhealthy skewness remains and appears far more pronounced than that of SAH. On the surface, this suggests there is some disconnect between SAH and clinical health. However, it is difficult to determine whether there is a strong relationship between the different health measures when only examining their univariate distributions.

4 Reporting Heterogeneity Results

4.1 Cluster Analysis

We first examine the relationship between income, self-reported health, and clinical measures of health using cluster analysis. When testing for reporting heterogeneity, we are particularly interested in how quickly cardinalized SAH clusters with clinical health measures relative to the income-to-poverty ratio. Figure 4 displays results from our cluster analysis. SAH clusters almost immediately with the income-to-poverty ratio, suggesting a great deal of similarity between these two variables. Physically and mentally healthy days also cluster very quickly and then cluster with SAH and the income-to-poverty ratio. The clinical measures of health form into clusters separate from the self-reported measures of health and the income-to-poverty ratio. Thus, the cluster analysis suggests that self-reported measures are much more

closely related to income than they are to clinical measures of health. Put differently, the hierarchical structure of the data shows that self-assessed health measures are not good index measures of clinical health because they are much more closely related to another, relevant variable: income. Thus, using something like SAH as a proxy for clinical health could introduce measurement error that correlates strongly with SES. We test the robustness of these results to using household income grouped into 11 categories and find nearly identical results.[15] Figures for these cluster results are found in the appendix.

4.2 Lagrange Multiplier Tests

We next test for cut-point and index shifts in the ordinal SAH variable according to the methods described in Section 2.2.2. Table 3 shows likelihood ratio tests for reporting heterogeneity by the SES or demographic category in the column title using allostasis as our health variable. We examine the same 11 categories of family income noted in footnote 15, four categories of education (less than high school, high school, some college, and bachelors degree or above), three subgroups of race/ethnicity (white, black, Hispanic), two subgroups for sex, and two subgroups for age (above and below age 45).[16] We also control for marital status and survey wave fixed effects. We report the likelihood value for the restricted model (λ^R), the sum of the likelihood values from the unrestricted models, ($\sum_m \lambda_m^U$), the sum of the likelihood values from the constrained models, ($\sum_m \lambda_m^C$), and the likelihood ratio, degrees of freedom, and p-value from each of the three tests for reporting heterogeneity.[17]

Similar to Lindeboom and van Doorslaer (2004), we find strong evidence of reporting heterogeneity by sex and age and that the differential reporting behavior is due to cut-point shifts, but we additionally find evidence of cut-point shifts by race/ethnicity. This suggests

[15] Income categories are: \$0-\$5K, \$5K-\$10K, \$10K-\$15K, \$15K-\$20K, \$20K-\$25K, \$25K-\$35K, \$35K-\$45K, \$45K-\$55K, \$55K-\$65K, \$65K-\$75K, and over \$75K. We use the middle dollar amount suggested by each category. For example an individual in the first category is assigned an income of \$2,500, in the second category, \$7,500. Individuals in the highest category are assigned an income of \$100,000.

[16] All models other than "Income" instead control for income quartiles. All models other than "Age" control for a continuous measure of age using a quadratic function. Other than these differences, the restricted model is the same for each column.

[17] Tests for index shifts are of little importance if the second test suggests the presence of cut-point shifts.

that not only are SAH values not comparable between these groups, but also that these response differences cannot be corrected by using index functions. Thus, when using SAH, researchers should run separate analyses for each subpopulation when possible.

In contrast to Lindeboom and van Doorslaer (2004), we find clear evidence of reporting heterogeneity by both income and education and that these response differences are due to cut-point shifts. We cannot definitively determine whether this difference in result is due to differences in sample or due to using different measures of objective health—although we suspect that it is the latter. Lindeboom and van Doorslaer (2004) use a generic health measure, the HUI3, which is based on subjective responses to a series of arbitrarily scaled questions. This raises concerns that the HUI3 is subject to the same types of reporting heterogeneity as SAH, and previous research has shown that generic health measures can suffer from income-related response bias (e.g. Shmueli, 2003). However, we cannot rule out the possibility that the HUI3 is simply a more comprehensive measure of health than allostasis. We are able to test some elements of this hypothesis in Section 4.4.

Assuming allostasis sufficiently captures the relevant dimensions of health, our results suggest that there is a significant amount of variation in SAH responses due to differences in income and education. Moreover, this variation is independent of changes in health, and estimates of SES-based health inequality using this measure, or other self- reported health measures, are likely incorrect. This is particularly problematic because there is no valid method in this context to deal with reporting heterogeneity by running separate analyses by group.

We next explore whether income-based reporting heterogeneity holds for various subpopulations. Table 4 shows group-specific results testing for differential response behavior by income quartiles. The model "Full" repeats the tests for income-based reporting heterogeneity for the full sample using four income categories (instead of 11), and regression results from all other columns come from using only the subsample in the column title. Evidence of income-based cut-point shifts continues to hold for individuals that are white, for both men

and women, and for individuals of all ages. The gender result is contrary to the findings of Schneider et al. (2012) who only find income-based reporting heterogeneity for men. The only group for which there is not clear evidence of income-based cut-point shifts are black individuals.[18]

4.3 Characterizing SES-Based Reporting Heterogeneity

We use the linear regression approach outlined in Equation (4) to better understand how subjective health, SES, and objective health are related. We use the cardinalized measure of SAH and two other self-reported health measures, the reported number of physically healthy days last month and the reported number of mentally healthy days last month, as dependent variables. A statistically significant coefficient on income, income squared, or any characteristic interacted with income suggests that income is correlated with health self-assessments in a manner that does not operate through differences in clinical health, which is indicative of income-related reporting heterogeneity. It is important to note that this cross-sectional analysis does not allow us to make any strong claims about causation or the dynamics of how an individual reacts to change in his or her income, health, education, etc.

Table 5 shows results from these regressions. A p-value of the Chow test that all income-interacted coefficients are jointly equal to zero for each model is found at the bottom of each column. The null hypothesis of this test is strongly rejected in all three models, indicating that the marginal effects of the determinants of subjective health are not drawn from the same distribution for all income levels. This is yet further evidence of income-based reporting heterogeneity—not just in SAH, but in other subjective health measures as well.

In the first column, allostasis is positively and significantly associated with SAH, as expected. At the mean level of income, a one unit increase in allostasis is associated with a 0.217 increase in an individual's SAH, which is a nearly unit elastic response at the mean

[18]We repeated these tests using the individual biomarkers instead of allostasis and found very similar results. These tables are available in the Appendix.

($\epsilon_H = 0.97$).[19] Income is directly associated with an individuals' reporting of SAH, given the positive and statistically significant coefficient on income, and indirectly associated through a variety of other channels. Income interacts positively with allostasis, suggesting that as income increases above the mean, the marginal effect of clinical health on SAH increases. The association between income and SAH conditional on clinical health is strong: the marginal effect of a one unit change in allostasis approximately triples from one end of the income distribution to the other.[20] This effect could be viewed in two ways: either lower income individuals do not recognize the importance of marginal changes to their clinical health, or higher income individuals overreact to changes in their clinical health (or some combination of both). Of course, since we do not know what the "appropriate" marginal response is, it is impossible to know for sure.

Consistent with the results in Section 4.2, we see evidence of positive reporting heterogeneity from higher levels of education, and the insignificant interaction terms suggest that this effect is independent of the reporting heterogeneity we detect for income. We also see evidence of reporting heterogeneity related to age, with older individuals tending to assess their health better than their younger, equivalently healthy counterparts, and gender, with females assessing their health lower than equivalently healthy men.

We do a large set of robustness checks to these results. These include using individual biomarkers in place of the allostasis metric, using a version of allostasis which simply counts the dimensions of good health (with equal weight), running regressions without NHANES samples weights, using alternate versions of SAH, running regressions for various demographic and socioeconomic subgroups, and running a generalized ordered probit regression. Finally, since the US has no universal health care coverage, which could confound the relationship between income and health, we run regressions excluding individuals without regular access to medical care and an additional set of results excluding those individuals

[19] The mean value of allostasis is 7.507 and the mean value of cardinalized SAH is 1.675. Thus, $\partial y/\partial x \times \bar{x}/\bar{y} = 0.217 \times 7.507/1.675 = 0.97$.

[20] $\partial SAH/\partial H = 0.08$ at an income-to-poverty ratio of 0.5 and $\partial SAH/\partial H = 0.29$ at an income-to-poverty ratio of 4.5. Marginal effects for allostasis are statistically significant for all income values.

taking regular prescription medications. Tables of these results along with brief discussions are available in the appendix. All results are consistent with our main analysis.

4.4 Discussion of Unobservable Health Dimensions in SAH

There is some concern that SAH is capturing a dimension of health that allostasis (or any of the single clinical health measures) does not. Our tests of reporting heterogeneity using regressions (both ordered response and linear) implicitly assume that clinical measures are "good" proxies for health in general. If an important dimension of health captured by SAH, but not by our clinical indicators, is significantly correlated with income, then our findings of income-based reporting heterogeneity may be unfounded (see Lindeboom & van Doorslaer, 2004; Ziebarth, 2010). One may worry that mental health, for example, is unlikely to be directly captured by clinical assessments. Given that the coefficient for the income-to-poverty ratio in the mentally healthy days regression in Table 5 is positive, this could be plausible.

We believe this is unlikely for a number of reasons. As previously stated, mental health problems often create physical symptoms, and allostatic load and allostasis are designed to capture the physiological consequences of mental stress (McEwen, 1998). Second, the income-related heterogeneity found by Shmueli (2003) was strongest for the mental health component of the SF-36; so the strong relationship between mental health and income is not surprising, and is possibly also affected by reporting heterogeneity. Third, the coefficients on the mentally and physically healthy days regressions are not consistent with they hypothesis that non-clinical health is driving the remaining correlations with income. The pattern of results looks remarkably different from the first column. To further alleviate concerns regarding dimensions of health not captured by clinical markers, we run an additional SAH regression that controls for both mentally and physically healthy days, shown in the last column of Table 5. The coefficients on income and the interaction between income and allostasis are very similar to the results in column one, which suggests that the reporting heterogeneity detected in our main results is not due to a lack of accounting for mental health.

We cannot completely rule out the possibility that SAH captures some income-correlated dimension of health that allostasis does not. However, the results discussed above suggest that such a dimension of health is not related to mental health or physical health.

5 Health Inequality Results

In Section 4, we find significant evidence of positive income-related reporting heterogeneity using a variety of tests. If these results hold generally, then using SAH as a health measure could cause researchers to overstate the degree of income-based health inequality. Furthermore, we find that changes in clinical health (whether positive or negative) have a greater impact on health self-assessments for higher-income people than for lower-income people. Thus, a marginal increase to the true health of the entire population would cause the variance of the SAH distribution to increase. This type of reporting heterogeneity could not only cause health inequality measurements to be overstated for healthy populations, but it could also tend to make income-based inequality appear to be a greater portion of overall inequality.

We examine the effects of the choice of health measure on the relationship between health inequality and SES using the concentration index framework described in Equation (2). Table 6 shows concentration index results using three measures of self-reported health and allostasis. We rank individuals' SES using the family income-to-poverty ratio, and the coefficient on this measure is the concentration index. The three measures of self-reported health (SAH, physically healthy days, and mentally healthy days) all suggest a strong relationship between health inequality and SES, as evidenced by the positive coefficients, statistically significant at the 1% level. Our concentration index measure of 0.058 for SAH is about half of that of van Doorslaer et al. (1997), who also use US data and a similar methodology. One difference in methodologies is that our results are adjusted for education. Comparing our unweighted concentration index results for SAH that omit education controls yields a very

comparable measure of health inequality to van Doorslaer et al. (1997). These results are shown in Table 7. Our multidimensional measure of clinical health, allostasis, suggests a much weaker relationship between health inequality and SES, with a concentration index of 0.005. Although the concentration index is also statistically significant for allostasis at the 1% level, it is roughly one-tenth the size of the concentration index from the SAH regression. Our results are comparable to other papers that estimate the concentration index using an objective health measure (e.g. Ziebarth, 2010; Makdissi et al., 2013; Makdissi & Yazbeck, 2014). These results are consistent with our hypothesis that income-related reporting heterogeneity in SAH could cause income-related health inequality to be overstated.

It is also worth noting that physically and mentally healthy days show a concentration index about one-third of the size of the one for SAH, which is consistent with the weaker observed relationship between these self-reported measures and income in Section 4.3. We speculate that this may be due to the fact that mentally and physically healthy days are measured on an objective scale, despite being self-reported. We again note that we cannot rule out that SAH is capturing a dimension of health that mentally healthy days, physically healthy days, or allostasis misses and that this unobserved dimension of health is driving the higher concentration index estimates we see for SAH. Even in that case, we can still say that calculating concentration indices with SAH will cause researchers to overstate the degree of income-related inequality in physical and mental health, which are often the health dimensions of primary policy interest.

Table 7 shows a number of different robustness checks to our concentration index results. We use two other income measures, family income and household income, in addition to our preferred income-to-poverty ratio measure. We compare each of these concentration indices to ones that are calculated without using NHANES sample weights. Additionally, we show each of these possible iterations both controlling for education and without controlling for education. Finally, we include a concentration index calculated using a count measure of allostasis, which is more common in the allostatic load literature. There are two important

patterns that emerge: First, the concentration index calculated using SAH is more sensitive to these changes than the other two self-reported measures, and all self-reported measures appear more sensitive to these changes than allostasis. Second, despite some sensitivity, the concentration index for allostatic load remains consistently less than one-tenth of the concentration index for SAH and about one-fifth of the concentration index for physically and mentally healthy days.

We further test the robustness of our results in a number of ways. We estimate separate concentration index results for each clinical measure of health, estimate results using the transformation suggested by Erreygers (2009a) instead of the standard concentration index, and estimate results using various alternative SAH measures and cardinalizations. Finally, in order to ensure that our results are not affected by differential access to the health care system, we replicate the results in Table 6 once excluding individuals without regular access to medical care and again excluding individuals regularly taking prescription medications. All of these results are found in the Appendix and are very similar to our main results.

6 Conclusion

The main objective of this paper is to understand the impact of reporting heterogeneity in SAH on the estimation of health inequality. We estimate a comprehensive set of tests of SES-based reporting heterogeneity, including hierarchical agglomerative clustering, specific tests of parallel regression assumptions using ordered probit models, and fully interacted linear regression models. Our results suggest that there is significant, positive income-based reporting heterogeneity and that this heterogeneity interacts positively with clinical health measures. Individuals not only tend to assess their health more positively as income increases, but higher income individuals react more strongly to a change in clinical health measures. We not only find significant evidence that SAH reporting is affected by income, but also that these effects hold for nearly every demographic subgroup of the data. Fur-

thermore, we also find evidence of reporting heterogeneity by sex, race/ethnicity, and age. These results suggest that using SAH to measure the health concentration index will lead to an over-statement of health inequality, relative to both other self-reported measures of health and especially clinical measures of health. We confirm this by comparing concentration indices calculated using self-reported measures to those calculated using clinical health measures. Our concentration index results show that income-based reporting heterogeneity has an extremely large effect on the estimates of the concentration index.

Parallel to and in support of the analysis described above, we propose the use of allostasis as an objective measure of clinical health in the context of measuring income-related health inequality. Based on the concept of allostatic load, a commonly-used measure in medicine of the cumulative physiological stresses on an individual, allostasis is a multidimensional, ratio scale measure of health. Although previous papers have suggested multidimensional measures of health (e.g. Makdissi & Yazbeck, 2014), allostasis suggests a simple framework for using clinical foundations to guide the choice of both clinical measures and also how healthy cut-points should be chosen. In addition to being an objective and comprehensive measure of health, allostasis is also a ratio scale variable, which makes it appropriate for evaluating health inequality with the concentration index.

There are a number of caveats worth mentioning. The most important thing to note is that our results are estimated using only one survey, NHANES. We cannot rule out the possibility that both the magnitude and the nature of reporting bias may be different in other surveys, particularly those done outside of the United States. Additionally, it is not clear from our analysis to what extent these same types of reporting bias are present in other self-reported health measures. While we find suggestive evidence of income-based reporting bias in both physically and mentally days, the pattern of bias was not as strong as the SAH measure and did not appear to have the same impact on health inequality.

Future research should focus on examining the relationship between alternative measures of self-reported health and clinical markers of health in other surveys to determine the extent

to which these results are generalizable. There are a growing number of surveys, both in the United States and in other countries, which contain clinical marker of health. For example, the Health and Retirement Study, a representative panel of about 20,000 aging adults in the United States, collected a variety of clinical health measures during two survey waves (Health and Retirement Survey, 2017). Also in the United States, the Midlife in the United States survey, another longitudinal survey, also collected a number of biomarkers (MIDAS, 2013). Internationally, the Health Survey of England, a repeated cross-section dataset, contains clinical markers of health since the 2002 wave (Carrieri & Jones, 2017), and surveys with clinical markers of health are also available in many other countries.[21] Constructing a clinical index of health may also be beneficial for any research into the complex relationship between health, income, and education.

[21]The Chicago Core for Biomarkers in Population-Based Aging Research at the University of Chicago has assembled a list of studies collecting biomarkers `http://biomarkers.bsd.uchicago.edu/studiescollectingbiomarkers.htm` (last accessed July 27, 2018). Additionally, the Biomarker Network at the University of Southern California has another list of surveys containing biomarkers. `http://gero.usc.edu/CBPH/network/resources/studies/index.html` (last accessed June 19, 2017).

References

Alkire, S., & Foster, J. (2011). Counting and multidimensional poverty measurement. *Journal of Public Economics*, *95*(7-8), 476-487.

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.

Bago d'Uva, T., van Doorslaer, E., Lindeboom, M., & O'Donnell, O. (2008). Does reporting heterogeneity bias the measurement of health disparities? *Health Economics*, *17*(3), 351-375.

Black, N., Johnston, D. W., Shields, M. A., & Suziedelyte, A. (2017). Who provides inconsistent reports of their health status? the importance of age, cognitive ability and socioeconomic status. *Social Science & Medicine*, *191*, 9 - 18.

Bleichrodt, H., Rohde, K. I., & Ourti, T. V. (2012). An experimental test of the concentration index. *Journal of Health Economics*, *31*(1), 86 - 98.

Borgen, F. H., & Barnett, D. C. (1987). Applying cluster analysis in counseling psychology research. *Journal of Counseling Psychology*, *34*(4), 456.

Bound, J. (1991). Self-reported versus objective measures of health in retirement models. *Journal of Human Resources*, *26*(1), 106–138.

Carrieri, V., & Jones, A. M. (2017). The income-health relationship 'beyond the mean': New evidence from biomarkers. *Health Economics*, *26*(7), 937–956.

Cawley, J., & Choi, A. (2017). Health disparities across education: The role of differential reporting error. *Health Economics,*, *forthcoming*, 1–29.

Clarke, P. M., Gerdtham, U.-G., Johannesson, M., Bingefors, K., & Smith, L. (2002). On the measurement of relative and absolute income-related health inequality. *Social Science & Medicine*, *55*(11), 1923 - 1928.

Collaboration, P. S. (2002). Age-specific relevance of usual blood pressure to vascular mortality: a meta-analysis of individual data for one million adults in 61 prospective studies. *Lancet*, *360*(9349), 1903–1913.

Contoyannis, P., Jones, A. M., & Rice, N. (2004). The dynamics of health in the british household panel survey. *Journal of Applied Econometrics*, *19*(4), 473-503.

Criqui, M. H., Heiss, G., Cohn, R., Cowan, L. D., Suchindran, C. M., Bangdiwala, S., ... Davis, C. (1993). Plasma triglyceride level and mortality from coronary heart disease. *New England Journal of Medicine*, *328*(17), 1220–1225.

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

Deaton, A. S., & Paxson, C. H. (1998). Aging and inequality in income and health. *American Economic Review*, *88*, 248–253.

Dolores Montoya Diaz, M. (2002). Socio-economic health inequalities in Brazil: gender and age effects. *Health Economics*, *11*(2), 141–154.

Dowd, J. B. (2012). Whiners, deniers, and self-rated health: What are the implications for measuring health inequalities? a commentary on layes, et al. *Social Science & Medicine*, *75*(1), 10 - 13.

Dowd, J. B., & Zajacova, A. (2007). Does the predictive power of self-rated health for subsequent mortality risk vary by socioeconomic status in the US? *International Journal*

of Epidemiology, *36*(6), 1214-1221.

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

Erreygers, G. (2009a). Correcting the concentration index. *Journal of Health Economics*, *28*(2), 504–515.

Erreygers, G. (2009b). Correcting the concentration index: A reply to Wagstaff. *Journal of Health Economics*, *28*(2), 521 - 524.

Erreygers, G., & Van Ourti, T. (2011a). Measuring socioeconomic inequality in health, health care and health financing by means of rank-dependent indices: A recipe for good practice. *Journal of Health Economics*, *30*(4), 685 - 694.

Erreygers, G., & Van Ourti, T. (2011b). Putting the cart before the horse. a comment on Wagstaff on inequality measurement in the presence of binary variables. *Health Economics*, *20*(10), 1161–1165.

Fernández-Val, I., Savchenko, Y., & Vella, F. (2017). Evaluating the role of income, state dependence and individual specific heterogeneity in the determination of subjective health assessments. *Economics & Human Biology*, *25*, 85 - 98.

Fleurbaey, M., & Schokkaert, E. (2009). Unfair inequalities in health and health care. *Journal of Health Economics*, *28*(1), 73 - 90.

Frijters, P., & Ulker, A. (2008). Robustness in health research: Do differences in health measures, techniques, and time frame matter? *Journal of Health Economics*, *27*(6), 1626 - 1644.

Geronimus, A. T., Hicken, M., Keene, D., & Bound, J. (2006). Weathering and age patterns of allostatic load scores among blacks and whites in the United States. *American Journal of Public Health*, *96*(5), 826-833.

Greenacre, M., & Blasius, J. (2006). *Multiple correspondence analysis and related methods*. Chapman and Hall/CRC.

Greene, W. H., Harris, M. N., & Hollingsworth, B. (2015). Inflated responses in measures of self-assessed health. *American Journal of Health Economics*, *1*, 461–493.

Groot, W. (2000). Adaptation and scale of reference bias in self-assessments of quality of life. *Journal of Health Economics*, *19*(3), 403 - 420.

Harris, M. C., & Kohn, J. L. (2017). Reference health and the demand for medical care. *The Economic Journal*, *forthcoming*.

Health and Retirement Survey. (2017). *Biomarker data 2006, 2008, 2010, 2012.* (Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant number NIA U01AG009740). Ann Arbor, MI)

Henry, D. B., Tolan, P. H., & Gorman-Smith, D. (2005). Cluster analysis in family psychology research. *Journal of Family Psychology*, *19*(1), 121.

Hirschberg, J. G., Maasoumi, E., & Slottje, D. J. (1991). Cluster analysis for measuring welfare and quality of life across countries. *Journal of Econometrics*, *50*(1), 131–150.

Hirschberg, J. G., Maasoumi, E., & Slottje, D. J. (2001a). Clusters of attributes and well-being in the usa. *Journal of Applied Econometrics*, *16*(3), 445–460.

Hirschberg, J. G., Maasoumi, E., & Slottje, D. J. (2001b). The environment and the quality of life in the United States over time. *Environmental Modelling & Software*, *16*(6), 525.

Humphries, K. H., & van Doorslaer, E. (2000). Income-related health inequality in Canada. *Social science & medicine*, *50*(5), 663–671.

Jenkins, S. (1988). Calculating income distribution indicies from micro-data. *National Tax Journal*, *41*(1), 139–42.

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.

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

Kakwani, N. (1977). Measurement of tax progressivity: an international comparison. *The Economic Journal*, *87*, 71–80.

Kakwani, N. (1980). *Income inequality and poverty: Methods of estimation and policy applications*. Oxford University Press, New York.

Kakwani, N., Wagstaff, A., & van Doorslaer, E. (1997). Socioeconomic inequalities in health: measurement, computation, and statistical inference. *Journal of Econometrics*, *77*(1), 87–103.

Karlamangla, A. S., Singer, B. H., McEwen, B. S., Rowe, J. W., & Seeman, T. E. (2002). Allostatic load as a predictor of functional decline: MacArthur studies of successful aging. *Journal of Clinical Epidemiology*, *55*(7), 696–710.

Kennedy, B. P., Kawachi, I., Glass, R., & Prothrow-Stith, D. (1998). Income distribution, socioeconomic status, and self rated health in the United States: multilevel analysis. *BMJ*, *317*(7163), 917–921.

Kjellsson, G., & Gerdtham, U.-G. (2013). On correcting the concentration index for binary variables. *Journal of Health Economics*, *32*(3), 659 - 670.

Kohn, J. L. (2012, Feb 01). What is health? a multiple correspondence health index. *Eastern Economic Journal*, *38*(2), 223–250.

Lambert, P., & Zheng, B. (2011). On the consistent measurement of attainment and shortfall inequality. *Journal of Health Economics*, *30*(1), 214 - 219.

Layes, A., Asada, Y., & Kephart, G. (2012). Whiners and deniers—what does self-rated health measure? *Social Science & Medicine*, *75*(1), 1 - 9.

Levey, A. S., Stevens, L. A., Schmid, C. H., Zhang, Y., Castro, A. F., III, Feldman, H. I., ... Coresh, J. (2009). A new equation to estimate glomerular filtration rate. *Annals of Internal Medicine*, *150*(9), 604-612.

Lindeboom, M., & van Doorslaer, E. (2004). Cut-point shift and index shift in self-reported health. *Journal of Health Economics*, *23*(6), 1083 - 1099.

Makdissi, P., Sylla, D., & Yazbeck, M. (2013). Decomposing health achievement and socioeconomic health inequalities in presence of multiple categorical information. *Economic Modelling*, *35*, 964 - 968.

Makdissi, P., & Yazbeck, M. (2014). Measuring socioeconomic health inequalities in presence of multiple categorical information. *Journal of Health Economics*, *34*, 84 - 95.

Makdissi, P., & Yazbeck, M. (2017). Robust rankings of socioeconomic health inequality using a categorical variable. *Health Economics*, *forthcoming*.

McEwen, B. S. (1998). Stress, adaptation, and disease: Allostasis and allostatic load. *Annals of the New York Academy of Sciences*, *840*(1), 33–44.

McEwen, B. S., & Stellar, E. (1993). Stress and the individual: mechanisms leading to disease. *Archives of Internal Medicine*, *153*(18), 2093–2101.

MIDAS. (2013). *Ryff, Carol D., Teresa Seeman and Maxine Weinstein. Midlife in the United States (MIDUS 2): Biomarker Project, 2004-2009.* (ICPSR29282-v6. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2013-12-20. https://doi.org/10.3886/ICPSR29282.v6)

Mokdad, A. H., Marks, J. S., Stroup, D. F., & Gerberding, J. L. (2004). Actual causes of death in the United States, 2000. *JAMA*, *291*(10), 1238–1245.

Mossey, J. M., & Shapiro, E. (1982). Self-rated health: a predictor of mortality among the elderly. *American Journal of Public Health*, *72*(8), 800-808.

NIDDK. (2016). *Estimating glomerular filtration rate. National Kidney Disease Education Program.* National Institute of Diabetes and Digestive and Kidney Diseases. U.S. Department of Health and Human Services `https://www.niddk.nih.gov/health-information/communication-programs/nkdep/laboratory-evaluation/glomerular-filtration-rate/estimating`.

O'Donnell, O. A., Van Doorslaer, E., Wagstaff, A., & Lindelow, M. (2008). *Analyzing health equity using household survey data: a guide to techniques and their implementation.* World Bank Publications.

Punj, G., & Stewart, D. W. (1983). Cluster analysis in marketing research: Review and suggestions for application. *Journal of Marketing Research*, *20*, 134–148.

Rosa Dias, P. (2009). Inequality of opportunity in health: evidence from a UK cohort study. *Health Economics*, *18*(9), 1057–1074.

Schneider, U., Pfarr, C., Schneider, B. S., & Ulrich, V. (2012). I feel good! Gender differences and reporting heterogeneity in self-assessed health. *The European Journal of Health Economics*, *13*(3), 251–265.

Seeman, T. E., McEwen, B. S., Rowe, J. W., & Singer, B. H. (2001). Allostatic load as a marker of cumulative biological risk: MacArthur studies of successful aging. *Proceedings of the National Academy of Sciences*, *98*(8), 4770–4775.

Shmueli, A. (2003). Socio-economic and demographic variation in health and in its measures: the issue of reporting heterogeneity. *Social Science & Medicine*, *57*(1), 125 - 134.

Stamler, J., Daviglus, M. L., Garside, D. B., Dyer, A. R., Greenland, P., & Neaton, J. D. (2000). Relationship of baseline serum cholesterol levels in 3 large cohorts of younger men to long-term coronary, cardiovascular, and all-cause mortality and to longevity. *JAMA*, *284*(3), 311–318.

Stommel, M., & Schoenborn, C. A. (2010). Variations in BMI and prevalence of health risks in diverse racial and ethnic populations. *Obesity*, *18*(9), 1821–1826.

Trannoy, A., Tubeuf, S., Jusot, F., & Devaux, M. (2010). Inequality of opportunities in health in France: a first pass. *Health Economics*, *19*(8), 921–938.

Trivelli, L. A., Ranney, H. M., & Lai, H.-T. (1971). Hemoglobin components in patients with diabetes mellitus. *New England Journal of Medicine*, *284*(7), 353–357.

van Doorslaer, E., & Gerdtham, U.-G. (2003). Does inequality in self-assessed health predict inequality in survival by income? evidence from Swedish data. *Social Science & Medicine*, *57*(9), 1621 - 1629.

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.

van Doorslaer, E., Wagstaff, A., Bleichrodt, H., Calonge, S., Gerdtham, U.-G., Gerfin, M., ... Winkelhake, O. (1997). Income-related inequalities in health: some international comparisons. *Journal of Health Economics*, *16*(1), 93 - 112.

Wagstaff, A. (2002). Inequality aversion, health inequalities and health achievement. *Journal of Health Economics*, *21*(4), 627–641.

Wagstaff, A. (2009). Correcting the concentration index: A comment. *Journal of Health Economics*, *28*(2), 516 - 520.

Wagstaff, A. (2011). The concentration index of a binary outcome revisited. *Health Economics*, *20*(10), 1155–1160.

Wagstaff, A., Paci, P., & van Doorslaer, E. (1991). On the measurement of inequalities in health. *Social Science & Medicine*, *33*(5), 545–557.

Wagstaff, A., & van Doorslaer, E. (1994). Measuring inequalities in health in the presence of multiple-category morbidity indicators. *Health Economics*, *3*(4), 281–291.

Wagstaff, A., van Doorslaer, E., & Watanabe, N. (2003). On decomposing the causes of health sector inequalities with an application to malnutrition inequalities in Vietnam. *Journal of Econometrics*, *112*(1), 207–223.

Ward, J., Joe H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, *58*(301), 236–244.

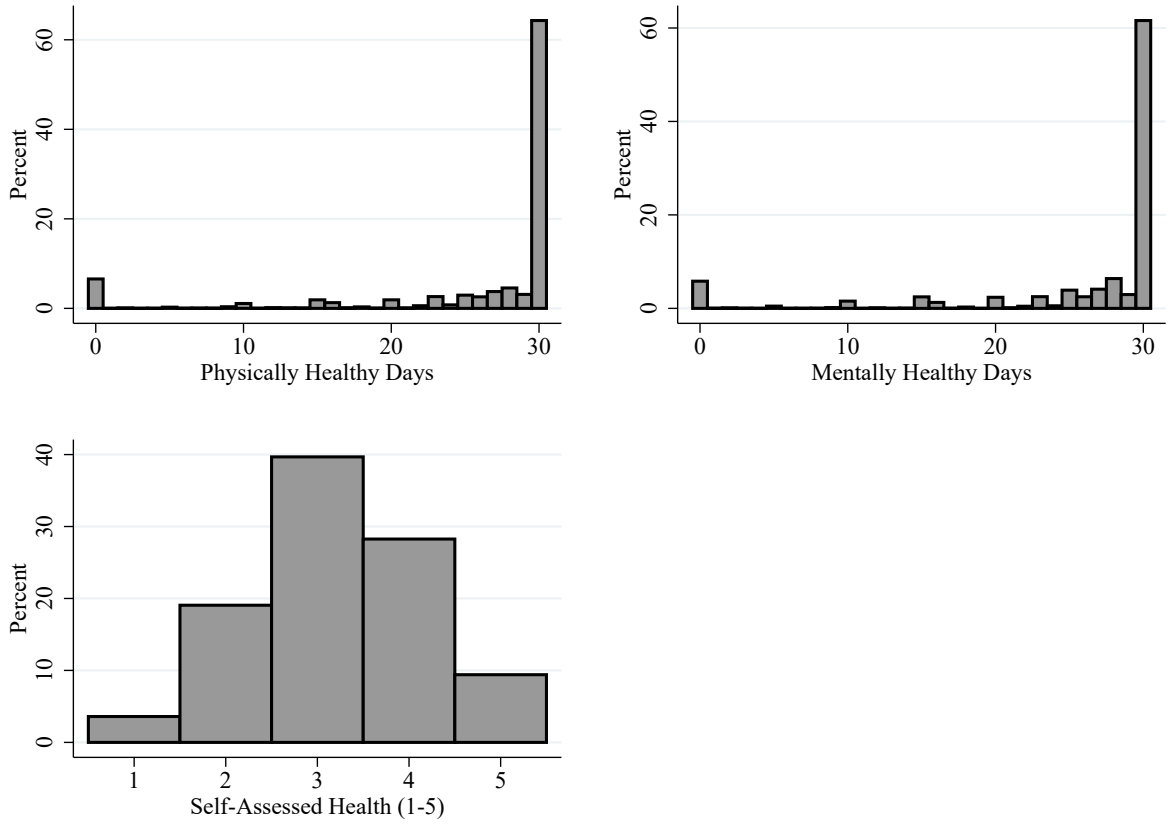
Whelton, P. K., Carey, R. M., Aronow, W. S., Casey, D. E., Collins, K. J., Himmelfarb, C. D., ... Wright, J. T. (2018). 2017 guideline for the prevention, detection, evaluation, and management of high blood pressure in adults: A report of the american college of cardiology/american heart association task force on clinical practice guidelines. *Hypertension*, *71*(6), e13-e115.

Zajacova, A., & Dowd, J. B. (2011). Reliability of self-rated health in US adults. *American Journal of Epidemiology*, *174*(8), 977-983.

Zhang, Q., & Wang, Y. (2004). Socioeconomic inequality of obesity in the United States: do gender, age, and ethnicity matter? *Social Science & Medicine*, *58*(6), 1171–1180.

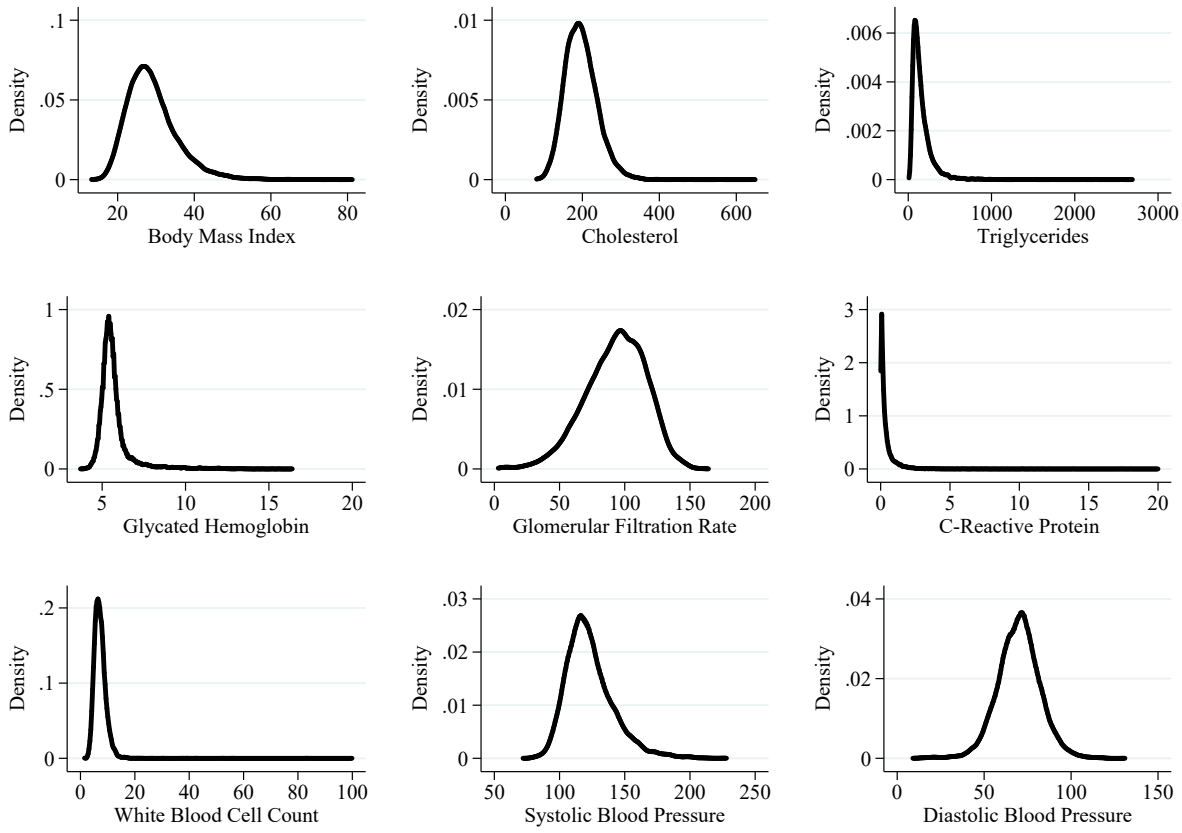
Ziebarth, N. (2010). Measurement of health, health inequality, and reporting heterogeneity. *Social Science & Medicine*, *71*(1), 116–124.

Figure 1: Relative Frequency Chart of Self-Reported Health Levels



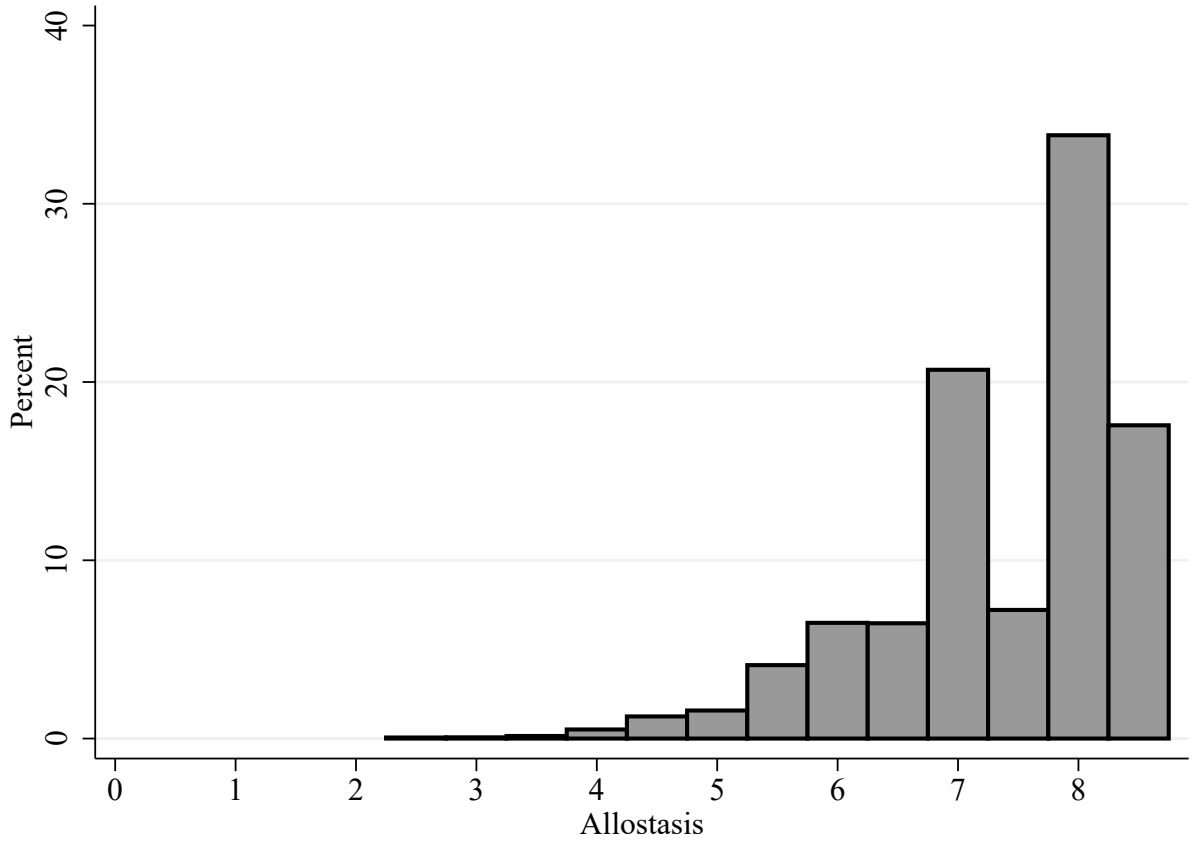
Notes: Data from the 2005-2006, 2007-2008 and 2009-2010 waves of NHANES including adults over age 20. Higher values for self-assessed health correspond to better self-assessed health.

Figure 2: Densities of Health Biomarkers



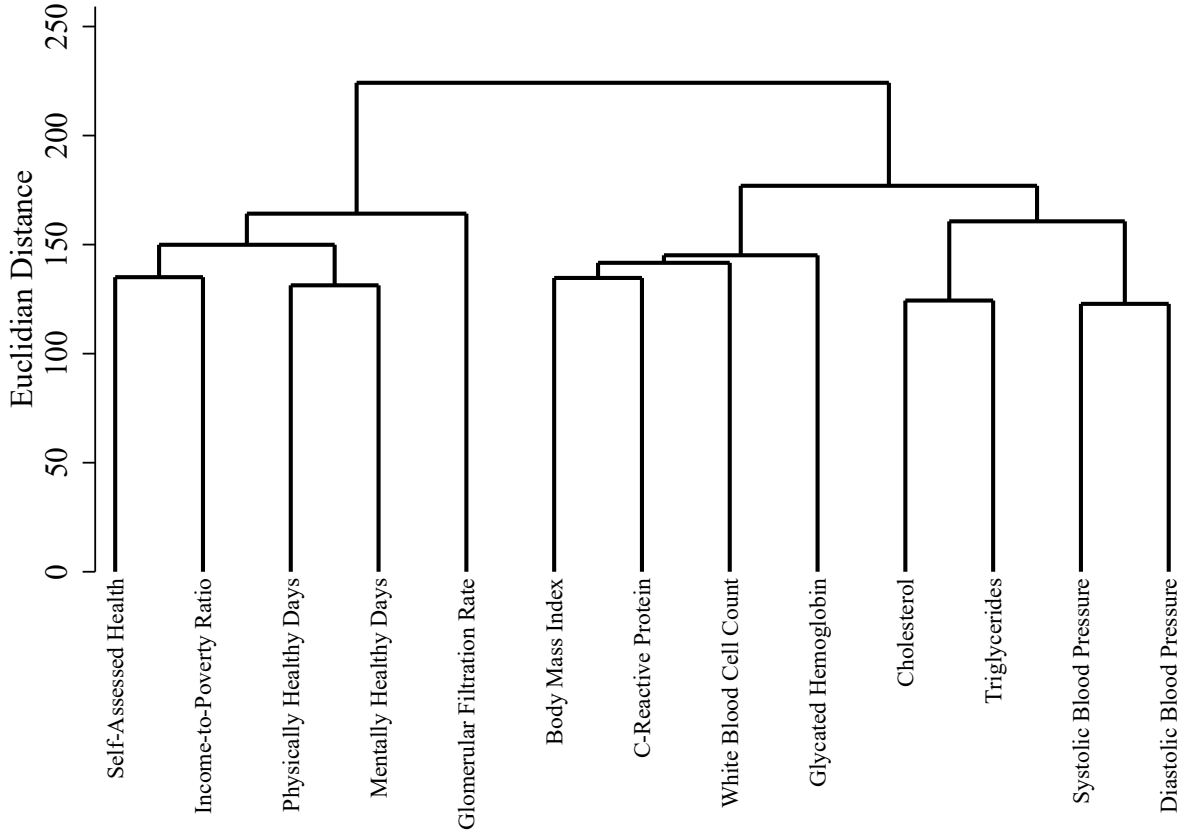
Notes: Data from the 2005-2006, 2007-2008 and 2009-2010 waves of NHANES including adults over age 20.

Figure 3: Relative Frequency Chart of Allostasis Health Levels



Notes: Data from the 2005-2006, 2007-2008 and 2009-2010 waves of NHANES including adults over age 20. Higher values for allostasis correspond to better health. For exposition, the values of allostasis have been rounded to the nearest 0.5.

Figure 4: Dendogram of Agglomative Clusters of Health and Income-to-Poverty Ratio



Notes: Data from the 2005-2006, 2007-2008 and 2009-2010 waves of NHANES including adults over age 20. Self-assessed health is cardinalized using the log-normal transformation method suggested in Wagstaff and van Doorslaer (1994), and clustering is done using Ward's method.

Table 1: Summary of Allostasis Biomarkers

	Mean	Std. Dev.	Clinical Threshold for Low Risk	Fraction Low Risk	Source of Clinical Threshold
Body Mass Index	28.64	6.50	<25	0.29	CDC, WHO
Cholesterol	198.12	41.27	<240 mg/dL	0.85	NHLBI, Mayo
Triglycerides	157.05	128.22	<200 mg/dL	0.77	NHLBI, Mayo
Glycated Hemoglobin	5.55	0.85	<6.5%	0.90	NLM, ADA
Glomerular Filtration Rate	94.03	21.20	$\geq$ 60 mL/min/1.73m^2	0.91	NKF & NIDDK
C-Reactive Protein	0.38	0.75	<2 mg/dL	0.97	Mayo
White Blood Cell Count	7.21	2.23	$\geq$ 3.5K & $\leq$ 10.5K cells/uL	0.92	Mayo
Systolic Blood Pressure	121.16	16.83	<140 mm Hg	0.83	AMA, Mayo
Diastolic Blood Pressure	70.42	11.68	<90 mm Hg	0.95	AMA, Mayo
N	11751				

Notes: Data from the 2005-2006, 2007-2008 and 2009-2010 waves of NHANES including adults over age 20. "Fraction Low Risk" indicates the proportion of our sample that falls within the normal clinical range for each biomarker. Source abbreviations are as follows: CDC—Centers for Disease Control and Preventions, WHO—World Health Organization, NHLBI—National Heart Lung and Blood Institute, Mayo—Mayo Medical Laboratories, NLM—National Library of Medicine, ADA—American Diabetes Association, NKF—National Kidney Foundation, NIDDK—National Institute of Diabetes and Digestive and Kidney Diseases, AMA—American Medical Association.

Table 2: Summary Statistics

Variable	Mean	Std.Dev	Min	Max
Income-to-Poverty Ratio	3.115	1.618	0.000	5.000
Allostasis	7.507	0.956	2.349	8.587
Self-Assessed Health (1-5)	3.386	0.939	1.000	5.000
Physically Healthy Days	26.541	7.615	0.000	30.000
Mentally Healthy Days	26.096	7.748	0.000	30.000
Mentally/Physically Healthy Days	28.273	5.563	0.000	30.000
Female	0.497	0.500	0.000	1.000
Age/100	0.470	0.163	0.200	0.850
Black	0.098	0.298	0.000	1.000
Hispanic	0.118	0.322	0.000	1.000
Married	0.657	0.475	0.000	1.000
Widowed	0.056	0.229	0.000	1.000
Divorced	0.128	0.334	0.000	1.000
Never Married	0.159	0.366	0.000	1.000
College Degree	0.275	0.447	0.000	1.000
Some College	0.308	0.462	0.000	1.000
High School Diploma	0.241	0.427	0.000	1.000
Less than High School	0.177	0.381	0.000	1.000
N	11751			

Notes: Data from the 2005-2006, 2007-2008 and 2009-2010 waves of NHANES including adults over age 20. Summary statistics are weighted using NHANES sampling weights.

Table 3: Likelihood Ratio Tests for Reporting Bias

	SES		Demographic		
	Income	Education	Race	Sex	Age
λ^R	-15088.7	-15111.2	-15111.2	-15111.2	-15131.3
Test for Reporting Bias					
$\sum_m \lambda_m^U$	-14935.5	-14998.1	-14973.8	-15077.5	-15074.3
$-2 \cdot (\lambda^R - \sum_{m=1}^M \lambda_m^U)$	306.5	226.2	274.9	67.5	73.9
Deg. of Freedom	170	51	36	19	19
χ^2 test p-value	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
Test for Cut-Point Shift					
$\sum_m \lambda_m^C$	-15039.5	-15046.4	-15020.4	-15106.1	-15104.2
$-2 \cdot (\lambda^R - \sum_{m=1}^M \lambda_m^C)$	98.4	129.7	181.6	10.2	54.2
Deg. of Freedom	40	12	8	4	4
χ^2 test p-value	**0.000**	**0.000**	**0.000**	**0.037**	**0.000**
Test for Index Shift					
$-2 \cdot (\sum_m \lambda_m^C - \sum_{m=1}^M \lambda_m^U)$	208.1	96.5	93.3	57.2	59.9
Deg. of Freedom	130	39	28	15	15
χ^2 test p-value	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**

Notes: Tests are for reporting bias by the SES or demographic category in the column title. Likelihood values taken from three sets of ordered probit regressions as follows: Likelihood values for the restricted models (λ^R) come from a single ordered probit regression with dummy variables to index SES and demographic categories. There are 11 subgroups of income (all models other than "Income" control for income quartiles), 4 subgroups of education, 3 subgroups of race, 2 subgroups for sex, and 2 subgroups for age. The likelihood value for the sum of the unrestricted models ($\sum_m \lambda_m^U$) comes from running a separate ordered probit model for each subgroup. The likelihood value for the sum of the constrained models ($\sum_m \lambda_m^C$) comes from running a separate ordered probit model for each subgroup while constraining the coefficient estimates to be equal to those of the single, restricted model and allowing the cut-points to vary. Degrees of freedom are determined by the number of restricted parameters in the likelihood ratio test. All regressions use robust standard errors.

Table 4: Likelihood Ratio Tests for Income-Based Reporting Bias by Sub-Population

		Race			Sex		Age	
	Full	White	Black	Hisp.	Male	Female	≤ 45	> 45
λ^R	-15111.2	-8360.9	-2801.3	-3811.5	-7864.9	-7212.6	-6312.6	-8761.6
Test for Reporting Bias								
$\sum_m \lambda_m^U$	-15007.8	-8293.4	-2770.2	-3776.9	-7809.6	-7147.6	-6262.0	-8713.6
$-2 \cdot (\lambda^R - \sum_{m=1}^M \lambda_m^U)$	206.8	135.1	62.1	69.3	110.7	130.0	101.3	96.1
Deg. of Freedom	51	45	45	45	48	48	45	45
χ^2 test p-value	**0.000**	**0.000**	**0.046**	**0.011**	**0.000**	**0.000**	**0.000**	**0.000**
Test for Cut-Point Shift								
$\sum_m \lambda_m^C$	-15067.5	-8339.3	-2791.8	-3792.8	-7842.1	-7189.7	-6286.8	-8738.9
$-2 \cdot (\lambda^R - \sum_{m=1}^M \lambda_m^C)$	87.5	43.3	19.0	37.4	45.8	45.8	51.7	45.5
Deg. of Freedom	12	12	12	12	12	12	12	12
χ^2 test p-value	**0.000**	**0.000**	0.089	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
Test for Index Shift								
$-2 \cdot (\sum_m \lambda_m^C - \sum_{m=1}^M \lambda_m^U)$	119.4	91.8	43.2	31.9	65.0	84.2	49.6	50.6
Deg. of Freedom	39	33	33	33	36	36	33	33
χ^2 test p-value	**0.000**	**0.000**	0.111	0.522	**0.002**	**0.000**	**0.032**	**0.025**

Notes: Tests are for reporting bias by income using income quartiles. Result from each column (except "Full") are from regressions using only the subsample of our data indicated by the column title. Likelihood values taken from three sets of ordered probit regressions as follows: Likelihood values for the restricted models (λ^R) come from a single ordered probit regression with dummy variables to index income quartiles. The likelihood value for the sum of the unrestricted models ($\sum_m \lambda_m^U$) comes from running a separate ordered probit model for each income quartile. The likelihood value for the sum of the constrained models ($\sum_m \lambda_m^C$) comes from running a separate ordered probit model for each income quartile while constraining the coefficient estimates to be equal to those of the single, restricted model and allowing the cut-points to vary. Degrees of freedom are determined by the number of restricted parameters in the likelihood ratio test. All regressions use robust standard errors.

Table 5: Regression Results of Self-Assessed Health on Income and Allostasis

	Self-Assessed Health	Physically Healthy Days	Mentally Healthy Days	Controlling for Physically and Mentally Healthy Days
Allostasis	0.217 (0.015) ***	0.321 (0.100) ***	0.262 (0.104) **	0.204 (0.015) ***
IPR x Allostasis	0.052 (0.010) ***	0.015 (0.061)	-0.031 (0.059)	0.052 (0.010) ***
Mentally Healthy Days				0.017 (0.002) ***
IPR x Mentally Healthy Days				0.002 (0.001)
Physically Healthy Days				0.027 (0.002) ***
IPR x Physically Healthy Days				0.001 (0.001)
Income-to-Poverty Ratio	0.098 (0.038) **	0.406 (0.197) **	0.864 (0.197) ***	0.072 (0.038) *
Income-to-Poverty Ratio Squared	-0.005 (0.009)	-0.227 (0.049) ***	-0.136 (0.052) ***	0.002 (0.009)
College Degree	0.706 (0.127) ***	1.031 (0.586) *	2.484 (0.614) ***	0.645 (0.124) ***
College Degree x IPR	-0.003 (0.038)	0.091 (0.188)	-0.396 (0.183) **	-0.001 (0.037)
Some College	0.285 (0.080) ***	-0.321 (0.483)	0.277 (0.521)	0.290 (0.077) ***
Some College x IPR	0.015 (0.031)	0.261 (0.176)	-0.094 (0.174)	0.009 (0.030)
High School Diploma	0.206 (0.072) ***	0.157 (0.498)	0.929 (0.503) *	0.185 (0.071) ***
High School Diploma x IPR	-0.004 (0.031)	0.087 (0.190)	-0.345 (0.181) *	0.000 (0.030)
Age/100	-0.243 (0.108) **	-3.972 (0.587) ***	4.719 (0.590) ***	-0.216 (0.106) **
Age/100 x IPR	0.188 (0.072) ***	0.417 (0.379)	0.387 (0.352)	0.165 (0.072) **
Age/100 Squared	2.161 (0.549) ***	15.134 (3.151) ***	25.171 (3.163) ***	1.375 (0.541) **
Age/100 Squared x IPR	-1.226 (0.371) ***	-8.273 (2.015) ***	-7.647 (1.830) ***	-0.918 (0.367) **
Female	-0.136 (0.059) **	-1.235 (0.349) ***	-2.873 (0.368) ***	-0.066 (0.058)
Female x IPR	0.010 (0.020)	0.154 (0.100)	0.356 (0.102) ***	0.003 (0.019)
Black	-0.024 (0.071)	0.835 (0.442) *	1.328 (0.484) ***	-0.067 (0.069)
Black x IPR	-0.074 (0.023) ***	-0.194 (0.134)	-0.341 (0.152) **	-0.063 (0.023) ***
Hispanic	-0.211 (0.067) ***	1.577 (0.390) ***	3.132 (0.440) ***	-0.298 (0.066) ***
Hispanic x IPR	-0.000 (0.024)	-0.571 (0.136) ***	-0.832 (0.148) ***	0.028 (0.024)
Married	-0.001 (0.092)	-0.296 (0.445)	0.631 (0.523)	-0.002 (0.090)
Married x IPR	-0.002 (0.033)	0.146 (0.144)	-0.102 (0.157)	-0.005 (0.032)
Widowed	0.024 (0.130)	-0.281 (1.026)	1.426 (0.865) *	0.007 (0.128)
Widowed x IPR	-0.026 (0.050)	-0.063 (0.357)	-0.601 (0.278) **	-0.013 (0.049)
Divorced	-0.066 (0.113)	-0.308 (0.644)	-0.666 (0.742)	-0.051 (0.110)
Divorced x IPR	0.015 (0.042)	-0.160 (0.226)	0.007 (0.226)	0.021 (0.042)
Adj. R^2	0.109	0.041	0.058	0.141
Num Obs	11751	11751	11751	11751
P-value for Chow test	0.000	0.000	0.000	0.000

Notes: Data from the 2005-2006, 2007-2008 and 2009-2010 waves of NHANES including adults over age 20. Each column shows coefficients from a regression using different measures of self-assessed health. In addition to the coefficients shown, all models include survey wave fixed effects. All regressions include NHANES sample weights. Robust standard errors are shown in parenthesis, and stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%

Table 6: Concentration Index Results Using Income-to-Poverty Ratio

	Self-Reported Health			
	Self-Assessed Health	Physically Healthy Days	Mentally Healthy Days	Allostasis
Income-to-Poverty Ratio	0.058***	0.017***	0.020***	0.005***
	(0.006)	(0.002)	(0.002)	(0.001)
Age/100	-0.059***	-0.026***	0.025***	-0.039***
	(0.010)	(0.003)	(0.003)	(0.001)
Age/100 Squared	0.225***	0.088***	0.150***	0.056***
	(0.053)	(0.018)	(0.018)	(0.007)
Female	-0.004	-0.004***	-0.010***	0.005***
	(0.003)	(0.001)	(0.001)	(0.000)
Black	-0.022***	0.002	0.002	-0.002***
	(0.003)	(0.001)	(0.001)	(0.001)
Hispanic	-0.022***	0.001	0.007***	-0.002***
	(0.004)	(0.001)	(0.001)	(0.001)
Married	0.000	0.001	0.003	0.000
	(0.005)	(0.001)	(0.002)	(0.001)
Widowed	-0.007	-0.003	-0.001	-0.003***
	(0.007)	(0.003)	(0.003)	(0.001)
Divorced	-0.004	-0.005**	-0.004*	-0.001
	(0.006)	(0.002)	(0.002)	(0.001)
College Degree	0.070***	0.007***	0.008***	0.005***
	(0.005)	(0.002)	(0.002)	(0.001)
Some College	0.029***	0.002	0.000	-0.000
	(0.004)	(0.002)	(0.002)	(0.001)
High School Diploma	0.018***	0.002	0.000	0.000
	(0.004)	(0.002)	(0.002)	(0.001)
Adj. R^2	0.084	0.035	0.051	0.139
Num Obs	11751	11751	11751	11751

Notes: Data from the 2005-2006, 2007-2008 and 2009-2010 waves of NHANES including adults over age 20. Each regression shows coefficients from a concentration index model, and robust standard errors are shown in parenthesis. The first three columns measure health using self-assessed health measures, and the remaining column measures allostasis. In addition to the coefficients shown, all models include survey wave fixed effects. All models include NHANES sample weights, and stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%

Table 7: Concentration Index Results
Other Income Measures, Education, and Sample Weights

	Self-Reported Health				
	Self-Assessed Health	Physically Healthy Days	Mentally Healthy Days	Allostasis	Allostasis (Count)
Income-to-Poverty Ratio Incl. Educ. (Weighted)	0.058***	0.017***	0.020***	0.005***	0.006***
	(0.006)	(0.002)	(0.002)	(0.001)	(0.001)
Income-to-Poverty Ratio Incl. Educ. (Unweighted)	0.071***	0.018***	0.020***	0.005***	0.006***
	(0.006)	(0.002)	(0.002)	(0.001)	(0.001)
Family Income (Category) Incl. Educ. (Weighted)	0.052***	0.018***	0.018***	0.004***	0.005***
	(0.006)	(0.002)	(0.002)	(0.001)	(0.001)
Family Income (Category) Incl. Educ. (Unweighted)	0.063***	0.019***	0.019***	0.004***	0.005***
	(0.006)	(0.002)	(0.002)	(0.001)	(0.001)
Household Income (Category) Incl. Educ. (Weighted)	0.052***	0.016***	0.018***	0.005***	0.006***
	(0.006)	(0.002)	(0.002)	(0.001)	(0.001)
Household Income (Category) Incl. Educ. (Unweighted)	0.063***	0.018***	0.020***	0.005***	0.006***
	(0.006)	(0.002)	(0.002)	(0.001)	(0.001)
Income-to-Poverty Ratio No Educ. (Weighted)	0.094***	0.020***	0.024***	0.007***	0.009***
	(0.006)	(0.002)	(0.002)	(0.001)	(0.001)
Income-to-Poverty Ratio No Educ. (Unweighted)	0.113***	0.023***	0.023***	0.007***	0.009***
	(0.005)	(0.002)	(0.002)	(0.001)	(0.001)
Family Income (Category) No Educ. (Weighted)	0.085***	0.022***	0.022***	0.006***	0.008***
	(0.006)	(0.002)	(0.002)	(0.001)	(0.001)
Family Income (Category) No Educ. (Unweighted)	0.103***	0.023***	0.023***	0.007***	0.008***
	(0.005)	(0.002)	(0.002)	(0.001)	(0.001)
Household Income (Category) No Educ. (Weighted)	0.086***	0.020***	0.022***	0.007***	0.009***
	(0.006)	(0.002)	(0.002)	(0.001)	(0.001)
Household Income (Category) No Educ. (Unweighted)	0.103***	0.023***	0.023***	0.007***	0.009***
	(0.005)	(0.002)	(0.002)	(0.001)	(0.001)

Notes: Data from the 2005-2006, 2007-2008 and 2009-2010 waves of NHANES including adults over age 20. Each cell shows coefficients from a concentration index model. Each row shows results using a different measure of income and either using NHANES sample weights or not. In addition to the income coefficient shown, all models additionally control for the same controls as in Table 6. The first three columns measure health using self-assessed health measures, and the remaining column measures allostasis. Robust standard errors are shown in parenthesis, and stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%.

A Appendix

Methodological Issues with the Concentration index

Despite its ubiquity in the health inequality literature, many potential methodological problems with the concentration index have been identified. These challenges include concerns regarding measurement error (Fleurbaey & Schokkaert, 2009), the relationship between income and health across the distribution of health (Carrieri & Jones, 2017), and evidence that experimental subjects do not have a preference for decreasing the correlation between health and SES (Bleichrodt, Rohde, & Ourti, 2012). Another debate concerns the appropriateness of the use of the concentration index for bounded variables (Clarke, Gerdtham, Johannesson, Bingefors, & Smith, 2002; Erreygers, 2009a; Wagstaff, 2009; Erreygers, 2009b; Erreygers & Van Ourti, 2011a; Lambert & Zheng, 2011; Wagstaff, 2011; Erreygers & Van Ourti, 2011b; Kjellsson & Gerdtham, 2013). The debate began with Clarke et al. (2002), who show that income-related inequality rankings between two countries can reverse depending on whether health is measured in terms of health achievement or shortfall from some maximum state of health. Erreygers (2009a) proposes a "corrected concentration index" which attempted to remedy this issue as well as the more general problem of correcting the mean-dependence of the bounds of the concentration index that occurs for bounded variables. Wagstaff (2009) challenges Erreygers (2009a) on the notion that his proposed correction is an absolute measure of inequality, not a relative one. In response, Erreygers (2009b) note that the terms "absolute" and "relative" inequality lose their traditional meaning when dealing with bounded variables. The debate continues and has become more nuanced (Erreygers & Van Ourti, 2011a; Lambert & Zheng, 2011; Wagstaff, 2011; Erreygers & Van Ourti, 2011b). Kjellsson and Gerdtham (2013) clarifies some of the difficulties of this debate—at least for binary health variables—by demonstrating that the differences in the measures are implicitly due to different assumptions about the most unequal society. Thus, the choice of SES-related inequality index is as much a normative judgment as it is a technical one. Finally, the very plausible scenario of reverse causality—that is, that SES is determined by health—is rarely considered or addressed with regard to the concentration index.

While we do not think many of these critiques should be taken lightly, addressing all of them is beyond the scope of this paper. Therefore, like the majority of papers in this literature, our analysis assumes that the concentration index is a reasonable measure of SES-related health inequality. However, we do test the robustness of our concentration index results to using the Erreygers transformation in Appendix Table A19.

Construction of Allostasis

As we discuss in Section 2.3, in order to accurately assess income-related health inequality, we need a representative measure of health that can be objectively measured, and, thus, is not affected by income-related reporting heterogeneity. For this, a natural place to look is for measures collected during clinical assessment since the objective assessment of health is precisely the purpose of these measures. However, because health is a complex concept, any single clinical measure is unlikely to capture the fullness of what makes a person healthy or not. Thus, we argue that a multivariate measure is needed. This decision creates its

own set of difficulties, though. What measures should be used to construct the multivariate measure? How is each measure used to assess healthiness? And how should the measures be combined? Below, we explore in greater detail how we address each of these problems.

Which measures?

To answer to first question of what measures should be included, we employ the concept of allostatic load from the medical literature. First developed by McEwen and Stellar (1993), allostatic load is a measure of cumulative physiological deterioration across a number of biological systems relevant to disease risk. Allostatic load is indicated by some combination of two categories of biomarkers: primary mediators and measures of the body's response to chronic exposure to the primary mediators. Primary mediators are substances that the body releases as an acute response to stress, such cortisol or epinephrine (Geronimus et al., 2006). (McEwen & Stellar, 1993) note that chronic exposure to stress (and the primary mediators produced to respond to stress) comes at a cost to the body's adaptive systems and can result in chronic illnesses, such as hypertension, obesity, diabetes, inflammation, and immunosuppression. Biomarkers used to identify these conditions make up the second category of indicators of allostatic load: the measures of the body's response to chronic exposure to the primary mediators.

We use (Geronimus et al., 2006), who calculated allostatic load using an earlier wave of NHANES, as our primary reference for which measures to include. (Geronimus et al., 2006) use diastolic and systolic blood pressure and homocysteine as indicators of the cardiovascular system; BMI, cholesterol, triglycerides, and glycated hemoglobin as indicators of the metabolic system; and c-reactive protein, albumin, and creatine clearance as indicators for immune system and kidney function. We utilize the same set of measures, with the exception of homocysteine and albumin. NHANES stopped collecting data on homocysteine in the 2007-2008 wave; so it was not available for our sample. Although our sample did include urinary albumin, there was very little variation in the measure for our sample. Additionally, we chose to add a direct measure of immune system function: white blood cell count. We elaborate on how we use these measures to classify healthy and and unhealthy individuals in the next section.

When calculating a typical measure of allostatic load, individuals receive one point for each clinical measure for which they fall over a high-risk threshold, and the cumulative number of points is the measure of allostatic load. We utilize this method for our count allostasis measure. Additionally, we calculate a weighted allostasis measure, which we discuss below.

What is healthy?

To address the question of how each measure should be used to assess healthiness, we utilize existing laboratory and clinical standards in the US for what constitutes a healthy range or "low risk" for each laboratory test or clinical assessment using the most relevant source. We provide the details for each measure below. The majority of this information is also briefly summarized in Table 1.

- BMI is calculated as an individual's weight in kilograms divided by the square of the

individual's height in meters. According to the CDC and World Health Organization, an individual can be classified into one of four categories based on their BMI: <18.5 is considered underweight, 18.5 to 24.9 is considered normal or healthy weight, 25.0 to 29.9 is considered overweight, and >30 is considered obese. [22] Following this standard, we classify an individual with a BMI below 25 as healthy.[23]

- We include two measures of blood lipid levels: total cholesterol and triglycerides. Higher cholesterol and triglyceride levels are linked to many heart conditions (Stamler et al., 2000; Criqui et al., 1993). Both measures were assessed via blood sample and measured in milligrams per deciliter (mg/dL). Standards at both the National Heart, Lung, and Blood Institute (NHLBI) and Mayo Clinic cite the US standard for high cholesterol to be a level of 240 mg/dL or greater. Therefore we classify individuals as healthy if their total cholesterol is below 240 mg/dL. Similarly, both the NHLBI and Mayo Clinic cite the US standard for high triglyceride levels to be to be a level of 200 mg/dL or greater. Therefore we classify individuals as healthy if their triglycerides are below 200 mg/dL.[24]

- Glycated Hemoglobin, also referred to as glycohemoglobin and A1C level, is a measure of blood glucose control for the previous four months (Trivelli et al., 1971). Hemoglobin—a protein in the red blood cells responsible for carrying oxygen around the body—will bind with glucose in the body stream and become glycated. The percentage of hemoglobin that is glycated is a stable measure of the body's ability to regulate blood glucose levels. According to the National Library of Medicine and the American Diabetes Association, a glycated hemoglobin percentage of 6.5 or above would constitute a diagnosis of Type 2 diabetes. We, therefore, classify an individual as healthy if their glycated hemoglobin percentage is below 6.5.[25]

- Glomerular filtration rate (GFR) is a measure of kidney function. Measuring GFR accurately is a difficult and time consuming; thus, the most common way to measure it is by using serum marker, typically creatinine. Creatinine, a metabolite of creatine, is released into the blood stream at a near constant rate, is filtered through the kidneys, and excreted in the urine. Thus, if GFR declines, then serum creatinine will begin to spike, but does so at a non-linear rate. This results in the possibility that a person within the normal reference interval for serum creatine will not have a normal GFR. Thus, the Nation Kidney Disease Education Program (NKDEP) recommends that laboratories use serum creatinine levels to calculate an estimated GFR (eGFR) based on established equations (NIDDK, 2016). The NKDEP recommends using one of two equations: the Modification of Diet in Renal Disease (MDRD) Study equation or the Chronic Kidney Disease Epidemiological Collaboration (CKD-EPI) equation. We used

[22] See `https://www.cdc.gov/healthyweight/assessing/bmi/adult_bmi/index.html`. Last accessed June 18, 2018.

[23] We could also consider an individual with a BMI below 18.5 as unhealthy; however, there were very few adults in our sample with a BMI below 18.5.

[24] See `https://www.mayoclinic.org/diseases-conditions/high-blood-cholesterol/diagnosis-treatment/drc-20350806`. Last accessed June 18, 2019.

[25] See `https://medlineplus.gov/ency/article/003640.htm`. Last accessed June 18, 2019.

the CKD-EPI equation because the NKDEP notes that it is slightly more accurate, particularly for patients with mild kidney damage. The equation was estimated using to two-slope spline model based on serum creatinine, age, gender, and race. It estimates an approximate GFR in milliliters per minute and assume an average adult surface area of 1.73 m^2 ($mL/min/1.73m^2$). The specifics of the equation can be found on the NKDEP website or in the original published article (NIDDK, 2016; Levey et al., 2009). According to the National Kidney Foundation, an individual with an eGFR less that 60 $mL/min/1.73m^2$ is considered to be at increased risk for chronic kidney disease. As such, we classify an individual as healthy if his or her eGFR is 60 $mL/min/1.73m^2$ or above.[26]

- C-reactive protein (CRP) levels are assessed via blood test and reported in mg/dL. CRP increases when there in inflammation in the body, which can be due to an infection or chronic inflammatory conditions. According to the Mayo Clinic, a CRP test above 10 mg/dL is a sign of serious infection; however, CRP levels above 2 mg/dL are associated with increased risk of heart disease. Thus, we classify an individual as being healthy if CRP levels are below 2 mg/dL.[27]

- White blood cells, or leukocytes, are used to fight infections and remove foreign objects from the body. A white cell count that is too low may indicate a serious medical condition such as an autoimmune disorder or cancer. A white cell count that is too high may be an indication of an infection or inflammation. At the time our analysis was preformed, the Mayo Clinic defined a normal range for white cell count to be between 3500 to 10,500 cells per microliter (uL), and we, therefore, defined an individual as healthy if his or her white cell count was in that range. It is worth noting that in May of 2018 the Mayo Clinic updated its reference values such that a normal white cell count is now considered to be between 3400 and 9600 cells/uL.[28] Had we used this standard, the proportion of our sample classified as healthy would have decreased by about 4 percentage points.

- Systolic and diastolic blood pressure are measures of the pressure in an individual's arteries when the heart beats and the pressure in between heart beats (respectively). They are measured in millimeters of mercury (mm Hg). NHANES examiners take three consecutive blood pressure readings, and we use the average of the second and third reading. According to the American Medical Association and the Mayo Clinic, a normal blood pressure reading is below 120 mm Hg for systolic and below 80 mm Hg for diastolic. Historically, an individual would be classified as having high blood pressure, or hypertension, if his or her systolic blood pressure was 140 mm Hg or above or his or her diastolic blood pressure was 90 mm Hg or above. In November of 2017, the American Heart Association published updated guidelines that classified a blood pressure reading of 130/80 mm Hg or higher as stage 1 hypertensive and a

[26]See `https://www.kidney.org/atoz/content/gfr`. Last accessed June 18, 2019.

[27]See `https://www.mayoclinic.org/tests-procedures/c-reactive-protein-test/about/pac-20385228`. Last accessed June 18, 2019.

[28]See `https://www.mayocliniclabs.com/test-notifications/attachment.php?id=55172`. Last accessed June 18, 2019.

reading of 140/90 mm Hg as stage 2 hypertensive. However, individuals with stage 1 hypertension and no other cardiovascular disease risks are not considered high risk enough to warrant pharmaceutical treatment (Whelton et al., 2018). Considering this along with the tendency for blood pressure to fluctuate throughout the day and become elevated during a clinical assessment (sometimes referred to as white coat hypertension), we chose to use the previous guidelines (now stage 2 hypertension) to define our cut-point.

How to combine?

As mentioned and described in Section 2.3, we use multiple correspondence analysis (MCA), which is analogous to principal components analysis for categorical variables, to aggregate the dimensions of health into one summary measure. The core piece of MCA is an $2q \times n$ indicator matrix, $\mathbf{Z}$, where q is the number of variables used for MCA (for our analysis, $q = 9$). All the results that one can obtain from MCA—including the standard coordinates used for graphical analysis and the row coordinates that we use for our health index—can be found through algebraic manipulations of $\mathbf{Z}$. To understand how this measure differs from a traditional count measure of allostasis, it is helpful to examine the standard coordinates resulting from MCA, which are the foundation for the weights used to calculate the index from $\mathbf{Z}$. The standard coordinates are calculated using the first eigenvector of the Burt's matrix, $\mathbf{Z}'\mathbf{Z}$, and a diagonal matrix of the column masses, which are themselves the result of an algebraic manipulation of the Burt's matrix.[29] To calculate the column masses, one would multiply the Burt's matrix by a vector of ones and divide by the sum total of all cross-tabulations in the Burt's matrix. The standard coordinates from our analysis are found in Table A1. The first column shows the weight used for individuals in the healthy range of the corresponding clinical measure, and the second column shows the weight used for individuals not in the healthy range of the corresponding clinical measure. The final column shows the difference between the weights and can be interpreted as the absolute (raw) change in the health index from an individual moving from the unhealthy range of a clinical measure to the healthy range. These weights make it clear that the critical assumption of a count allostasis measure—equal weighting—is not satisfied. Some measures seem relatively less important (e.g. c-reactive protein and white blood cell count), while others are relatively much more important (glycated hemoglobin and systolic and diastolic blood pressure).

To replicate the predicted row score generated by most statistical software (we use Stata v.14), one would need to sum weights from either the healthy (1) column or unhealthy (0) column—according to each individuals' health status—then divide by the product of the number of variables used to construct the measure (9) and the first eigenvalue of the Burt's matrix. Finally, we scale the resulting row score—which is constructed to have a mean of zero—by the mean of the count allostasis measure to make the two directly comparable. This gives the resulting MCA index the interpretation of a weighted allostasis measure.

[29]The Burt's matrix must be centered before the eigen decomposition to remove the trivial solution (see Kohn, 2012).

Cluster Analysis Robustness Checks

We begin with demonstrating the robustness of our cluster analysis results. We test the robustness of these results to using household income grouped into 11 categories, instead of the income-to-poverty ratio used in our main results. This analysis, summarized in Appendix Figure A1, finds nearly identical results, with the exception that the glomular filtration rate clusters with the self-reported health and income variables instead of clustering with the other clinical markers of health.[30]

Ordered Probit Models Robustness Checks

We next demonstrate the robustness of our ordered probit results. To this end, we reestimate our results using allostasis as calculated by summing the dimensions of health clinical health marker levels and also using the individual biomarkers instead of allostasis and found very similar results. These tables are reported in Appendix Tables A2 - A7. We see a very similar pattern of results to our main results.

Characterizing SES-based reporting heterogeneity in SAH Robustness Checks

We first examine whether our results in Table 5 are affected by using the count measure of allostasis or the aggregation of clinical health measures. To this end, we run regressions using the count allostasis metric, and we also run regressions where we include a single clinical measure of health in place of the allostasis metric. Results from these regressions, found in Appendix Tables A8 and A9, show a pattern of estimates that is very similar to our main results. We additionally test the robustness of our results to not using NHANES sample weights, shown in Appendix Table A10, and to using alternate versions of SAH, shown in Appendix Table A11. Specifically, we use the SAH response from the initial home interview (the HUQ questionnaire) in addition to than SAH response from the physical examination (used in our main results), as well as the two SAH measures transformed into dichotomous measures[31] We also estimate results for various demographic subsamples (sex, race, age, and educational attainment), removing individuals with top-coded income-to-poverty ratios, and by NHANES survey wave, all of which are shown in Appendix Table A12. All of the robustness checks show results that are very similar to our main results. Finally, since these results are based on calculating conditional correlations and not concentration indices, we are able to check the robustness of these results to using the original 1-5 SAH measure in a generalized ordered probit model. These results are found in Appendix Table A13. These results are consistent with our main results and show a positive income reporting bias.

[30]Income categories are: \$0-\$5K, \$5K-\$10K, \$10K-\$15K, \$15K-\$20K, \$20K-\$25K, \$25K-\$35K, \$35K-\$45K, \$45K-\$55K, \$55K-\$65K, \$65K-\$75K, and over \$75K. We use the middle dollar amount suggested by each category. For example an individual in the first category is assigned an income of \$2,500, in the second category, \$7,500. Individuals in the highest category are assigned an income of \$100,000.

[31]The distribution of SAH as assessed in the HUQ questionnaire is very similar to SAH from the physical examination. The distribution of SAH as assessed in the HUQ questionnaire can be seen in Appendix Figure A2.

Another potential concern is that, because there is no universal health care coverage in the U.S., lower income individuals likely have less access to medical care. Individuals without regular medical care may be less aware of health problems, which could potentially lead to differences in self-reported health. Relatedly, our results may be affected by individuals who are on prescription medications, which may affect some clinical measures of health like blood pressure. The NHANES includes questions about access to care and the use of prescription drugs, which we can utilize to investigate these concerns. In our sample, only 15 percent of respondents did not have a routine place they visit for medical care, and only 17 percent of respondents did not see a doctor or health care professional in the last year. Appendix Table A14 show results excluding individuals that do not have a routine place to visit for medical care and did not see a health care professional in the last year. The estimates are nearly identical to our main results, which mitigates concerns that heterogeneity in access to care is affecting our results. Appendix Table A15 show results excluding individuals reporting prescription drug use. Here again the estimates are nearly identical to our main results, which mitigates concerns that prescription drug use affects our results.

Concentration Index Robustness Checks

Finally, we examine the robustness of the condition index results.

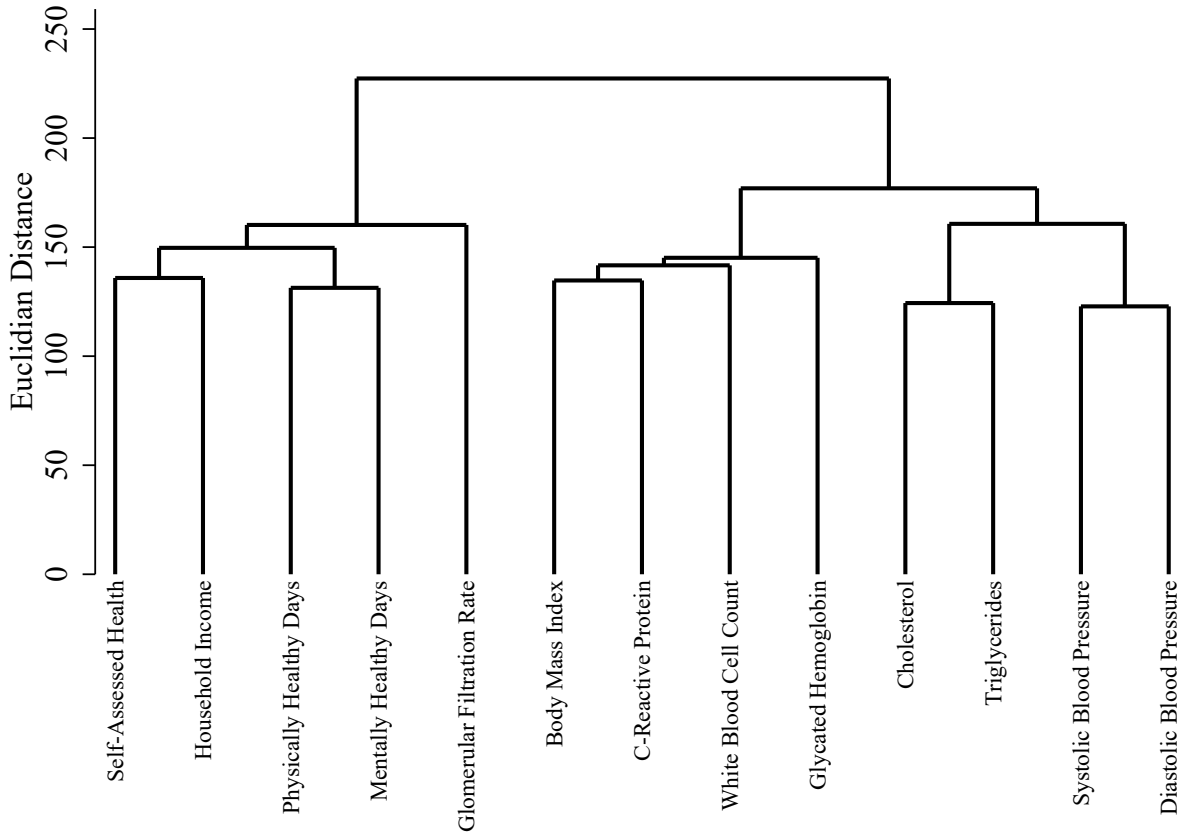
To check that these results are not due to the aggregation of different clinical measures into allostasis, we estimate concentration indexes separately for each clinical measure of health. These results are displayed in Appendix Table A16. All nine clinical measures produce concentration index results suggesting a lesser relationship between health inequality and income than suggested by the self-reported measures of health. Second, we ensure that access to medical care or prescription drug use are not driving our results, excluding the same groups of individuals as we did in Appendix Tables A14 and A15. These results, shown in Appendix Tables A17 and A18, again show a similar pattern to our main concentration index results.

Since all of our measure except cardinalized SAH are bounded, it might be more appropriate to use the Erreygers transformation; these results are found in Appendix Table A19. Finally, we examine different measures of SAH, including a SAH collected from a different NHANES questionnaire asked on a different date (cardinalized using the log-normal transform), dichotomized measures both SAH variables, and both SAH variables cardinalized using the predicted values of an interval regression on the bounds of the cumulative density function of the Canadian HUI-III health utility index mapped onto SAH (see (Ziebarth, 2010) for more information).[32] Appendix Table A20 shows these results. Results for the alternative SAH measure are similar to our main results. Both the dichotomized measures and the interval regression predictions show much higher levels of healthy inequality. The result on the dichotomized measures is consistent with (Ziebarth, 2010). The result on the HUI measures is unsurprising given the fact the dependent variables capture only the explained variation in SAH from a prediction that was made using income as an explanatory variable. If we instead estimate the predicted HUI index without using income, then the concentration index on the resulting measure is essentially zero ($= -2.16 \times 10^{-16}$) and not

[32]Mapping of SAH values on to the CDF of the HUI-III are taken from (Ziebarth, 2010).

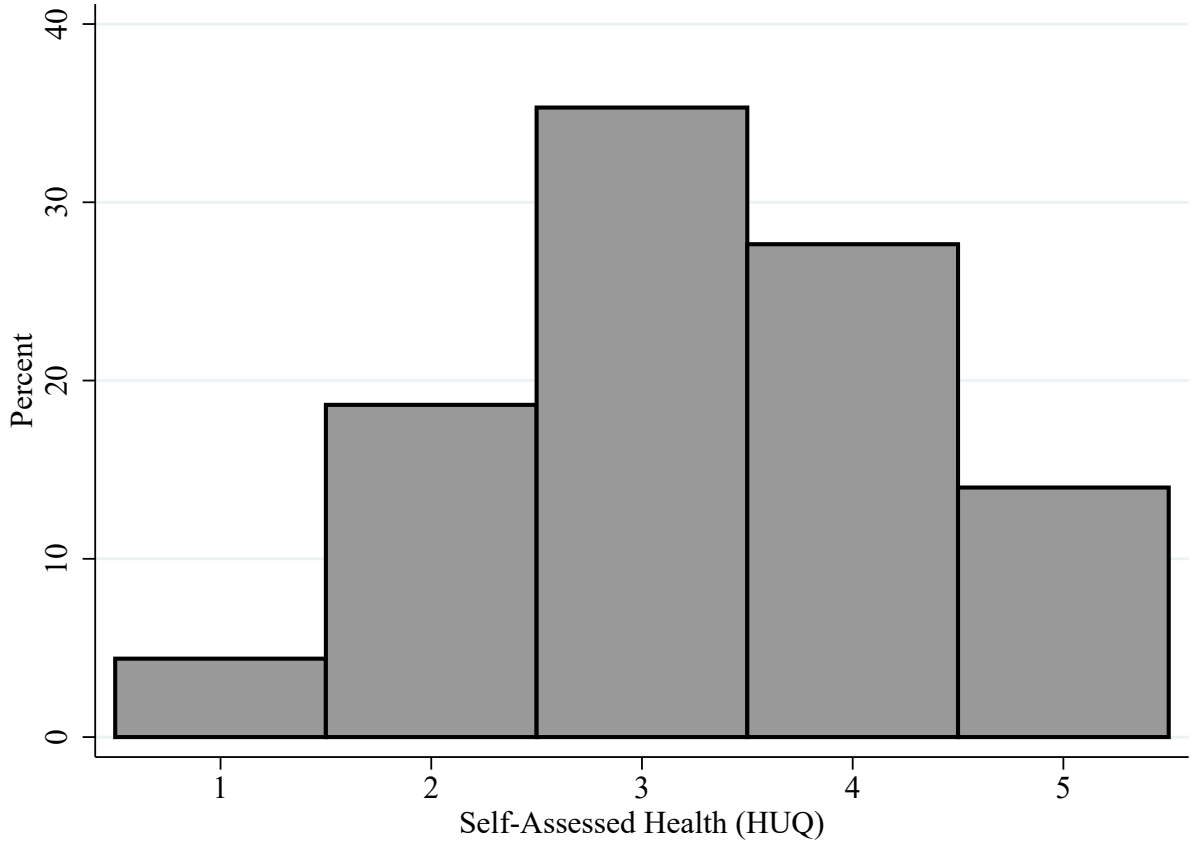
statistically significant.

Figure A1: Dendogram of Agglomative Clusters of Health and Household Income Categories



Notes: Data from the 2005-2006, 2007-2008 and 2009-2010 waves of NHANES including adults over age 20. Self-assessed health is transformed to fit an inverse log-normal distribution, and household income is transformed to show the middle dollar amounts for each category. Clustering is done using Ward's method.

Figure A2: Relative Frequency Chart of Self-Assessed Health Levels
From Health Utilization Questionnaire



Notes: Data from the 2005-2006, 2007-2008 and 2009-2010 waves of NHANES including adults over age 20. Higher values for SAH correspond to better self-assessed health.

Table A1: MCA Weights for Allostasis

Variable	MCA Weight (1)	MCA Weight (0)	Δ from 0 to 1
Body Mass Index	1.73	-0.69	2.42
Cholesterol	0.38	-2.22	2.61
Triglycerides	0.72	-2.39	3.11
Glycated Hemoglobin	0.38	-3.34	3.73
Glomerular Filtration Rate	0.22	-2.31	2.54
C-Reactive Protein	0.03	-1.09	1.12
White Blood Cell Count	0.10	-1.21	1.31
Systolic Blood Pressure	0.63	-3.11	3.74
Diastolic Blood Pressure	0.27	-4.95	5.22

Notes: This table shows weights from a MCA analysis. Each clinical marker of health is coded to be a 1 if the marker is in a healthy level and 0 if it is in an unhealthy level according to the cutoffs in Table 1. The health index is constructed by adding weights from either the healthy (1) column or unhealthy (0) column—according to each individuals' health status—then dividing by the product of the number of variables used to construct the measure (9) and the first eigenvalue of the Burt's matrix.

Table A2: Likelihood Ratio Tests for Reporting Bias using Count Allostasis

	SES		Demographic		
	Income	Education	Race	Sex	Age
λ^R	-15078.7	-15101.5	-15101.5	-15101.5	-15121.0
Test for Reporting Bias					
$\sum_m \lambda_m^U$	-14924.4	-14988.6	-14964.4	-15067.8	-15063.1
$-2 \cdot (\lambda^R - \sum_{m=1}^M \lambda_m^U)$	308.6	225.8	274.2	67.3	76.8
Deg. of Freedom	170	51	36	19	19
χ^2 test p-value	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
Test for Cut-Point Shift					
$\sum_m \lambda_m^C$	-15029.4	-15036.4	-15010.6	-15096.1	-15092.6
$-2 \cdot (\lambda^R - \sum_{m=1}^M \lambda_m^C)$	98.7	130.1	181.8	10.8	56.7
Deg. of Freedom	40	12	8	4	4
χ^2 test p-value	**0.000**	**0.000**	**0.000**	**0.028**	**0.000**
Test for Index Shift					
$-2 \cdot (\sum_m \lambda_m^C - \sum_{m=1}^M \lambda_m^U)$	209.9	95.8	92.3	56.4	59.1
Deg. of Freedom	130	39	28	15	15
χ^2 test p-value	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**

Notes: Tests are for reporting bias by the SES or demographic category in the column title. Likelihood values taken from three sets of ordered probit regressions as follows: Likelihood values for the restricted models (λ^R) come from a single ordered probit regression with dummy variables to index SES and demographic categories. There are 11 subgroups of income (all models other than "Income" control for income quartiles), 4 subgroups of education, 3 subgroups of race, 2 subgroups for sex, and 2 subgroups for age. The likelihood value for the sum of the unrestricted models ($\sum_m \lambda_m^U$) comes from running a separate ordered probit model for each subgroup. The likelihood value for the sum of the constrained models ($\sum_m \lambda_m^C$) comes from running a separate ordered probit model for each subgroup while constraining the coefficient estimates to be equal to those of the single, restricted model and allowing the cut-points to vary. Degrees of freedom are determined by the number of restricted parameters in the likelihood ratio test. All regressions use robust standard errors.

Table A3: Likelihood Ratio Tests for Income-Based Reporting Bias by Sub-Population using Count Allostasis

		Race			Sex		Age	
	Full	White	Black	Hisp.	Male	Female	≤ 45	> 45
λ^R	-15101.5	-8353.9	-2799.3	-3811.2	-7863.4	-7204.4	-6304.3	-8758.8
Test for Reporting Bias								
$\sum_m \lambda_m^U$	-14997.3	-8286.6	-2768.5	-3775.6	-7809.1	-7139.3	-6252.5	-8711.4
$-2 \cdot (\lambda^R - \sum_{m=1}^M \lambda_m^U)$	208.4	134.6	61.7	71.2	108.7	130.3	103.5	94.7
Deg. of Freedom	51	45	45	45	48	48	45	45
χ^2 test p-value	**0.000**	**0.000**	**0.049**	**0.008**	**0.000**	**0.000**	**0.000**	**0.000**
Test for Cut-Point Shift								
$\sum_m \lambda_m^C$	-15057.7	-8332.2	-2790.0	-3792.4	-7840.6	-7181.3	-6278.4	-8736.0
$-2 \cdot (\lambda^R - \sum_{m=1}^M \lambda_m^C)$	87.6	43.3	18.8	37.6	45.6	46.3	51.8	45.6
Deg. of Freedom	12	12	12	12	12	12	12	12
χ^2 test p-value	**0.000**	**0.000**	0.094	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
Test for Index Shift								
$-2 \cdot (\sum_m \lambda_m^C - \sum_{m=1}^M \lambda_m^U)$	120.8	91.3	42.9	33.6	63.1	84.0	51.7	49.1
Deg. of Freedom	39	33	33	33	36	36	33	33
χ^2 test p-value	**0.000**	**0.000**	0.115	0.437	**0.003**	**0.000**	**0.020**	**0.035**

Notes: Tests are for reporting bias by income using income quartiles. Result from each column (except "Full") are from regressions using only the subsample of our data indicated by the column title. Likelihood values taken from three sets of ordered probit regressions as follows: Likelihood values for the restricted models (λ^R) come from a single ordered probit regression with dummy variables to index income quartiles. The likelihood value for the sum of the unrestricted models ($\sum_m \lambda_m^U$) comes from running a separate ordered probit model for each income quartile. The likelihood value for the sum of the constrained models ($\sum_m \lambda_m^C$) comes from running a separate ordered probit model for each income quartile while constraining the coefficient estimates to be equal to those of the single, restricted model and allowing the cut-points to vary. Degrees of freedom are determined by the number of restricted parameters in the likelihood ratio test. All regressions use robust standard errors.

Table A4: Likelihood Ratio Tests for Reporting Bias, No Education Controls

	SES		Demographic		
	Income	Education	Race	Sex	Age
λ^R	-15333.4	-15363.9	-15363.9	-15363.9	-15380.8
Test for Reporting Bias					
$\sum_m \lambda_m^U$	-15168.8	-15105.4	-15234.0	-15334.1	-15324.9
$-2 \cdot (\lambda^R - \sum_{m=1}^M \lambda_m^U)$	329.3	223.1	259.8	59.6	78.0
Deg. of Freedom	140	39	30	16	16
χ^2 test p-value	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
Test for Cut-Point Shift					
$\sum_m \lambda_m^C$	-15283.7	-15148.2	-15273.6	-15358.8	-15361.3
$-2 \cdot (\lambda^R - \sum_{m=1}^M \lambda_m^C)$	99.5	431.5	180.7	10.2	39.0
Deg. of Freedom	40	12	8	4	4
χ^2 test p-value	**0.000**	**0.000**	**0.000**	**0.037**	**0.000**
Test for Index Shift					
$-2 \cdot (\sum_m \lambda_m^C - \sum_{m=1}^M \lambda_m^U)$	229.8	85.6	79.1	49.4	72.7
Deg. of Freedom	100	27	22	12	12
χ^2 test p-value	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**

Notes: Tests are for reporting bias by the SES or demographic category in the column title. Likelihood values taken from three sets of ordered probit regressions as follows: Likelihood values for the restricted models (λ^R) come from a single ordered probit regression with dummy variables to index SES and demographic categories. There are 11 subgroups of income (all models other than "Income" control for income quartiles), 4 subgroups of education, 3 subgroups of race, 2 subgroups for sex, and 2 subgroups for age. The likelihood value for the sum of the unrestricted models ($\sum_m \lambda_m^U$) comes from running a separate ordered probit model for each subgroup. The likelihood value for the sum of the constrained models ($\sum_m \lambda_m^C$) comes from running a separate ordered probit model for each subgroup while constraining the coefficient estimates to be equal to those of the single, restricted model and allowing the cut-points to vary. Degrees of freedom are determined by the number of restricted parameters in the likelihood ratio test. All regressions use robust standard errors.

Table A5: Likelihood Ratio Tests for Income-Based Reporting Bias by Sub-Population, No Education Controls

		Race			Sex		Age	
	Full	White	Black	Hisp.	Male	Female	≤ 45	> 45
λ^R	-15363.9	-8489.5	-2823.9	-3920.7	-7996.5	-7337.7	-6425.8	-8899.1
Test for Reporting Bias								
$\sum_m \lambda_m^U$	-15245.6	-8414.4	-2799.0	-3883.4	-7935.5	-7263.5	-6369.5	-8849.0
$-2 \cdot (\lambda^R - \sum_{m=1}^M \lambda_m^U)$	236.6	150.1	49.8	74.5	121.9	148.3	112.6	100.3
Deg. of Freedom	42	36	36	36	39	39	36	36
χ^2 test p-value	**0.000**	**0.000**	0.063	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
Test for Cut-Point Shift								
$\sum_m \lambda_m^C$	-15319.7	-8467.5	-2814.3	-3901.0	-7973.0	-7314.9	-6399.2	-8876.7
$-2 \cdot (\lambda^R - \sum_{m=1}^M \lambda_m^C)$	88.5	44.0	19.2	39.5	47.0	45.5	53.2	44.9
Deg. of Freedom	12	12	12	12	12	12	12	12
χ^2 test p-value	**0.000**	**0.000**	0.085	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
Test for Index Shift								
$-2 \cdot (\sum_m \lambda_m^C - \sum_{m=1}^M \lambda_m^U)$	148.0	106.1	30.6	35.0	74.9	102.9	59.4	55.4
Deg. of Freedom	30	24	24	24	27	27	24	24
χ^2 test p-value	**0.000**	**0.000**	0.165	0.068	**0.000**	**0.000**	**0.000**	**0.000**

Notes: Tests are for reporting bias by income using income quartiles. Result from each column (except "Full") are from regressions using only the subsample of our data indicated by the column title. Likelihood values taken from three sets of ordered probit regressions as follows: Likelihood values for the restricted models (λ^R) come from a single ordered probit regression with dummy variables to index income quartiles. The likelihood value for the sum of the unrestricted models ($\sum_m \lambda_m^U$) comes from running a separate ordered probit model for each income quartile. The likelihood value for the sum of the constrained models ($\sum_m \lambda_m^C$) comes from running a separate ordered probit model for each income quartile while constraining the coefficient estimates to be equal to those of the single, restricted model and allowing the cut-points to vary. Degrees of freedom are determined by the number of restricted parameters in the likelihood ratio test. All regressions use robust standard errors.

Table A6: Likelihood Ratio Tests for Reporting Bias

	SES		Demographic		
	Income	Education	Race	Sex	Age
Body Mass Index					
Test for Reporting Bias (p-value)	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
Test for Cut-Point Shift (p-value)	**0.020**	**0.000**	**0.000**	**0.004**	**0.000**
Test for Index Shift (p-value)	**0.000**	**0.000**	**0.000**	**0.004**	**0.000**
Cholesterol					
Test for Reporting Bias (p-value)	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
Test for Cut-Point Shift (p-value)	**0.038**	**0.000**	**0.000**	**0.041**	**0.000**
Test for Index Shift (p-value)	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
Triglycerides					
Test for Reporting Bias (p-value)	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
Test for Cut-Point Shift (p-value)	**0.030**	**0.000**	**0.000**	**0.040**	**0.000**
Test for Index Shift (p-value)	**0.000**	**0.000**	**0.000**	**0.000**	**0.002**
Glycated Hemoglobin					
Test for Reporting Bias (p-value)	**0.000**	**0.000**	**0.000**	**0.000**	**0.001**
Test for Cut-Point Shift (p-value)	**0.028**	**0.000**	**0.000**	**0.044**	**0.000**
Test for Index Shift (p-value)	**0.000**	**0.000**	**0.000**	**0.000**	**0.033**
Glomerular Filtration Rate					
Test for Reporting Bias (p-value)	**0.000**	**0.000**	**0.000**	**0.000**	**0.011**
Test for Cut-Point Shift (p-value)	**0.038**	**0.000**	**0.000**	**0.041**	**0.000**
Test for Index Shift (p-value)	**0.000**	**0.000**	**0.000**	**0.000**	**0.011**
C-Reactive Protein					
Test for Reporting Bias (p-value)	**0.000**	**0.000**	**0.000**	**0.000**	**0.002**
Test for Cut-Point Shift (p-value)	**0.039**	**0.000**	**0.000**	**0.033**	**0.000**
Test for Index Shift (p-value)	**0.000**	**0.000**	**0.000**	**0.000**	**0.003**
White Blood Cell Count					
Test for Reporting Bias (p-value)	**0.000**	**0.000**	**0.000**	**0.000**	**0.001**
Test for Cut-Point Shift (p-value)	**0.033**	**0.000**	**0.000**	**0.036**	**0.000**
Test for Index Shift (p-value)	**0.000**	**0.000**	**0.000**	**0.000**	**0.007**
Systolic Blood Pressure					
Test for Reporting Bias (p-value)	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
Test for Cut-Point Shift (p-value)	**0.034**	**0.000**	**0.000**	**0.032**	**0.000**
Test for Index Shift (p-value)	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
Diastolic Blood Pressure					
Test for Reporting Bias (p-value)	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
Test for Cut-Point Shift (p-value)	**0.037**	**0.000**	**0.000**	**0.040**	**0.000**
Test for Index Shift (p-value)	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**

Notes: Tests are for reporting bias by the SES or demographic category in the column title. Likelihood values taken from three sets of ordered probit regressions as follows: Likelihood values for the restricted models (λ^R) come from a single ordered probit regression with dummy variables to index SES and demographic categories. There are 11 subgroups of income (all models other than "Income" control for income quartiles), 4 subgroups of education, 3 subgroups of race, 2 subgroups for sex, and 2 subgroups for age. The likelihood value for the sum of the unrestricted models ($\sum_m \lambda_m^U$) comes from running a separate ordered probit model for each subgroup. The likelihood value for the sum of the constrained models ($\sum_m \lambda_m^C$) comes from running a separate ordered probit model for each subgroup while constraining the coefficient estimates to be equal to those of the single, restricted model and allowing the cut-points to vary. Degrees of freedom are determined by the number of restricted parameters in the likelihood ratio test. All regressions use robust standard errors.

Table A7: Likelihood Ratio Tests for Income-Based Reporting Bias by Sub-Population

		Race			Sex		Age	
	Full	White	Black	Hisp.	Male	Female	≤ 45	> 45
Body Mass Index								
Test for Reporting Bias (p-value)	**0.000**	**0.000**	**0.013**	**0.003**	**0.000**	**0.000**	**0.000**	**0.004**
Test for Cut-Point Shift (p-value)	**0.000**	**0.000**	0.274	**0.021**	**0.002**	**0.007**	**0.000**	**0.036**
Test for Index Shift (p-value)	**0.000**	**0.000**	**0.011**	**0.018**	**0.000**	**0.000**	**0.000**	**0.021**
Cholesterol								
Test for Reporting Bias (p-value)	**0.000**	**0.000**	0.143	0.127	**0.004**	**0.000**	**0.001**	0.050
Test for Cut-Point Shift (p-value)	**0.000**	**0.000**	0.281	**0.031**	**0.004**	**0.017**	**0.000**	**0.046**
Test for Index Shift (p-value)	**0.000**	**0.000**	0.164	0.451	0.073	**0.000**	0.089	0.177
Triglycerides								
Test for Reporting Bias (p-value)	**0.000**	**0.000**	0.133	**0.049**	**0.000**	**0.000**	**0.000**	**0.017**
Test for Cut-Point Shift (p-value)	**0.000**	**0.000**	0.295	**0.032**	**0.004**	**0.012**	**0.000**	**0.046**
Test for Index Shift (p-value)	**0.000**	**0.000**	0.146	0.209	**0.002**	**0.000**	**0.001**	0.063
Glycated Hemoglobin								
Test for Reporting Bias (p-value)	**0.000**	**0.000**	0.136	0.142	**0.002**	**0.000**	**0.001**	**0.002**
Test for Cut-Point Shift (p-value)	**0.000**	**0.000**	0.258	**0.027**	**0.003**	**0.016**	**0.000**	**0.040**
Test for Index Shift (p-value)	**0.000**	**0.000**	0.165	0.515	**0.044**	**0.000**	0.108	**0.008**
Glomerular Filtration Rate								
Test for Reporting Bias (p-value)	**0.000**	**0.000**	0.122	**0.034**	**0.000**	**0.000**	**0.001**	**0.042**
Test for Cut-Point Shift (p-value)	**0.000**	**0.000**	0.289	**0.032**	**0.004**	**0.017**	**0.000**	**0.050**
Test for Index Shift (p-value)	**0.000**	**0.000**	0.135	0.153	**0.004**	**0.000**	0.062	0.145
C-Reactive Protein								
Test for Reporting Bias (p-value)	**0.000**	**0.000**	**0.029**	0.061	**0.001**	**0.000**	**0.000**	**0.028**
Test for Cut-Point Shift (p-value)	**0.000**	**0.000**	0.291	**0.029**	**0.004**	**0.017**	**0.000**	**0.049**
Test for Index Shift (p-value)	**0.000**	**0.000**	**0.026**	0.264	**0.017**	**0.000**	**0.006**	0.100
White Blood Cell Count								
Test for Reporting Bias (p-value)	**0.000**	**0.000**	**0.039**	**0.021**	**0.006**	**0.000**	**0.000**	0.058
Test for Cut-Point Shift (p-value)	**0.000**	**0.000**	0.263	**0.031**	**0.004**	**0.012**	**0.000**	**0.046**
Test for Index Shift (p-value)	**0.000**	**0.000**	**0.041**	0.100	0.098	**0.000**	**0.007**	0.202
Systolic Blood Pressure								
Test for Reporting Bias (p-value)	**0.000**	**0.000**	0.055	0.099	**0.000**	**0.000**	**0.000**	**0.011**
Test for Cut-Point Shift (p-value)	**0.000**	**0.000**	0.272	**0.029**	**0.004**	**0.016**	**0.000**	**0.049**
Test for Index Shift (p-value)	**0.000**	**0.000**	0.058	0.385	**0.009**	**0.000**	**0.041**	**0.040**
Diastolic Blood Pressure								
Test for Reporting Bias (p-value)	**0.000**	**0.000**	0.137	0.120	**0.006**	**0.000**	**0.001**	**0.029**
Test for Cut-Point Shift (p-value)	**0.000**	**0.000**	0.279	**0.031**	**0.004**	**0.017**	**0.000**	**0.048**
Test for Index Shift (p-value)	**0.000**	**0.000**	0.157	0.434	0.105	**0.000**	0.096	0.103

Notes: Tests are for reporting bias by income using income quartiles. Result from each column (except "Full") are from regressions using only the subsample of our data indicated by the column title. Likelihood values taken from three sets of ordered probit regressions as follows: Likelihood values for the restricted models (λ^R) come from a single ordered probit regression with dummy variables to index income quartiles. The likelihood value for the sum of the unrestricted models ($\sum_m \lambda_m^U$) comes from running a separate ordered probit model for each income quartile. The likelihood value for the sum of the constrained models ($\sum_m \lambda_m^C$) comes from running a separate ordered probit model for each income quartile while constraining the coefficient estimates to be equal to those of the single, restricted model and allowing the cut-points to vary. Degrees of freedom are determined by the number of restricted parameters in the likelihood ratio test. All regressions use robust standard errors.

Table A8: Regression Results of Self-Assessed Health on Income and Count Allostasis

	Self-Assessed Health	Physically Healthy Days	Mentally Healthy Days
Allostasis (Count)	0.185 (0.013) ***	0.356 (0.083) ***	0.249 (0.083) ***
Allostasis (Count) x IPR	0.044 (0.008) ***	0.018 (0.050)	-0.035 (0.048)
Income-to-Poverty Ratio	0.100 (0.038) ***	0.410 (0.196) **	0.865 (0.197) ***
Income-to-Poverty Ratio Squared	-0.006 (0.009)	-0.227 (0.049) ***	-0.136 (0.052) ***
Age/100	-0.240 (0.107) **	-3.793 (0.585) ***	4.772 (0.586) ***
Age/100 x IPR	0.195 (0.072) ***	0.432 (0.377)	0.374 (0.351)
Age/100 Squared	2.104 (0.548) ***	14.764 (3.149) ***	25.008 (3.155) ***
Age/100 Squared x IPR	-1.240 (0.371) ***	-8.260 (2.011) ***	-7.601 (1.826) ***
Female	-0.132 (0.059) **	-1.226 (0.349) ***	-2.870 (0.367) ***
Female x IPR	0.009 (0.020)	0.146 (0.101)	0.354 (0.102) ***
Black	-0.031 (0.070)	0.817 (0.442) *	1.313 (0.483) ***
Black x IPR	-0.074 (0.023) ***	-0.189 (0.134)	-0.338 (0.152) **
Hispanic	-0.215 (0.067) ***	1.572 (0.390) ***	3.130 (0.440) ***
Hispanic x IPR	0.001 (0.024)	-0.566 (0.136) ***	-0.831 (0.148) ***
Married	0.008 (0.092)	-0.268 (0.444)	0.650 (0.524)
Married x IPR	-0.004 (0.033)	0.138 (0.143)	-0.106 (0.157)
Widowed	0.029 (0.130)	-0.252 (1.026)	1.448 (0.865) *
Widowed x IPR	-0.024 (0.050)	-0.063 (0.357)	-0.603 (0.278) **
Divorced	-0.061 (0.112)	-0.291 (0.643)	-0.648 (0.743)
Divorced x IPR	0.014 (0.042)	-0.162 (0.225)	0.003 (0.226)
College Degree	0.703 (0.127) ***	1.001 (0.587) *	2.455 (0.614) ***
College Degree x IPR	-0.005 (0.038)	0.089 (0.188)	-0.391 (0.183) **
Some College	0.285 (0.080) ***	-0.331 (0.482)	0.271 (0.521)
Some College x IPR	0.014 (0.031)	0.262 (0.176)	-0.094 (0.174)
High School Diploma	0.204 (0.072) ***	0.156 (0.497)	0.929 (0.503) *
High School Diploma x IPR	-0.003 (0.031)	0.087 (0.189)	-0.345 (0.181) *
Adj. R^2	0.111	0.042	0.059
Num Obs	11751	11751	11751
P-value for Chow test	0.000	0.000	0.000

Notes: Data from the 2005-2006, 2007-2008 and 2009-2010 waves of NHANES including adults over age 20. Each column shows coefficients from a regression using different measures of self-assessed health. In addition to the coefficients shown, all models include survey wave fixed effects. All regressions include NHANES sample weights. Robust standard errors are shown in parenthesis, and stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%

Table A9: Regression Results of Self-Assessed Health on Income and Individual Biomarkers

	Body Mass Index	Choles-terol	Trigly-cerides	Glycohe-moglobin	Creat-inine	C-Reactive Protein	White Blood Cell Count	Systolic Blood Pressure	Diastolic Blood Pressure
Biomarker	0.265***	-0.016	0.144***	0.192***	-0.025	0.115***	0.138***	0.105***	0.053***
	(0.014)	(0.015)	(0.015)	(0.015)	(0.023)	(0.020)	(0.021)	(0.017)	(0.016)
Biomarker x IPR	0.063***	-0.007	0.054***	0.053***	-0.032**	0.033***	0.025	0.040***	0.020*
	(0.009)	(0.010)	(0.009)	(0.010)	(0.015)	(0.011)	(0.016)	(0.011)	(0.011)
Income-to-Poverty Ratio	0.128***	0.095**	0.097**	0.090**	0.092**	0.091**	0.095**	0.101***	0.099**
	(0.038)	(0.039)	(0.039)	(0.039)	(0.039)	(0.039)	(0.039)	(0.039)	(0.039)
Income-to-Poverty Ratio Squared	-0.007	-0.004	-0.005	-0.005	-0.005	-0.003	-0.004	-0.005	-0.003
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Age/100	-0.505***	-0.648***	-0.586***	-0.324***	-0.729***	-0.601***	-0.702***	-0.410***	-0.654***
	(0.103)	(0.105)	(0.104)	(0.108)	(0.143)	(0.105)	(0.105)	(0.112)	(0.105)
Age/100 x IPR	0.181***	0.118*	0.134*	0.212***	0.003	0.131*	0.122*	0.200***	0.111
	(0.070)	(0.071)	(0.071)	(0.074)	(0.093)	(0.071)	(0.072)	(0.076)	(0.072)
Age/100 Squared	1.243**	2.843***	2.104***	2.299***	2.699***	2.586***	2.621***	2.839***	2.151***
	(0.553)	(0.562)	(0.553)	(0.555)	(0.554)	(0.551)	(0.550)	(0.550)	(0.578)
Age/100 Squared x IPR	-1.584***	-1.132***	-1.396***	-1.205***	-1.203***	-1.190***	-1.173***	-1.111***	-1.396***
	(0.373)	(0.380)	(0.375)	(0.376)	(0.376)	(0.374)	(0.375)	(0.374)	(0.393)
Female	-0.085	-0.148**	-0.152**	-0.148**	-0.152**	-0.149**	-0.151**	-0.143**	-0.144**
	(0.060)	(0.060)	(0.060)	(0.060)	(0.060)	(0.061)	(0.060)	(0.060)	(0.061)
Female x IPR	0.005	0.032	0.017	0.028	0.034*	0.038*	0.033*	0.021	0.026
	(0.019)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
Black	0.009	-0.023	-0.036	-0.011	-0.039	-0.019	-0.057	-0.019	-0.019
	(0.070)	(0.071)	(0.072)	(0.071)	(0.072)	(0.071)	(0.074)	(0.071)	(0.071)
Black x IPR	-0.063***	-0.081***	-0.097***	-0.066***	-0.074***	-0.079***	-0.089***	-0.075***	-0.082***
	(0.023)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.025)	(0.024)	(0.024)
Hispanic	-0.200***	-0.211***	-0.221***	-0.210***	-0.224***	-0.219***	-0.224***	-0.207***	-0.211***
	(0.066)	(0.068)	(0.068)	(0.068)	(0.069)	(0.068)	(0.068)	(0.068)	(0.068)
Hispanic x IPR	-0.002	-0.010	-0.003	0.006	-0.004	-0.004	-0.008	-0.011	-0.010
	(0.024)	(0.024)	(0.024)	(0.025)	(0.025)	(0.024)	(0.024)	(0.024)	(0.024)
Married	0.035	-0.020	-0.010	-0.008	-0.018	-0.012	-0.006	-0.015	-0.017
	(0.091)	(0.093)	(0.092)	(0.092)	(0.093)	(0.092)	(0.092)	(0.092)	(0.092)
Married x IPR	-0.020	0.006	0.004	-0.002	0.006	0.002	0.001	0.001	0.005
	(0.032)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)
Widowed	0.025	0.003	0.004	0.007	0.009	0.006	0.021	0.022	0.008
	(0.128)	(0.131)	(0.131)	(0.131)	(0.131)	(0.131)	(0.131)	(0.131)	(0.131)
Widowed x IPR	-0.038	-0.025	-0.019	-0.027	-0.028	-0.026	-0.029	-0.035	-0.027
	(0.049)	(0.050)	(0.050)	(0.050)	(0.050)	(0.050)	(0.050)	(0.050)	(0.050)
Divorced	-0.039	-0.059	-0.064	-0.053	-0.060	-0.059	-0.032	-0.060	-0.062
	(0.111)	(0.114)	(0.113)	(0.114)	(0.114)	(0.114)	(0.114)	(0.114)	(0.114)
Divorced x IPR	-0.009	0.010	0.015	0.005	0.012	0.013	0.010	0.009	0.012
	(0.042)	(0.044)	(0.043)	(0.043)	(0.044)	(0.043)	(0.043)	(0.043)	(0.043)
College Degree	0.760***	0.719***	0.725***	0.695***	0.723***	0.718***	0.699***	0.727***	0.725***
	(0.125)	(0.129)	(0.128)	(0.128)	(0.129)	(0.128)	(0.128)	(0.129)	(0.129)
College Degree x IPR	-0.025	0.008	-0.002	0.004	0.005	0.004	0.003	-0.001	0.005
	(0.037)	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)
Some College	0.316***	0.301***	0.284***	0.278***	0.308***	0.310***	0.291***	0.295***	0.301***
	(0.079)	(0.080)	(0.080)	(0.079)	(0.081)	(0.080)	(0.080)	(0.081)	(0.080)
Some College x IPR	0.007	0.008	0.016	0.014	0.005	0.005	0.007	0.011	0.010
	(0.030)	(0.031)	(0.031)	(0.030)	(0.031)	(0.031)	(0.030)	(0.031)	(0.031)
High School Diploma	0.213***	0.204***	0.201***	0.192***	0.210***	0.202***	0.203***	0.202***	0.208***
	(0.072)	(0.073)	(0.073)	(0.072)	(0.073)	(0.072)	(0.072)	(0.073)	(0.073)
High School Diploma x IPR	-0.003	-0.002	-0.002	-0.001	-0.006	-0.000	-0.003	-0.001	-0.004
	(0.030)	(0.031)	(0.031)	(0.030)	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)
Adj. R^2	0.125	0.086	0.099	0.099	0.087	0.094	0.095	0.092	0.088
Num Obs	11751	11751	11751	11751	11751	11751	11751	11751	11751
P-value for Chow test	0.000	0.002	0.000	0.000	0.001	0.000	0.002	0.000	0.001

Notes: Data from the 2005-2006, 2007-2008 and 2009-2010 waves of NHANES including adults over age 20. Each regression shows coefficients from a regression using self-assessed health as the dependent variable and the biomarker listed in the column heading as the biological level of health. In addition to the coefficients shown, all models include survey wave fixed effects. Robust standard errors are shown in parenthesis. All models include NHANES sample weights. Stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%

Table A10: Regression Results of Self-Reported Health on Income and Allostasis
No Sample Weights

	Self-Assessed Health	Physically Healthy Days	Mentally Healthy Days
Allostasis	0.188 (0.013) ***	0.272 (0.084) ***	0.160 (0.079) **
Allostasis x IPR	0.048 (0.008) ***	-0.022 (0.053)	-0.025 (0.049)
Income-to-Poverty Ratio	0.109 (0.031) ***	0.584 (0.183) ***	0.998 (0.183) ***
Income-to-Poverty Ratio Squared	-0.003 (0.007)	-0.185 (0.043) ***	-0.165 (0.043) ***
Age/100	-0.377 (0.090) ***	-4.321 (0.505) ***	4.642 (0.484) ***
Age/100 x IPR	0.147 (0.059) **	0.698 (0.332) **	0.550 (0.312) *
Age/100 Squared	1.943 (0.465) ***	18.389 (2.750) ***	22.580 (2.550) ***
Age/100 Squared x IPR	-1.306 (0.312) ***	-8.568 (1.765) ***	-8.723 (1.599) ***
Female	-0.135 (0.045) ***	-1.232 (0.314) ***	-2.753 (0.311) ***
Female x IPR	0.005 (0.016)	0.176 (0.092) *	0.349 (0.091) ***
Black	0.010 (0.062)	1.118 (0.420) ***	1.706 (0.434) ***
Black x IPR	-0.079 (0.022) ***	-0.266 (0.127) **	-0.436 (0.134) ***
Hispanic	-0.171 (0.052) ***	1.442 (0.379) ***	3.302 (0.380) ***
Hispanic x IPR	-0.016 (0.021)	-0.552 (0.128) ***	-0.839 (0.124) ***
Married	0.026 (0.070)	0.197 (0.413)	0.914 (0.470) *
Married x IPR	0.007 (0.026)	-0.004 (0.134)	-0.188 (0.151)
Widowed	0.005 (0.101)	0.654 (0.786)	1.534 (0.725) **
Widowed x IPR	0.015 (0.042)	-0.365 (0.288)	-0.664 (0.250) ***
Divorced	-0.022 (0.084)	-0.426 (0.575)	-0.072 (0.612)
Divorced x IPR	0.015 (0.033)	-0.039 (0.190)	-0.078 (0.198)
College Degree	0.670 (0.101) ***	1.318 (0.544) **	2.322 (0.521) ***
College Degree x IPR	-0.001 (0.031)	0.061 (0.174)	-0.385 (0.161) **
Some College	0.357 (0.064) ***	0.226 (0.435)	0.158 (0.446)
Some College x IPR	-0.005 (0.025)	0.167 (0.161)	-0.046 (0.153)
High School Diploma	0.165 (0.057) ***	0.705 (0.432)	0.827 (0.423) *
High School Diploma x IPR	0.015 (0.025)	-0.011 (0.170)	-0.231 (0.155)
Adj. R^2	0.117	0.042	0.062
Num Obs	11751	11751	11751
P-value for Chow test	0.000	0.000	0.000

Notes: Data from the 2005-2006, 2007-2008 and 2009-2010 waves of NHANES including adults over age 20. Each column shows coefficients from a regression using different measures of self-assessed health. In addition to the coefficients shown, all models include survey wave fixed effects. All regressions include NHANES sample weights. Robust standard errors are shown in parenthesis, and stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%

Table A11: Regression Results of Self-Reported Health on Income and Allostasis
Alternate Measures of Self-Assessed Health

	Self-Assessed Health	Self-Assessed Health (HUQ)	Self-Assessed Health (0/1)	HUQ Self-Assessed Health (0/1)
Allostasis	0.217 (0.015) ***	0.187 (0.014) ***	0.048 (0.005) ***	0.041 (0.005) ***
Allostasis x IPR	0.052 (0.010) ***	0.038 (0.009) ***	-0.002 (0.003)	-0.004 (0.003)
Income-to-Poverty Ratio	0.098 (0.038) **	0.126 (0.037) ***	0.049 (0.010) ***	0.058 (0.010) ***
Income-to-Poverty Ratio Squared	-0.005 (0.009)	-0.025 (0.008) ***	-0.010 (0.002) ***	-0.012 (0.002) ***
Age/100	-0.243 (0.108) **	-0.594 (0.100) ***	-0.167 (0.027) ***	-0.224 (0.027) ***
Age/100 x IPR	0.188 (0.072) ***	0.178 (0.066) ***	0.067 (0.016) ***	0.065 (0.016) ***
Age/100 Squared	2.161 (0.549) ***	2.417 (0.530) ***	0.700 (0.145) ***	0.715 (0.145) ***
Age/100 Squared x IPR	-1.226 (0.371) ***	-1.112 (0.340) ***	-0.398 (0.087) ***	-0.578 (0.088) ***
Female	-0.136 (0.059) **	-0.108 (0.056) *	-0.035 (0.017) **	-0.018 (0.017)
Female x IPR	0.010 (0.020)	0.006 (0.018)	0.008 (0.004) *	0.003 (0.005)
Black	-0.024 (0.071)	0.138 (0.067) **	-0.046 (0.022) **	-0.040 (0.022) *
Black x IPR	-0.074 (0.023) ***	-0.104 (0.022) ***	-0.005 (0.007)	-0.004 (0.007)
Hispanic	-0.211 (0.067) ***	-0.099 (0.063)	-0.104 (0.021) ***	-0.106 (0.021) ***
Hispanic x IPR	-0.000 (0.024)	-0.042 (0.022) *	0.011 (0.006) *	0.015 (0.006) **
Married	-0.001 (0.092)	0.079 (0.080)	0.020 (0.022)	0.030 (0.022)
Married x IPR	-0.002 (0.033)	-0.006 (0.029)	0.001 (0.007)	-0.004 (0.007)
Widowed	0.024 (0.130)	0.040 (0.120)	0.044 (0.042)	0.041 (0.041)
Widowed x IPR	-0.026 (0.050)	0.010 (0.049)	-0.013 (0.013)	-0.010 (0.013)
Divorced	-0.066 (0.113)	0.085 (0.101)	-0.009 (0.032)	0.035 (0.031)
Divorced x IPR	0.015 (0.042)	-0.013 (0.037)	0.001 (0.010)	-0.010 (0.009)
College Degree	0.706 (0.127) ***	0.551 (0.110) ***	0.191 (0.029) ***	0.225 (0.028) ***
College Degree x IPR	-0.003 (0.038)	0.024 (0.036)	-0.015 (0.009)	-0.022 (0.009) **
Some College	0.285 (0.080) ***	0.295 (0.078) ***	0.143 (0.024) ***	0.126 (0.024) ***
Some College x IPR	0.015 (0.031)	0.020 (0.031)	-0.010 (0.009)	-0.005 (0.009)
High School Diploma	0.206 (0.072) ***	0.205 (0.071) ***	0.126 (0.025) ***	0.122 (0.025) ***
High School Diploma x IPR	-0.004 (0.031)	-0.010 (0.031)	-0.018 (0.009) *	-0.016 (0.009) *
Adj. R^2	0.109	0.114	0.126	0.129
Num Obs	11751	11751	11751	11751
P-value for Chow test	0.000	0.000	0.000	0.000

Notes: Data from the 2005-2006, 2007-2008 and 2009-2010 waves of NHANES including adults over age 20. Each column shows coefficients from a regression using different measures of self-assessed health. In addition to the coefficients shown, all models include survey wave fixed effects. All regressions include NHANES sample weights. Robust standard errors are shown in parenthesis, and stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%

Table A12: Regression Results of Self-Reported Health on Income and Allostasis Subsets of NHANES Data

	Men	Women	Non-Blacks	Blacks	Age≤45	Age>45	BA Degree	Less than BA Degree	Inc-Poverty Ratio<5	NHANES 2005-2006	NHANES 2007-2008	NHANES 2009-2010
Allostasis	0.172***	0.275***	0.220***	0.177***	0.263***	0.193***	0.388***	0.172***	0.177***	0.194***	0.267***	0.184***
	(0.021)	(0.021)	(0.016)	(0.027)	(0.026)	(0.018)	(0.040)	(0.016)	(0.016)	(0.028)	(0.025)	(0.026)
Allostasis x IPR	0.054***	0.055***	0.054***	0.026	0.057***	0.050***	0.041	0.035***	0.027**	0.028	0.068***	0.059***
	(0.014)	(0.014)	(0.011)	(0.018)	(0.016)	(0.013)	(0.029)	(0.011)	(0.013)	(0.018)	(0.017)	(0.016)
Income-to-Poverty Ratio	0.145***	0.060	0.103**	0.010	0.104**	0.119*	0.151*	0.076*	0.119**	0.093	0.155**	0.016
	(0.053)	(0.053)	(0.042)	(0.066)	(0.053)	(0.070)	(0.086)	(0.042)	(0.048)	(0.077)	(0.066)	(0.060)
Income-Poverty Ratio Squared	-0.002	-0.009	-0.007	0.012	-0.004	0.000	0.042*	-0.016*	-0.016	-0.028*	-0.017	0.025*
	(0.013)	(0.012)	(0.010)	(0.016)	(0.012)	(0.012)	(0.025)	(0.009)	(0.012)	(0.016)	(0.015)	(0.014)
Age/100	-0.540***	0.154	-0.180	-0.981***	-2.762*	0.747	0.135	-0.368***	-0.410***	-0.158	-0.491***	-0.108
	(0.149)	(0.155)	(0.116)	(0.218)	(1.569)	(0.511)	(0.261)	(0.117)	(0.116)	(0.199)	(0.187)	(0.172)
intage	0.238**	0.180*	0.187**	0.071	0.259	-0.114	0.256	0.126	0.092	0.148	0.168	0.217*
	(0.098)	(0.107)	(0.077)	(0.147)	(1.004)	(0.334)	(0.193)	(0.084)	(0.094)	(0.132)	(0.123)	(0.122)
Age/100 Squared	2.320***	2.033***	2.243***	0.399	-5.523	-0.874	-0.040	2.908***	2.885***	2.325**	1.306	2.591***
	(0.791)	(0.754)	(0.595)	(1.174)	(4.607)	(1.719)	(1.525)	(0.573)	(0.572)	(0.950)	(1.010)	(0.897)
intage2	-1.256**	-1.101**	-1.267***	-1.152	-0.741	-0.585	-0.386	-1.153***	-1.044**	-1.072*	-1.682**	-0.874
	(0.528)	(0.520)	(0.398)	(0.851)	(2.959)	(1.187)	(1.056)	(0.416)	(0.476)	(0.629)	(0.657)	(0.642)
Black	0.110	-0.126			0.044	-0.119	-0.433*	0.007	0.027	-0.085	0.056	-0.039
	(0.111)	(0.091)			(0.109)	(0.082)	(0.252)	(0.074)	(0.076)	(0.127)	(0.123)	(0.120)
intblack	-0.071**	-0.082***			-0.083**	-0.061**	0.000	-0.070***	-0.104***	-0.050	-0.097**	-0.080*
	(0.036)	(0.030)			(0.036)	(0.029)	(0.065)	(0.026)	(0.029)	(0.040)	(0.041)	(0.041)
Hispanic	-0.233**	-0.189**	-0.189***		-0.270***	-0.096	-0.325	-0.178**	-0.213***	-0.271*	-0.169*	-0.241**
	(0.097)	(0.093)	(0.069)		(0.085)	(0.111)	(0.259)	(0.071)	(0.074)	(0.163)	(0.101)	(0.103)
inthisp	0.014	-0.018	-0.005		-0.002	-0.007	0.023	-0.010	0.008	0.022	-0.038	0.032
	(0.036)	(0.032)	(0.024)		(0.031)	(0.038)	(0.066)	(0.027)	(0.032)	(0.053)	(0.035)	(0.039)
Married	0.181	-0.217*	-0.029	0.079	-0.021	0.007	-0.017	-0.046	-0.018	-0.116	-0.031	0.105
	(0.128)	(0.131)	(0.107)	(0.153)	(0.114)	(0.161)	(0.309)	(0.096)	(0.098)	(0.178)	(0.157)	(0.147)
intmarried	-0.050	0.056	-0.001	0.029	0.011	-0.025	-0.033	0.038	0.011	-0.005	-0.017	0.026
	(0.045)	(0.047)	(0.037)	(0.055)	(0.040)	(0.064)	(0.080)	(0.038)	(0.042)	(0.061)	(0.058)	(0.053)
Widowed	-0.046	-0.156	0.008	-0.035	-0.494	-0.013	-0.324	0.021	0.003	-0.177	-0.124	0.302
	(0.202)	(0.171)	(0.151)	(0.204)	(0.401)	(0.181)	(0.571)	(0.130)	(0.140)	(0.248)	(0.224)	(0.211)
intwidowed	-0.052	0.022	-0.038	0.161*	0.055	-0.032	0.027	-0.007	-0.010	-0.002	0.029	-0.083
	(0.075)	(0.067)	(0.055)	(0.091)	(0.097)	(0.073)	(0.155)	(0.052)	(0.061)	(0.086)	(0.097)	(0.075)
Divorced	0.036	-0.198	-0.083	0.010	-0.037	-0.076	-0.139	-0.089	-0.058	-0.092	-0.198	0.056
	(0.166)	(0.153)	(0.132)	(0.181)	(0.157)	(0.176)	(0.399)	(0.118)	(0.120)	(0.220)	(0.189)	(0.184)
intdivorced	-0.007	0.045	0.016	0.021	-0.026	0.019	0.000	0.047	0.012	-0.049	0.040	0.061
	(0.062)	(0.058)	(0.048)	(0.063)	(0.059)	(0.071)	(0.106)	(0.048)	(0.053)	(0.076)	(0.075)	(0.070)
College Degree	0.837***	0.579***	0.760***	0.343	0.841***	0.569***			0.967***	0.440*	0.804***	0.813***
	(0.184)	(0.176)	(0.140)	(0.243)	(0.178)	(0.180)			(0.141)	(0.259)	(0.229)	(0.186)
intedba	-0.030	0.019	-0.010	0.022	-0.068	0.045			-0.130***	0.091	-0.066	-0.015
	(0.054)	(0.052)	(0.041)	(0.074)	(0.058)	(0.051)			(0.050)	(0.074)	(0.067)	(0.058)
Some College	0.347***	0.240**	0.277***	0.415**	0.342***	0.305***		0.272***	0.341***	0.270	0.224*	0.357***
	(0.116)	(0.110)	(0.089)	(0.165)	(0.111)	(0.116)		(0.080)	(0.087)	(0.165)	(0.134)	(0.123)
intedsomecoll	-0.012	0.036	0.022	-0.068	-0.036	0.034		0.021	-0.024	0.026	0.024	0.001
	(0.045)	(0.041)	(0.034)	(0.060)	(0.048)	(0.040)		(0.031)	(0.038)	(0.059)	(0.053)	(0.048)
High School Diploma	0.241**	0.184**	0.262***	-0.134	0.224**	0.190**		0.197***	0.155*	0.272*	0.251**	0.108
	(0.112)	(0.093)	(0.081)	(0.138)	(0.110)	(0.093)		(0.072)	(0.079)	(0.143)	(0.117)	(0.119)
intedhs	-0.004	-0.017	-0.018	0.092	-0.019	0.007		0.001	0.018	-0.016	-0.053	0.059
	(0.046)	(0.040)	(0.034)	(0.058)	(0.051)	(0.038)		(0.031)	(0.039)	(0.056)	(0.052)	(0.052)
Female			-0.100	-0.289**	-0.225***	-0.027	-0.295	-0.126**	-0.141**	-0.049	-0.146	-0.209**
			(0.066)	(0.117)	(0.085)	(0.082)	(0.236)	(0.061)	(0.065)	(0.116)	(0.100)	(0.095)
intfemale			0.006	-0.010	0.014	-0.006	0.034	0.013	0.016	-0.011	0.000	0.043
			(0.021)	(0.039)	(0.029)	(0.027)	(0.056)	(0.022)	(0.027)	(0.036)	(0.034)	(0.032)
Adj. R^2	0.089	0.135	0.109	0.076	0.094	0.124	0.055	0.072	0.083	0.099	0.125	0.108
Num Obs	6053	5698	9569	2182	5070	6681	2453	9298	9536	3276	4164	4311

Notes: Data from the 2005-2006, 2007-2008 and 2009-2010 waves of NHANES including adults over age 20. Each column shows coefficients from a regression using self-assessed health as the dependent variable. Robust standard errors are shown in parenthesis, and stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%

Table A13: Generalized Ordered Probit Regression Results of Self-Reported Health on Income and Allostasis

	$P(SAH > 1)$	$P(SAH > 2)$	$P(SAH > 3)$	$P(SAH > 4)$
Allostasis	0.005 (0.001) ***	0.040 (0.004) ***	0.077 (0.007) ***	0.036 (0.004) ***
Allostasis x IPR	0.000 (0.001)	0.006 (0.002) ***	0.011 (0.004) ***	0.007 (0.003) ***
Income-to-Poverty Ratio	0.002 (0.002)	0.025 (0.008) ***	0.040 (0.015) ***	0.009 (0.008)
Income-to-Poverty Ratio Squared	-0.001 (0.000) ***	-0.006 (0.002) ***	-0.003 (0.003)	0.000 (0.002)
Age/100	-0.023 (0.007) ***	-0.141 (0.024) ***	-0.110 (0.041) ***	-0.004 (0.024)
Age/100 x IPR	0.006 (0.006)	0.022 (0.018)	0.030 (0.027)	0.031 (0.015) **
Age/100 Squared	0.081 (0.033) **	0.487 (0.129) ***	1.015 (0.215) ***	0.277 (0.121) **
Age/100 Squared x IPR	-0.068 (0.044)	-0.466 (0.122) ***	-0.458 (0.139) ***	-0.134 (0.081) *
Female	0.008 (0.000)	-0.001 (0.000)	-0.001 (0.000)	-0.036 (0.000)
Female x IPR	0.002 (0.001)	0.008 (0.005) *	0.012 (0.007)	-0.001 (0.004)
Black	-0.003 (0.000)	0.071 (0.000)	-0.027 (0.000)	-0.032 (0.000)
Black x IPR	-0.001 (0.001)	-0.008 (0.007)	-0.022 (0.009) **	-0.012 (0.005) **
Hispanic	0.006 (0.000)	0.073 (0.000)	-0.029 (0.000)	-0.015 (0.000)
Hispanic x IPR	-0.002 (0.002)	0.007 (0.007)	0.010 (0.009)	-0.003 (0.005)
Married	-0.006 (0.000)	-0.028 (0.000)	0.004 (0.000)	-0.015 (0.000)
Married x IPR	0.001 (0.002)	0.004 (0.007)	0.005 (0.011)	-0.003 (0.007)
Widowed	-0.022 (0.000)	-0.007 (0.000)	0.003 (0.000)	-0.038 (0.000)
Widowed x IPR	0.000 (0.002)	-0.006 (0.011)	0.014 (0.021)	-0.010 (0.010)
Divorced	-0.001 (0.000)	0.005 (0.000)	-0.002 (0.000)	0.002 (0.000)
Divorced x IPR	0.000 (0.002)	0.002 (0.009)	0.002 (0.015)	0.003 (0.008)
College Degree	-0.039 (0.000)	-0.143 (0.000)	-0.007 (0.000)	0.121 (0.000)
College Degree x IPR	-0.002 (0.003)	-0.013 (0.009)	-0.006 (0.013)	0.003 (0.010)
Some College	-0.020 (0.000)	-0.102 (0.000)	0.006 (0.000)	0.047 (0.000)
Some College x IPR	0.000 (0.002)	-0.009 (0.008)	-0.002 (0.012)	0.006 (0.007)
High School Diploma	-0.010 (0.000)	-0.054 (0.000)	0.005 (0.000)	0.036 (0.000)
High School Diploma x IPR	0.000 (0.002)	-0.013 (0.007) *	-0.008 (0.013)	0.003 (0.008)
Pseudo R^2	0.079			
Num Obs	11751			

Notes: Data from the 2005-2006, 2007-2008 and 2009-2010 waves of NHANES including adults over age 20. Each column shows marginal effects at the mean from a generalized ordered probit regression using self-assessed health as the dependent variable. Robust standard errors are shown in parenthesis, and stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%

Table A14: Regression Results of Self-Assessed Health on Income and Allostasis Excluding Individuals without Regular Medical Care

	Self-Assessed Health	Physically Healthy Days	Mentally Healthy Days
Allostasis	0.210 (0.017) ***	0.379 (0.118) ***	0.255 (0.116) **
Allostasis x IPR	0.058 (0.011) ***	0.039 (0.072)	-0.045 (0.069)
Income-to-Poverty Ratio	0.089 (0.045) **	0.535 (0.254) **	1.110 (0.244) ***
Income-to-Poverty Ratio Squared	-0.011 (0.010)	-0.319 (0.060) ***	-0.185 (0.063) ***
Age/100	-0.252 (0.123) **	-3.908 (0.680) ***	4.708 (0.686) ***
Age/100 x IPR	0.209 (0.081) ***	0.421 (0.440)	0.216 (0.416)
Age/100 Squared	2.848 (0.615) ***	20.136 (3.707) ***	27.835 (3.742) ***
Age/100 Squared x IPR	-1.536 (0.411) ***	-10.162 (2.316) ***	-10.128 (2.167) ***
Female	-0.120 (0.069) *	-0.678 (0.449)	-2.542 (0.453) ***
Female x IPR	0.020 (0.022)	0.065 (0.124)	0.294 (0.121) **
Black	-0.026 (0.078)	1.113 (0.542) **	1.845 (0.585) ***
Black x IPR	-0.077 (0.025) ***	-0.284 (0.159) *	-0.516 (0.178) ***
Hispanic	-0.143 (0.073) **	1.118 (0.526) **	2.525 (0.558) ***
Hispanic x IPR	-0.018 (0.027)	-0.515 (0.174) ***	-0.672 (0.175) ***
Married	-0.099 (0.109)	-0.553 (0.621)	0.816 (0.685)
Married x IPR	0.018 (0.039)	0.207 (0.188)	-0.184 (0.198)
Widowed	-0.038 (0.145)	-0.141 (1.086)	1.207 (1.005)
Widowed x IPR	-0.016 (0.055)	-0.210 (0.374)	-0.583 (0.311) *
Divorced	-0.169 (0.129)	-0.739 (0.831)	-1.209 (0.940)
Divorced x IPR	0.017 (0.048)	-0.078 (0.276)	0.095 (0.273)
College Degree	0.775 (0.152) ***	1.039 (0.774)	2.818 (0.747) ***
College Degree x IPR	-0.017 (0.043)	0.115 (0.236)	-0.522 (0.212) **
Some College	0.201 (0.085) **	-0.330 (0.613)	0.120 (0.633)
Some College x IPR	0.032 (0.033)	0.267 (0.215)	-0.129 (0.200)
High School Diploma	0.188 (0.081) **	0.342 (0.635)	1.201 (0.616) *
High School Diploma x IPR	-0.010 (0.034)	0.036 (0.232)	-0.444 (0.209) **
Adj. R^2	0.121	0.048	0.068
Num Obs	8974	8974	8974
P-value for Chow test	0.000	0.000	0.000

Notes: Data from the 2005-2006, 2007-2008 and 2009-2010 waves of NHANES including adults over age 20 and excludes individuals that did not have a regular medical care facility or did not visit a medical professional at least once in the last year. Each column shows coefficients from a regression using different measures of self-assessed health. In addition to the coefficients shown, all models include survey wave fixed effects. All regressions include NHANES sample weights. Robust standard errors are shown in parenthesis, and stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%

Table A15: Regression Results of Self-Assessed Health on Income and Allostasis Excluding Individuals on Prescription Drugs

	Self-Assessed Health	Physically Healthy Days	Mentally Healthy Days
Allostasis	0.218 (0.028) ***	0.145 (0.133)	0.105 (0.158)
Allostasis x IPR	0.073 (0.018) ***	-0.008 (0.084)	0.045 (0.084)
Income-to-Poverty Ratio	0.113 (0.060) *	0.402 (0.181) **	0.391 (0.267)
Income-to-Poverty Ratio Squared	0.014 (0.014)	-0.082 (0.058)	0.008 (0.065)
Age/100	0.744 (0.239) ***	-2.054 (1.017) **	3.054 (1.064) ***
Age/100 x IPR	0.656 (0.175) ***	1.188 (0.654) *	0.689 (0.633)
Age/100 Squared	2.949 (1.145) **	5.058 (4.902)	13.833 (4.883) ***
Age/100 Squared x IPR	0.286 (0.827)	-1.978 (3.283)	-2.990 (3.035)
Female	-0.186 (0.092) **	-0.560 (0.370)	-2.164 (0.448) ***
Female x IPR	0.013 (0.033)	0.064 (0.112)	0.283 (0.133) **
Black	-0.072 (0.117)	0.693 (0.498)	0.278 (0.641)
Black x IPR	-0.064 (0.039)	-0.136 (0.151)	-0.174 (0.216)
Hispanic	-0.417 (0.100) ***	0.768 (0.440) *	2.390 (0.542) ***
Hispanic x IPR	0.038 (0.038)	-0.395 (0.160) **	-0.682 (0.197) ***
Married	0.131 (0.128)	0.202 (0.431)	0.720 (0.586)
Married x IPR	-0.012 (0.048)	-0.013 (0.148)	-0.036 (0.194)
Widowed	0.094 (0.262)	-0.988 (2.783)	0.616 (1.536)
Widowed x IPR	-0.110 (0.108)	0.285 (0.886)	-0.564 (0.587)
Divorced	0.062 (0.170)	1.076 (0.735)	0.251 (0.894)
Divorced x IPR	0.011 (0.069)	-0.618 (0.300) **	-0.166 (0.297)
College Degree	0.921 (0.198) ***	0.870 (0.579)	2.483 (0.781) ***
College Degree x IPR	-0.028 (0.066)	-0.096 (0.188)	-0.397 (0.254)
Some College	0.497 (0.131) ***	-0.570 (0.513)	0.399 (0.642)
Some College x IPR	-0.041 (0.056)	0.174 (0.179)	0.082 (0.237)
High School Diploma	0.302 (0.119) **	0.255 (0.522)	1.098 (0.609) *
High School Diploma x IPR	-0.013 (0.056)	0.017 (0.190)	-0.152 (0.242)
Adj. R^2	0.123	0.020	0.030
Num Obs	4898	4898	4898
P-value for Chow test	0.000	0.008	0.006

Notes: Data from the 2005-2006, 2007-2008 and 2009-2010 waves of NHANES including adults over age 20 and excludes individuals who reported taking prescription drugs in the previous year. Each column shows coefficients from a regression using different measures of self-assessed health. In addition to the coefficients shown, all models include survey wave fixed effects. All regressions include NHANES sample weights. Robust standard errors are shown in parenthesis, and stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%

Table A16: Concentration Index Results Using Income-to-Poverty Ratio and Biomarkers

	Body Mass Index	Choles-terol	Trigly-cerides	Glycohe-moglobin	Creat-inine	C-Reactive Protein	White Blood Cell Count	Systolic Blood Pressure	Diastolic Blood Pressure
Income-to-Poverty Ratio	0.003	0.002	0.030***	0.004***	0.002*	0.047***	0.013***	0.003***	-0.000
	(0.002)	(0.001)	(0.006)	(0.001)	(0.001)	(0.016)	(0.002)	(0.001)	(0.001)
Age/100	-0.017***	-0.018***	-0.047***	-0.047***	0.152***	-0.109***	0.025***	-0.054***	0.007***
	(0.003)	(0.002)	(0.008)	(0.001)	(0.002)	(0.019)	(0.003)	(0.002)	(0.002)
Age/100 Squared	0.186***	0.254***	0.550***	0.070***	0.050***	0.461***	0.041**	-0.024***	0.298***
	(0.014)	(0.013)	(0.046)	(0.008)	(0.010)	(0.117)	(0.018)	(0.008)	(0.010)
Female	0.001	-0.003***	0.039***	0.002***	-0.002***	-0.037***	0.000	0.006***	0.006***
	(0.001)	(0.001)	(0.003)	(0.000)	(0.001)	(0.007)	(0.001)	(0.000)	(0.001)
Black	-0.011***	0.007***	0.049***	-0.008***	-0.011***	-0.043***	0.020***	-0.006***	-0.003***
	(0.001)	(0.001)	(0.003)	(0.001)	(0.001)	(0.009)	(0.001)	(0.001)	(0.001)
Hispanic	-0.004***	-0.000	-0.006	-0.006***	-0.010***	-0.011	0.004***	-0.000	0.002**
	(0.001)	(0.001)	(0.004)	(0.001)	(0.001)	(0.009)	(0.001)	(0.001)	(0.001)
Married	0.001	0.001	-0.005	0.001**	0.000	0.014	-0.000	0.001**	0.001
	(0.001)	(0.001)	(0.004)	(0.001)	(0.001)	(0.009)	(0.002)	(0.001)	(0.001)
Widowed	0.001	-0.007***	-0.014**	0.001	0.002	0.006	-0.002	-0.001	0.000
	(0.002)	(0.002)	(0.006)	(0.001)	(0.001)	(0.017)	(0.002)	(0.001)	(0.001)
Divorced	0.003	-0.001	-0.005	0.001	0.000	-0.014	-0.010***	0.001	0.000
	(0.002)	(0.001)	(0.006)	(0.001)	(0.001)	(0.017)	(0.002)	(0.001)	(0.001)
College Degree	0.007***	0.001	0.026***	0.006***	0.004***	0.027**	0.013***	0.004***	0.001
	(0.001)	(0.001)	(0.005)	(0.001)	(0.001)	(0.011)	(0.002)	(0.001)	(0.001)
Some College	-0.003**	-0.000	0.001	0.002***	0.004***	0.002	0.005***	0.000	-0.002***
	(0.001)	(0.001)	(0.005)	(0.001)	(0.001)	(0.011)	(0.002)	(0.001)	(0.001)
High School Diploma	-0.002*	-0.000	0.005	0.002***	0.002***	-0.008	0.001	0.000	-0.000
	(0.001)	(0.001)	(0.005)	(0.001)	(0.001)	(0.010)	(0.002)	(0.001)	(0.001)
Adj. R^2	0.048	0.071	0.064	0.134	0.537	0.017	0.041	0.189	0.127
Num Obs	11751	11751	11751	11751	11751	11751	11751	11751	11751

Notes: Data from the 2005-2006, 2007-2008 and 2009-2010 waves of NHANES including adults over age 20. Each regression shows coefficients from a concentration index model, and robust standard errors are shown in parenthesis. In addition to the coefficients shown, all models include survey wave fixed effects. All models include NHANES sample weights. Stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%

Table A17: Concentration Index Results Using Income-to-Poverty Ratio Excluding Individuals without Regular Medical Care

	Self-Reported Health				
	Self-Assessed Health	Physically Healthy Days	Mentally Healthy Days	Allostasis	Allostasis (Count)
Income-to-Poverty Ratio	0.062***	0.022***	0.024***	0.005***	0.006***
	(0.006)	(0.003)	(0.002)	(0.001)	(0.001)
Age/100	-0.055***	-0.025***	0.025***	-0.037***	-0.044***
	(0.011)	(0.004)	(0.004)	(0.001)	(0.002)
Age/100 Squared	0.269***	0.116***	0.160***	0.056***	0.071***
	(0.057)	(0.021)	(0.021)	(0.008)	(0.010)
Female	0.001	-0.002*	-0.009***	0.004***	0.005***
	(0.003)	(0.001)	(0.001)	(0.000)	(0.001)
Black	-0.024***	0.002	0.002	-0.002***	-0.002***
	(0.004)	(0.001)	(0.002)	(0.001)	(0.001)
Hispanic	-0.020***	-0.002	0.004**	-0.002***	-0.002**
	(0.004)	(0.002)	(0.002)	(0.001)	(0.001)
Married	-0.004	0.001	0.002	0.000	-0.000
	(0.006)	(0.002)	(0.002)	(0.001)	(0.001)
Widowed	-0.010	-0.004	-0.002	-0.003**	-0.004***
	(0.008)	(0.003)	(0.003)	(0.001)	(0.001)
Divorced	-0.013*	-0.006**	-0.006**	-0.001	-0.001
	(0.007)	(0.002)	(0.003)	(0.001)	(0.001)
College Degree	0.069***	0.008***	0.007***	0.005***	0.007***
	(0.005)	(0.002)	(0.002)	(0.001)	(0.001)
Some College	0.026***	0.002	-0.001	0.000	0.001
	(0.004)	(0.002)	(0.002)	(0.001)	(0.001)
High School Diploma	0.014***	0.002	0.000	0.001	0.001
	(0.004)	(0.002)	(0.002)	(0.001)	(0.001)
Adj. R^2	0.094	0.041	0.061	0.141	0.139
Num Obs	8974	8974	8974	8974	8974

Notes: Data from the 2005-2006, 2007-2008 and 2009-2010 waves of NHANES including adults over age 20 and excludes individuals that did not have a regular medical care facility or did not visit a medical professional at least once in the last year. Each regression shows coefficients from a concentration index model, and robust standard errors are shown in parenthesis. The first three columns measure health using self-assessed health measures, and the remaining column measures allostasis. In addition to the coefficients shown, all models include survey wave fixed effects. All models include NHANES sample weights, and stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%

Table A18: Concentration Index Results Using Income-to-Poverty Ratio Excluding Individuals on Prescription Drugs

	Self-Reported Health				
	Self-Assessed Health	Physically Healthy Days	Mentally Healthy Days	Allostasis	Allostasis (Count)
Income-to-Poverty Ratio	0.046***	0.008***	0.009***	0.002	0.002
	(0.009)	(0.002)	(0.003)	(0.001)	(0.001)
Age/100	0.051**	-0.010*	0.020***	-0.037***	-0.040***
	(0.021)	(0.006)	(0.006)	(0.003)	(0.004)
Age/100 Squared	0.295***	0.030	0.084***	0.043***	0.064***
	(0.101)	(0.028)	(0.029)	(0.014)	(0.017)
Female	-0.005	-0.002	-0.008***	0.006***	0.007***
	(0.005)	(0.001)	(0.001)	(0.001)	(0.001)
Black	-0.020***	0.002	-0.002	-0.000	0.000
	(0.005)	(0.001)	(0.002)	(0.001)	(0.001)
Hispanic	-0.031***	-0.001	0.006***	-0.003***	-0.004***
	(0.005)	(0.001)	(0.002)	(0.001)	(0.001)
Married	0.010	0.002	0.004**	0.001	0.001
	(0.006)	(0.001)	(0.002)	(0.001)	(0.001)
Widowed	-0.024*	-0.004	-0.005	-0.002	-0.004
	(0.013)	(0.006)	(0.005)	(0.002)	(0.003)
Divorced	0.005	-0.004	-0.002	-0.001	-0.002
	(0.009)	(0.003)	(0.003)	(0.001)	(0.002)
College Degree	0.081***	0.003	0.007***	0.004***	0.005***
	(0.008)	(0.002)	(0.002)	(0.001)	(0.001)
Some College	0.033***	-0.001	0.004*	-0.001	-0.000
	(0.006)	(0.002)	(0.002)	(0.001)	(0.001)
High School Diploma	0.023***	0.001	0.004*	-0.001	-0.002*
	(0.006)	(0.002)	(0.002)	(0.001)	(0.001)
Adj. R^2	0.099	0.015	0.026	0.125	0.117
Num Obs	4898	4898	4898	4898	4898

Notes: Data from the 2005-2006, 2007-2008 and 2009-2010 waves of NHANES including adults over age 20 and excludes individuals who reported taking prescription drugs in the previous year. Each regression shows coefficients from a concentration index model, and robust standard errors are shown in parenthesis. The first three columns measure health using self-assessed health measures, and the remaining column measures allostasis. In addition to the coefficients shown, all models include survey wave fixed effects. All models include NHANES sample weights, and stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%

Table A19: Erreygers Index Results Using Income-to-Poverty Ratio

	Self-Reported Health				
	Self-Assessed Health	Physically Healthy Days	Mentally Healthy Days	Allostasis	Allostasis (Count)
Income-to-Poverty Ratio	0.075***	0.060***	0.069***	0.022***	0.025***
	(0.008)	(0.008)	(0.008)	(0.004)	(0.005)
Age/100	-0.076***	-0.093***	0.088***	-0.188***	-0.196***
	(0.013)	(0.011)	(0.011)	(0.006)	(0.007)
Age/100 Squared	0.289***	0.312***	0.521***	0.271***	0.305***
	(0.068)	(0.065)	(0.064)	(0.036)	(0.040)
Female	-0.005	-0.014***	-0.036***	0.022***	0.022***
	(0.004)	(0.003)	(0.004)	(0.002)	(0.002)
Black	-0.028***	0.006	0.007	-0.009***	-0.006**
	(0.004)	(0.004)	(0.005)	(0.003)	(0.003)
Hispanic	-0.028***	0.005	0.025***	-0.010***	-0.010***
	(0.005)	(0.004)	(0.005)	(0.002)	(0.003)
Married	0.000	0.005	0.009	0.000	-0.001
	(0.006)	(0.005)	(0.006)	(0.003)	(0.003)
Widowed	-0.009	-0.011	-0.004	-0.013***	-0.018***
	(0.009)	(0.010)	(0.010)	(0.005)	(0.006)
Divorced	-0.006	-0.017**	-0.015*	-0.004	-0.006
	(0.008)	(0.007)	(0.008)	(0.004)	(0.005)
College Degree	0.090***	0.024***	0.027***	0.022***	0.027***
	(0.006)	(0.006)	(0.006)	(0.003)	(0.004)
Some College	0.038***	0.007	0.001	-0.000	0.001
	(0.005)	(0.005)	(0.006)	(0.003)	(0.003)
High School Diploma	0.023***	0.007	0.001	0.001	0.001
	(0.005)	(0.006)	(0.006)	(0.003)	(0.003)
Adj. R^2	0.084	0.035	0.051	0.139	0.134
Num Obs	11751	11751	11751	11751	11751

Notes: Data from the 2005-2006, 2007-2008 and 2009-2010 waves of NHANES including adults over age 20. Each regression shows coefficients from a concentration index model, and robust standard errors are shown in parenthesis. The first three columns measure health using self-assessed health measures, and the remaining column measures allostasis. Dependent variables are transformed using the method suggested by Erreygers (2009a). In addition to the coefficients shown, all models include survey wave fixed effects. All models include NHANES sample weights, and stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%

Table A20: Concentration Index Results Using Alternative Measures of Self-Assessed Health

	Self-Assessed Health	Self-Assessed Health (HUQ)	Self-Assessed Health (0/1)	HUQ Self-Assessed Health (0/1)	Self-Assessed Health (HUI)	HUQ Self-Assessed Health (HUI)
Income-to-Poverty Ratio Incl. Educ. (Weighted)	0.058***	0.062***	0.040***	0.043***	0.017***	0.019***
	(0.006)	(0.006)	(0.003)	(0.003)	(0.000)	(0.000)
Income-to-Poverty Ratio Incl. Educ. (Unweighted)	0.071***	0.076***	0.051***	0.055***	0.019***	0.022***
	(0.006)	(0.005)	(0.003)	(0.003)	(0.000)	(0.000)
Family Income (Category) Incl. Educ. (Weighted)	0.052***	0.056***	0.035***	0.038***	0.015***	0.018***
	(0.006)	(0.006)	(0.003)	(0.003)	(0.000)	(0.000)
Family Income (Category) Incl. Educ. (Unweighted)	0.063***	0.068***	0.044***	0.049***	0.017***	0.020***
	(0.006)	(0.005)	(0.003)	(0.003)	(0.000)	(0.000)
Household Income (Category) Incl. Educ. (Weighted)	0.052***	0.052***	0.036***	0.039***	0.015***	0.017***
	(0.006)	(0.006)	(0.003)	(0.003)	(0.000)	(0.000)
Household Income (Category) Incl. Educ. (Unweighted)	0.063***	0.065***	0.044***	0.049***	0.017***	0.020***
	(0.006)	(0.005)	(0.003)	(0.003)	(0.000)	(0.000)
Income-to-Poverty Ratio No Educ. (Weighted)	0.094***	0.094***	0.055***	0.058***	0.024***	0.027***
	(0.006)	(0.005)	(0.003)	(0.003)	(0.000)	(0.000)
Income-to-Poverty Ratio No Educ. (Unweighted)	0.113***	0.115***	0.074***	0.078***	0.027***	0.032***
	(0.005)	(0.005)	(0.003)	(0.003)	(0.000)	(0.000)
Family Income (Category) No Educ. (Weighted)	0.085***	0.087***	0.050***	0.053***	0.022***	0.025***
	(0.006)	(0.005)	(0.003)	(0.003)	(0.000)	(0.000)
Family Income (Category) No Educ. (Unweighted)	0.103***	0.105***	0.066***	0.071***	0.025***	0.029***
	(0.005)	(0.005)	(0.003)	(0.003)	(0.000)	(0.000)
Household Income (Category) No Educ. (Weighted)	0.086***	0.083***	0.050***	0.053***	0.022***	0.024***
	(0.006)	(0.005)	(0.003)	(0.003)	(0.000)	(0.000)
Household Income (Category) No Educ. (Unweighted)	0.103***	0.103***	0.066***	0.071***	0.025***	0.029***
	(0.005)	(0.005)	(0.003)	(0.003)	(0.000)	(0.000)

Notes: Data from the 2005-2006, 2007-2008 and 2009-2010 waves of NHANES including adults over age 20. Each cell shows coefficients from a concentration index model. Each row shows results using a different measure of income and either using NHANES sample weights or not. In addition to the income coefficient shown, all models additionally control for the same controls as in Table 6, except for education controls as noted. Each column shows results using a different measure of self-assessed health: (1) Cardinalized SAH from the Current Health Status Questionnaire; (2) Cardinalized SAH from the Hospital Utilization and Access to Care Questionnaire; (3) Untransformed SAH from the current health questionnaire; (4) Untransformed SAH from the Hospital Utilization and Access to Care Questionnaire; (5) Predicted health score from an interval regression after mapping the cdf of SAH from the Current Health Status Questionnaire onto the cdf of the Canadian HUI; and (6) Predicted health score from an interval regression after mapping the cdf of SAH from the Hospital Utilization and Access to Care Questionnaire onto the cdf of the Canadian HUI. Robust standard errors are shown in parenthesis, and stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%.