The Measurement of Health and the Connection Between Health Inequality and Income*

Erik T. Nesson†
Joshua J. Robinson‡

June 27, 2017

Abstract

We examine the extent to which self-reported health measures suffer from reporting bias and then characterize how this reporting bias affects the estimation of income-related health inequality as measured by the concentration index. We run a comprehensive set of tests of reporting bias using several self-reported health measures and several clinical measures of health from the National Health and Nutritional Examination Surveys. Our results confirm the existence of significant, positive, income-related reporting bias and also suggest that higher income individuals react more strongly to a change in objective, clinical health measures. We find that self-reported health measures significantly overstate the degree of income-related health inequality relative to clinical health measures. Parallel to and in support of the analysis described above, we propose the use of a multidimensional measure of clinical health in the context of measuring income-related health inequality.

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 (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 large issue is the measure of health used in these models. To measure income-related health inequality accurately requires a representative measure of health which does not suffer from reporting bias. 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 bias individuals' perception of their health status.

We examine the extent to which SAH is affected by reporting bias and how this bias affects the degree of income-related health inequality as measured by the condition index. We use data from the National Health and Nutritional Examination Surveys (NHANES) which contain measures of self-reported health, including SAH, and results from physical examinations and blood, saliva, and urine samples. Using multiple different methods, we find evidence of income-related reporting bias in SAH. We also find that SAH suggests there is significantly more income-related health inequality than clinical health measures. 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. Finally, we propose simple method for constructing an objective, multivariate measure of health based on widely available clinical measures.

Our paper makes three main contributions. First, we use multiple different methods to show that, holding clinical measures of health constant, individuals with different incomes will report their health differently. Using cluster analysis to group health measures and income based on the "closeness" or similarity of their distributions, we find that SAH and physically and mentally health days are much more closely related to income than to the clinical measures of health. Additionally, we estimate a series of ordered probit regressions of

SAH on an objective health measure and test for different types of reporting bias using the method suggested by Lindeboom and van Doorslaer (2004). We find evidence of reporting bias by income, education, race, sex and age. To better characterize the nature of the income-related reporting bias, we regress measures of self-reported health on an objective health measure and demographic variables that are fully interacted with income. These results show that individuals of identical objective health rate their health better as their income increases. Furthermore, higher income individuals will make larger adjustments to their SAH after a change in clinical health. This type of bias is concerning for the health inequality literature, where the main research question concerns the relationship between SES and health.

Second, we characterize the effect of reporting bias on income-related health inequality by comparing the concentration indices of multiple clinical and self-reported measures of health. Although previous studies show that SES, culture, or other variables unrelated to health may also affect an individual's SAH (e.g. Crossley and Kennedy, 2002; Johnston et al., 2009; Greene et al., 2014; Zajacova and Dowd, 2011), most do not attempt to characterize what impact this bias has on health inequality. Although a few studies have compared subjective and objective measures of health in the condition index, these papers examine narrow objective health measures, such as grip strength, or another health measure based on a subjective assessment (e.g. Ziebarth, 2010; Lindeboom and van Doorslaer, 2004). In comparison, we examine SAH, self-reported physically and mentally healthy days over the past 30 days, nine common univariate clinical measures of health, and a multivariate measure of clinical health.

Third, we propose a simple framework for using a multidimensional health measure based on the clinical health of an individual. Our measure has the advantage of being based on objective health assessment, being representative of a broad spectrum of health conditions, and being a ratio scale measure that can be used with the condition index to measure health inequality. Traditionally, measures of health inequality have relied on a single attribute and cannot account for the multivariate nature of health status. Our measure uses multiple clinical health status 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 and Stellar, 1993; McEwen, 1998; Karlamangla et al., 2002; Seeman et al., 2001). Similar methods have recently been used in the health inequality literature to capture the breadth of health using a multivariate framework (e.g. Makdissi and Yazbeck, 2014). However, our framework addresses two common problems encountered by previous studies: the arbitrary choices of attributes and the arbitrary choice of healthy cut-points.

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, Section 4 summarizes our results, and Section 5 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 Lorenz curve applied to any positive and continuously differentiable function of income (Kakwani, 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 $C_h = {}^{2}/_{\mu_h}\text{cov}(h_i, R_i)$. Wagstaff et al. (1991) used this observation to show that when using individual-level data, the concentration index can be estimated with a simple linear regression,

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

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). The coefficient on R_i, β_1, is equivalent to the concentration index, and the standard error for β_1 allows researchers to conduct statistical inference.[1,2] Researchers have used these related measurements of the relationship between SES and health inequality in a wide range of settings (e.g. Zhang and Wang, 2004; Kennedy et al., 1998; Deaton and Paxson, 1998; Trannoy et al., 2010; Rosa Dias, 2009; Dolores Montoya Diaz, 2002).

Despite its ubiquity in the health inequality literature, many potential methodological problems with the concentration index have been identified. Bleichrodt et al. (2012) find little to no evidence that experimental subjects had a preference for decreasing the correlation between health and SES, which violates the social welfare function implied by using the concentration index (Bleichrodt and van Doorslaer, 2006). Another challenge is raised by Fleurbaey and Schokkaert (2009) who point out that the concentration index is not necessarily a measure of SES-related inequality because the measurement is likely affected

[1]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 and Jones, 2003).

[2]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.

by (potentially sample specific) correlations depending on the variables used. Ignoring the other sources of inequality increases the likelihood of mismeasurement. Furthermore, Carrieri and Jones (2017) show that the relationship between income and health varies across the distribution of health and may be quite different for men and women.

Others have challenged the appropriateness of the use of the concentration index for bounded variables. Clarke et al. (2002) 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.[3] Finally, the very plausible scenario of reverse causality—that is, that SES is determined by health—is rarely considered or addressed.

While we do not think the critiques should be taken lightly, addressing 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.

2.2 Self-Assessed Health Cardinalization

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 individuals' health (e.g. Zhang and Wang, 2004; Kennedy et al., 1998; Deaton and Paxson, 1998; Trannoy et al., 2010; Rosa Dias, 2009; Dolores Montoya Diaz, 2002). The use of SAH as a proxy for health status presents a number of challenges, particularly in the context

[3] 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 and Van Ourti, 2011a; Lambert and Zheng, 2011; Wagstaff, 2011; Erreygers and Van Ourti, 2011b). Finally, 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.

of health inequality. The first challenge is to transform SAH into a cardinal measure—or preferably, in the case of the standard concentration index, a ratio scale measure.[4] This inevitably involves assumptions about the nature of the relationship between individuals' subjective evaluation of their health and their actual health status.[5]

There are three common methods of cardinalizing SAH for the purpose of measuring health inequality: dichotomization, log-normal transformation, and prediction from interval regression. Dichotomization involves setting a cut-point for "good" health (or "poor" health) and then measuring health achievement (or shortfall) as a binary indicator. This method dates back to Wagstaff et al. (1991)'s seminal paper on SES-related health inequality. This method is very simple, easy to implement, and has a natural similarity to many clinical measures of health that are judged against a standard cut-point. However, it suffers from a number of shortcomings. The most obvious problem is that significant distributional information about the health variable is lost. Furthermore, Wagstaff and van Doorslaer (1994) point out that the concentration index is highly sensitive to the cut-points used. Finally, there are the previously mentioned challenges by Clarke et al. (2002) and Erreygers (2009a) (see Section 2.1).

The log-normal transformation, first proposed by Wagstaff and van Doorslaer (1994), involves assuming the distribution of the latent health variable (multiplied by -1) is log-normal and uses the cumulative distribution of the ordinal SAH values to parametrize the log-normal distribution. The primary shortcoming of this method is that health may not follow a log-normal distribution. However, Gerdtham et al. (1999) show that this method yields the same conclusions about health inequality as using a more objective quality adjusted life year (QALY) score to measure health outcomes.

Finally, some researchers map SAH values onto a health utility index (HUI). For ex-

[4] 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 use to estimate a concentration index.

[5] 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.

ample, van Doorslaer and Jones (2003) compare multiple methods of transforming SAH to a more comprehensive health assessment (the Canadian HUI) and show that an interval regression-based predicted health score yields inequality results that more consistently mirror the results of the generic health measures, which are assumed to be superior. However, while generic measures like the HUI are more comprehensive than a typical 5-point health self assessment, they are still inherently based on subjective evaluations. Furthermore, since this cardinalization method involves using predicted values from a regression, there will be significantly less overall variation in the health score, which is particularly problematic when examining inequality. The most problematic feature of this method, however, is that, by construction, most of the variation in the predicted health score will be determined by the covariates of the interval regression. Since the SES measure used to compute the concentration index will almost certainly be one of these covariates—and is likely to be one of the few continuous measures—this makes looking for SES-related inequality with this health measure somewhat of a tautology.

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).[6] 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.

2.3 Reporting Bias in Self-Assessed Health

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, most importantly, SES—a problem commonly referred to as *reporting heterogeneity* or

[6] In addition to what is discussed above, the controversy surrounding dichotomizing SAH values is further discussed in Footnote 3. Concerning the interval regression approach, Ziebarth (2010) found evidence that the transformed data is highly sensitive to the generic health measure used.

reporting bias. The primary focus of this paper is to determine whether reporting bias exists, to measure the extent that it effects the measurement of health inequality, and to evaluate an alternative measure to deal with this problem.

2.3.1 Background

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., 2014; Crossley and Kennedy, 2002; Zajacova and Dowd, 2011). Groot (2000) finds evidence that adaptations to chronic conditions and pain can change an individuals reference points for SAH, and Frijters and Ulker (2008) finds that controlling for individual fixed effects can dramatically change the statistical relationship between SAH and its determinants.

Reporting errors are more problematic if different groups of people—whether we define groups by gender, ethnicity, age, or SES—systematically self-report their health in different ways. When using an ordinal, subjective measure, like SAH, as a proxy for objective health, the researcher must assume that the choice of SAH response is based on a stable mapping between the individual's unobservable true health status and the ordered responses. Formally, 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 a different criteria to choose the ordinal value based on their true health.

Lindeboom and van Doorslaer (2004) outline two different types of reporting bias and propose tests for detecting them. An *index shift* is a form of reporting bias in which differential reporting behavior across subgroups leads to a parallel shift of the threshold parameters that determined the response categories 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 for 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, $\alpha_{j,m}$. Cut-point shifts are considered to be more problematic because an index shift can be corrected by including indicator variables for the relevant groups, which allows for a parallel shift (α_m) in the cut points. Detecting either type of reporting bias reliably requires conditioning on an objective measure of health (Lindeboom and van Doorslaer, 2004).

With regard to using SAH to evaluate a health concentration index, SES-related reporting bias 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 great deal of evidence in the literature that reporting bias exists across age and gender, but the evidence for SES-related reporting bias is mixed.

Some researchers test for reporting bias by conditioning a standard concentration index regression on a generic health measure.[7] Lindeboom and van Doorslaer (2004) find no evidence of reporting bias by income or education after conditioning SAH on the Canadian HUI. Layes et al. (2012), on the other hand, find evidence that lower income individuals overstate their health status, also conditioning on the HUI. Similarly Bago d'Uva et al. (2008) show that better educated individuals are less prone to rate their health highly—leading to an underestimation of health inequality—by conditioning on health vignettes.[8] However, Shmueli (2003) uses a multiple indicators–multiple causes model to estimate the latent health variable and finds that both SAH and SF-36 (a generic health measure) responses are positively

[7] Generic health measures are based on responses to questionnaires about multiple dimensions of health combined using a utility-based index function. It is worth noting that, despite being more comprehensive, these measures are nonetheless still subjective in nature

[8] Respondents are asked to rate the health of a hypothetical person to serve as a reference point for their own self-assessment.

biased by income.[9]

Other researchers condition on a clinical measure of health to test for reporting bias. 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. Ziebarth (2010) finds that both SAH and the SF-12 (a generic health measure) are still positively correlated with income while holding grip strength, an objective health measure, constant. Cawley and Choi (2015) compare self-reports of health status to clinical measures of health and find that individuals with higher education self-report their health more accurately. Dowd and Zajacova (2007) show that the ability of SAH to predict mortality varies significantly by income and education levels. However, van Doorslaer and Gerdtham (2003) find no evidence of income or education related reporting bias. When conditioning on mortality, Jürges (2008) finds some evidence of income based heterogeneity, but only for women.

From the variety of different results in the literature, it is clear that issue of reporting bias in SAH is far from settled. We extend the literature of assessing the reliability and bias in SAH by conducting an extensive set of tests to measure the existence and extent of reporting bias in SAH. The first, cluster analysis, is a novel method in this literature. Second, we test for reporting bias using ordered probit regression-based techniques similar to Lindeboom and van Doorslaer (2004). Finally, we use regression techniques where we regress SAH on an objective measure of health, income, and demographics, and interactions between income and objective health and demographics, allowing us to further characterize the nature of the bias.

2.3.2 Detecting reporting bias using cluster analysis

Cluster analysis involves visually mapping out the hierarchical structure of a set of variables using the characteristics of the univariate and multivariate distributions. Cluster analysis

[9] 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”).

has been widely used in psychology, marketing, management and the economic analysis of multidimensional well-being (e.g. Punj and Stewart, 1983; Borgen and Barnett, 1987; Henry et al., 2005; Hirschberg et al., 1991, 2001a,b), but to the authors knowledge, we are the first to use this method in an effort to detect reporting bias in SAH. The basic notion of clustering is to form groups of variables based on the "closeness" or similarity of their distributions. Specifically, we focus on the hierarchical agglomerative clustering technique proposed by Ward (1963). Ward's method involves reducing k clusters into $k-1$ mutually exclusive clusters by considering every possible combination of cluster pairs and combining one pair such that the amount of information lost by the combination is minimized. Information lost from combining two clusters is quantified using the error sum of squares (ESS).[10] ESS is assessed for all clusters, and at each stage of amalgamation the next two clusters are combined such that the sum of the ESS measures for all clusters is minimized. To begin, each variable forms its own cluster. The process of combining clusters is performed until there is only one cluster left contains all the variables. Once a variable is combined into a cluster with at least one other variable it cannot be re-associated into another cluster. The result is a full mapping of the hierarchical structure of the data.

To use cluster analysis to detect reporting bias in SAH, we start with a set of variables containing cardinalized SAH, multiple objective health measures based on clinical assessments, and a SES measure. We map out the hierarchical structure of the data as described above. 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, for example, that would be an indication that the distribution of SAH shares more information with the distribution of SES than with actual objective health measures. This would be indicative of

[10] While not a comprehensive measure of distributional information, the ESS is the most relevant measure of distributional similarity for assessing the impact of potential bias on a regression-based measure, like the concentration index.

SES-related reporting bias.

2.3.3 Detecting reporting bias using ordered probit models

We also test for reporting heterogeneity by more conventional methods. We first use the methods described by Lindeboom and van Doorslaer (2004). This involves using ordered probit models of the ordinal SAH measure regressed on an objective health measure and implementing various tests of the parallel regression assumption to establish the existence of reporting shifts. The first step is to estimate a restricted model that imposes the parallel regression assumption for an objective health measure, SES, and other relevant covariates. Formally, we estimate the probably of an individual choosing $SAH_i = j$ as

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

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 do to differences in actual health (we will discuss 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 categories dummies). If there are M SES categories, k parameters in each model, and J cut-points, then the likelihood ratio test, $L(\lambda) = -2 \cdot (\lambda^R - \sum_{m=1}^{M} \lambda_m^U)$, is χ^2 distributed with $(k+J-1)(M-1)$ degrees of freedom (Lindeboom and van Doorslaer, 2004). 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. The null hypothesis is that there is no reporting bias. Rejection of the null hypothesis indicates that reporting thresholds are change based on an individual's SES.

If the null hypothesis is rejected, then we can determine whether the change in reporting behavior is due to cut-point or index shifts. To accomplish this, we again run Equation (3) separately for each SES group, m, except this time the β parameters are constrained to be the equal to the full, restricted model and only the cut-point parameters, $\alpha_{j,m}$, are allowed to vary by SES group. We can then run a likelihood ratio test between the restricted model and the set of constrained models, $L(\lambda) = -2 \cdot (\lambda^C - \sum_{m=1}^{M} \lambda_m^U)$, 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 that would mean that 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 $L(\lambda) = -2 \cdot (\sum_{m=1}^{M} \lambda_m^C - \sum_{m=1}^{M} \lambda_m^U)$, 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 certain 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 a SES-based index shift.

2.3.4 Detecting reporting bias using linear regression models

While the methods described above are well suited to identify different types of reporting bias, they do very little to shed light on the nature of any reporting bias that is detected and cannot be used to evaluate whether reporting bias is present in self-reported health measures that are not ordinal. To gain greater intuition concerning how SES and clinical health are related to subjective health measures, we run linear regression models that are fully interacted with measures of SES on the cardinalized measure of SAH and other self-reported measures. That is, we estimate

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

where SHR_i is a measure of self-reported health for individual i; 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 health, which is consistent with SES-based reporting bias. Thus, statistically significant coefficients on the SES variable and the interactions between SES and the objective 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 bias 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.[11] A Chow test on the interacted coefficients provides an additional test of whether the determinants of self-reported health differ by SES.

Our primary contribution to the literature here is in the way we control for objective health. In the previous literature using these regression-based techniques, researchers have chosen their objective health measure to be either a generic health measure, which could also suffer from reporting bias, or a single clinical measure, which is likely too narrow to capture all of the variability in a person's health.[12] We condition on nine different clinical health

[11] van Doorslaer and Gerdtham (2003) suggest that evidence of income-based reporting bias 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 bias and that the interaction term indicates whether the bias changes over the distribution of health.

[12] An exception to this might by Schneider et al. (2012) who use multiple objective measures to create a "disease index." However, the index itself is constructed using predicted probabilities of a binary self-reported health variable. Additionally, the objective health measures appear to be somewhat arbitrarily chosen—including items like, back pain, cancer, and mental illness all the same index.

measures, covering multiple biological systems in the body. Focusing only on clinical measurements reduces concerns of reporting bias, and including a variety of measures increases the likelihood of capturing the full variation in a person's health status. Additionally, we use these clinical health measures to create a multivariate clinical health measure based on the concept of *allostatic load*, which is discussed in the next section.

2.4 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 bias, 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 bias 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 system, and thus the biomarkers, return to normal after experiencing a period of physiological stress is called *allostatsis*. 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 and 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.

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. This technique has been previously adapted for use in health inequality measurement by Makdissi et al. (2013) and Makdissi and Yazbeck (2014). There are two potential problems with the framework in Makdissi et al. (2013) and Makdissi and Yazbeck (2014), however. First, the set of attributes used to measure health is possibly arbitrarily chosen and sample dependent, and second, the cut-points used to define an outcome as "good" or "bad" are often arbitrarily defined. 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, allostasis is a ratio-scale variable, making it compatible with condition-index measures of income-related health inequality.

3 Data

We use the 2005-2006, 2007-2008, and 2009-2010 waves of NHANES, a cross-sectional survey of health and nutritional information conducted by the CDC which combines surveys, physical examinations, and laboratory measurements. The NHANES surveys include multiple measures of self-reported health. 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.[13] 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 measures 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 with a host of adverse health conditions and is one of the leading causes of preventable death in the United States (Stommel and Schoenborn, 2010; Mokdad et al., 2004).[14] Second and third, we use resting diastolic and systolic blood pressure measurements collected during the examination. Higher blood pressure is linked with many negative cardiovascular events such as strokes, heart attacks, aneurisms, 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

[13] 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?”

[14] 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.

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 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. Finally, 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. Thus our measure of allostasis varies from zero (high-risk values for all health measures) to nine (low-risk values for all health measures). 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 government health organizations. Table 1 contains summary of 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.

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

[15] 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.

4 Results

Figures 1, 2, and 3 show the distributions of SAH, clinical measures of health, 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.

4.1 The Relationship between Income and Health Measures

4.1.1 Cluster Analysis

We first examine the relationship between income, self-reported health, and clinical measures of health using cluster analysis. When testing for reporting bias, we are particularly interested in how quickly cardinalized SAH clusters with clinical health measures relative to a SES measure. Figure 4 displays results from our cluster analysis. SAH clusters almost immediately with 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 income-to-poverty ratio. The clinical measures of health form into clusters separate from self-reported measures of health and income-to-poverty ratio. Thus, the cluster analysis suggests that self-reported measures are much more closely related to SES than they are to clinical measures of health. Put differently, the hierarchal 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 would 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.[16] Figures for these cluster results are found in the appendix.

4.1.2 Likelihood ratio tests

While the cluster analysis makes a compelling case for the existence of reporting bias in SAH, it may be criticized on the grounds that the measures of similarity are based on unconditional covariances. Therefore, we also test for reporting bias in SAH using more conventional methods. We first test for cut-point and index shifts in the ordinal SAH variable according to the methods described in Section 2.3.3. Table 3 shows the results from likelihood ratio test for reporting by the SES or demographic category in the column title. Results are taken from three sets of ordered probit regressions for each column—one restricted model, M unrestricted models, and M constrained models—where M is the number of subgroups in the variable of interest. We estimate Equation 3 using allostasis as our health variable, 11 categories of family income, 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).[17] 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

[16] 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.

[17] 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. 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.

bias. Rejecting the null hypothesis of the first likelihood ratio test indicates evidence of reporting bias. The second likelihood ratio test indicates whether bias (if present) is due to cut-point shifts, and the third test indicates whether bias is due to an index shift.[18]

We find clear evidence of reporting bias by both income and education, and our results suggest that these response difference are due to cut-point shifts. That is, individuals with identical clinical health but different levels of income or education will, on average, report different SAH values. This result is in contrast to Lindeboom and van Doorslaer (2004), and it casts doubt on the validity of using SAH to measure SES-based health inequality. Our results suggest that there is a significant amount of variation in SAH responses due to differences in income and education—variation that is independent of changes in health—which means that estimates of SES-base health inequality using this measure, or other self reported health measures, are likely bias. This is particularly problematic because there is no valid method in this context to deal with reporting bias by running separate analyses by group.

Similar to Lindeboom and van Doorslaer (2004), we find strong evidence of reporting bias by sex and age and that the differential reporting behavior is due to cut-point shifts. We additionally find evidence of cut-point shifts by race/ethnicity. This suggests that not only are SAH values are 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.

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

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

hold for individuals that are white or Hispanic, for both men and women, and for individuals of all ages. The only group for which there is not clear evidence of income-based cut-point shifts is black individuals. While the test for overall reporting bias is marginally significant, there is no clear evidence of a cut-point shift or an index shift at a 0.05 significance level.

We repeated these tests using individual biomarkers instead of allostasis and found very similar results. These tables are available upon request.

4.1.3 Characterizing SES-based reporting bias in SAH

After finding significant evidence of SES-related reporting bias in SAH responses, we turn to the question of how this bias affects SES-based inequality measurement. To answer this questions and to better understand how subjective health, SES, and objective health are related, we use the linear regression approach outlined in Equation 4. We use as dependent variables a 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 independent variables, we use allostasis, our measure of clinical health, the income-to-poverty ratio, a continuous measure of income, and controls for age, sex, race, and marital status. We fully interact our model with the income-to-poverty ratio to allow all coefficient estimates to vary by income. This allows a test for income-based reporting bias. We mean-differenced continuous variables that are squared or interacted with other continuous variables to preserve the interpretation of the coefficient on the non-interacted term.

Table 5 shows results from these regressions. A p-value of the 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 bias—not just in SAH, but in other subjective health measures as well.

In the first column, allostasis has a positive, statistically significant effect on SAH, as expected. At the mean level of income, a one unit increase in allostasis increases an individual's SAH by 0.185, which is a nearly unit elastic response at the mean ($\epsilon_H = 0.93$). Income affects individuals' reporting of SAH through a variety of channels. First, the coefficient on the income-to-poverty ratio is positive. This indicates that, at the mean, if two individuals have identical clinical health levels and different incomes, then the one with the higher income will rate her health higher than the lower income individual. Furthermore, income interacts positively with allostasis, suggesting that as income increases above the mean, the marginal effect of clinical health on SAH increases. That is, an individual with above average income will view a marginal improvement (decline) in clinical health as having a greater positive (negative) impact on her subjective health than an average income person would. On the other hand, an individual with below average income would view a marginal change in clinical health as being less consequential in her assessment of her own health than the an average income person would. 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 for sure what the "appropriate" marginal response is, it is impossible to know for sure. In either case, the effect of income on SAH conditional on objective 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.[19]

Consistent with the results in Section 4.1.2, we find evidence of reporting bias 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. With respect to age, there is negative reporting bias for people of most ages, which only becomes positive beyond age 55. Income also interacts with age. As income increases it counteracts the gradient on the marginal effect of age, and as income

[19]: ${}^{\partial SAH}/_{\partial H} = 0.0932$ at an income-to-poverty ratio of 0.5 and ${}^{\partial SAH}/_{\partial H} = 0.268$ at an income-to-poverty ratio of 4.5. Marginal effects for allostasis are statistically significant for all income values.

decreases it increases the gradient. When the income-to-poverty ratio is about one standard deviation above the mean, age-based reporting bias is essentially zero for all ages.[20] On the other hand, when the income-to-poverty ratio is one standard deviation below the mean, there is generally an even greater discrepancy between how the young and old view their health. The small and insignificant coefficient on the female/income interaction term suggests that income-based reporting bias for is similar for both men and women. This is contrary to the findings of Schneider et al. (2012) who only find income-based reporting bias for men. Also noteworthy, the coefficient on the interaction between the income-to-poverty ratio and the indicator variable for being black suggests that the marginal effect of income on SAH decreases in magnitude by 74 percent for black individuals relative to white individuals at the mean. This is consistent with the results from the likelihood ratio test in Section 4.1.2. We also see evidence of positive reporting bias from higher levels of education, and the insignificant interaction terms suggest that this effect is independent of the reporting bias we detect for income.

To test whether our results are affected by the aggregation of clinical health measures, 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 Table A1, 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 A2, and to using alternate versions of SAH, shown in Appendix Table A3.[21] 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 A4. 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

[20] At even higher income levels, the gradient reverses, eventually becoming a positive age bias that declines with age.

[21] Specifically, we examine SAH asked in another NHANES questionnaire (the HUQ) and the two SAH measures transformed into dichotomous measures.

of these results to using the original 1-5 SAH measure in a generalized ordered probit model. These results are found in Appendix Table A5. These results are consistent with our main results and show a positive income reporting bias.

There is some concern that SAH is capturing a dimension of health that allostasis (or any of the single clinical health measures) does not. One may worry that mental health, for example, is unlikely to be directly captured by clinical assessments. If a non-clinical dimension of health captured by SAH is significantly correlated with income, then our claims of that SAH suffers from income-based reporting bias may be unfounded (Lindeboom and van Doorslaer, 2004; Ziebarth, 2010). One way to assess this is to look at the more specific health self-assessments. Looking at the coefficients for the mentally healthy days regression, there is a strong positive correlation between income and mental health. It is tempting to conclude that the reporting bias is actually driven by the variation in SAH caused by changes in mental health. There are several things to note about this, however: First, even if this is true, it should completely change the conversation about what health inequality, as measured by SAH, means. Second, the income related bias found by Shmueli (2003) was strongest for the mental health component of the SF-36. Third, allostatic load and allostasis are designed to capture the physiological consequences of mental stress (McEwen, 1998). Furthermore, the coefficients on the physically healthy days regression raise serious doubts that non-clinical health is driving the remaining correlations with income. One would expect that if physically healthy days was unbiasedly reported, then it should be highly correlated with our clinical assessment of health, allostasis. This is indeed the case; however, the pattern of coefficients is remarkably similar to those on SAH, which strongly suggests the presence of SES related reporting bias. We also ran additional SAH regressions that controlled for both mentally and physically healthy days (available in Appendix Table A6) that shows results that are indistinguishable from the results discussed above. Finally, one has to ask what dimensions of health, apart from clinical health, could these self-assessments capturing and whether this is something that is policy ameliorable or even policy relevant.

4.2 Health Inequality Estimated by the Concentration Index

In Section 4.1, we find significant evidence of positive income-based reporting bias 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 objective 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 bias would not only cause health inequality measurements to be overstated for healthy populations, but it would 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 condition 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. However, both physically and mentally healthy days show a concentration index about one-third of the size of the one for SAH. Our multidimensional measure of clinical health, allostasis, suggests a much weaker relationship between health inequality and SES. 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. These results are consistent with our hypothesis that income-based reporting bias in SAH could cause income-based inequality to be overstated.

To check that these results are not due to the aggregation of different clinical measures into allostasis, we estimate condition indexes separately for each clinical measure of health. These results are displayed in Appendix Table A7. All nine clinical measures produce condition index results suggesting a lesser relationship between health inequality and income than suggested by the self-reported measures of health.

We further test the robustness of our results in a number of ways. We check that our results are robust to using the other measures of income, verify that our results are robust to excluding education controls from our regression, and also verify that our results are consistent when we do not use NHANES sample weights (results shown in Appendix Table A8). 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 A9. 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 Section 2.2 and Ziebarth (2010)).[22] Appendix Table A10 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 statistically significant.

5 Conclusion

The main objectives of this paper are to understand the impact of reporting bias in SAH on the estimation of health inequality and to suggest a simple framework for constructing an objective, multidimensional measure of health. We estimate a comprehensive set of tests

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

of SES-based reporting bias, 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 bias and that this bias 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 is biased by income, but also that this bias holds for nearly every demographic subgroup of the data. Furthermore, we also find evidence of response bias by sex, race/ethnicity, and age, with the age bias actually having an interacted effect with income bias. These results suggest that using SAH to measure the health concentration index will lead to an over-statement of health inequality. We confirm this by comparing concentration indices calculated using self-reported measures to those calculated using a clinical health measures. Our concentration index results show that income-based reporting bias 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 and 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 in other countries. Additionally, it is not clear from our analysis to what extend 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 and Jones, 2017), and surveys with clinical markers of health are also available in many other countries.[23] More broadly, our results call for an examination of SAH in the context of the relationship between socioeconomic status in health. SES-based reporting bias could prove problematic for papers which attempt to estimate a causal relationship between education, income and health if SAH is the measure of health, even if the papers find convincing exogenous variation in education or income.

[23] 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 June 19, 2017). 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. and J. Foster (2011). Counting and multidimensional poverty measurement. *Journal of Public Economics 95*(7-8), 476–487.

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

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

Bleichrodt, H. and E. van Doorslaer (2006). A welfare economics foundation for health inequality measurement. *Journal of Health Economics 25*(5), 945–957.

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

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

Cawley, J. and A. Choi (2015). Health disparities across education: The role of differential reporting error. *NBER Working Paper No. 21317*.

Clarke, P. M., U.-G. Gerdtham, M. Johannesson, K. Bingefors, and L. Smith (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.

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

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

Deaton, A. S. and C. H. Paxson (1998). Aging and inequality in income and health. *American Economic Review*, 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. and A. Zajacova (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.

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. and T. Van Ourti (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. and T. Van Ourti (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.

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

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

Gerdtham, U.-G., M. Johannesson, L. Lundberg, and D. Isacson (1999). A note on validating Wagstaff and van Doorslaer's health measure in the analysis of inequalities in health. *Journal of Health Economics 18*(1), 117–124.

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

Greene, W. H., M. N. Harris, and B. Hollingsworth (2014). Inflated responses in measures of self-assessed health. *NYU Working Paper*.

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

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

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

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

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

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

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

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

Kakwani, N. C. (1980). *Income inequality and Poverty: Methods of Estimation and Policy Applications*. Oxford University Press, New York.

Karlamangla, A. S., B. H. Singer, B. S. McEwen, J. W. Rowe, and T. E. Seeman (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., I. Kawachi, R. Glass, and D. Prothrow-Stith (1998). Income distribution, socioeconomic status, and self rated health in the United States: multilevel analysis. *BMJ 317*(7163), 917–921.

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

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

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

Levey, A. S., L. A. Stevens, C. H. Schmid, Y. Zhang, A. F. Castro, III, H. I. Feldman, J. W. Kusek, P. Eggers, F. Van Lente, T. Greene, and J. Coresh (2009). A new equation to estimate glomerular filtration rate. *Annals of Internal Medicine 150*(9), 604–612.

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

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

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

Makdissi, P. and M. Yazbeck (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. and E. Stellar (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., J. S. Marks, D. F. Stroup, and J. L. Gerberding (2004). Actual causes of death in the United States, 2000. *JAMA 291*(10), 1238–1245.

O'Donnell, O. A., A. Wagstaff, et al. (2008). *Analyzing health equity using household survey data: a guide to techniques and their implementation*. World Bank Publications.

Punj, G. and D. W. Stewart (1983). Cluster analysis in marketing research: Review and suggestions for application. *Journal of Marketing Research*, 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., C. Pfarr, B. S. Schneider, and V. Ulrich (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., B. S. McEwen, J. W. Rowe, and B. H. Singer (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., M. L. Daviglus, D. B. Garside, A. R. Dyer, P. Greenland, and J. D. Neaton (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. and C. A. Schoenborn (2010). Variations in BMI and prevalence of health risks in diverse racial and ethnic populations. *Obesity 18*(9), 1821–1826.

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

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

van Doorslaer, E. and U.-G. Gerdtham (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. and A. M. Jones (2003). Inequalities in self-reported health: validation of a new approach to measurement. *Journal of Health Economics 22*(1), 61–87.

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

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

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

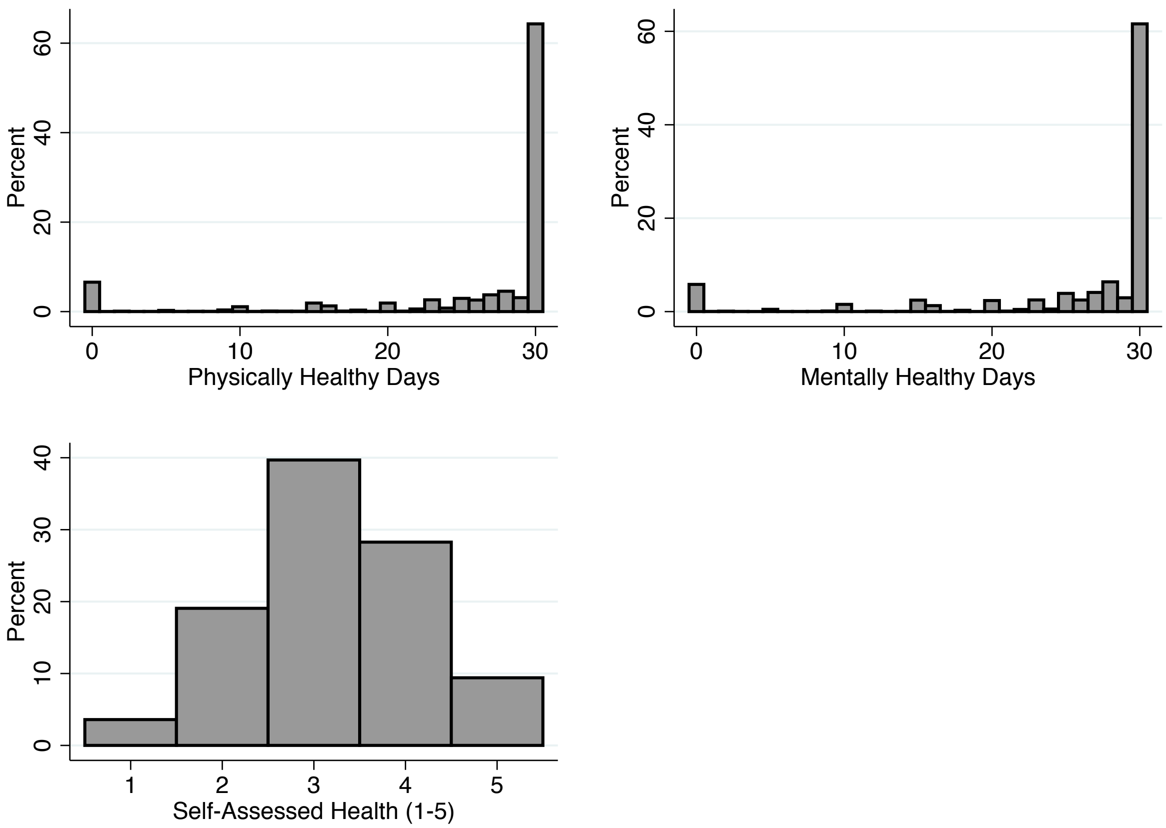
Ward, Joe H., J. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association 58*(301), 236–244.

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

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

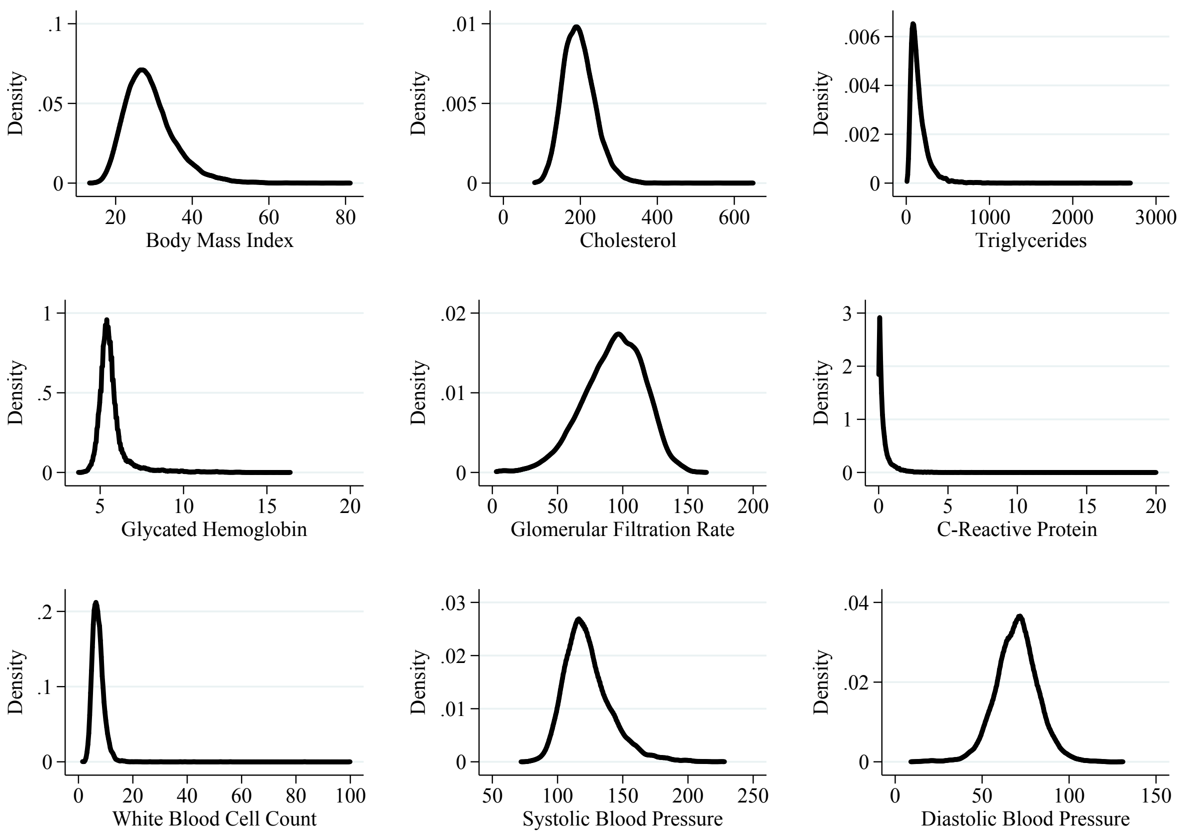
Ziebarth, N. (2010). Measurement of health, health inequality, and reporting heterogeneity. *Social Science & Medicine 71*(1), 116–124.

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



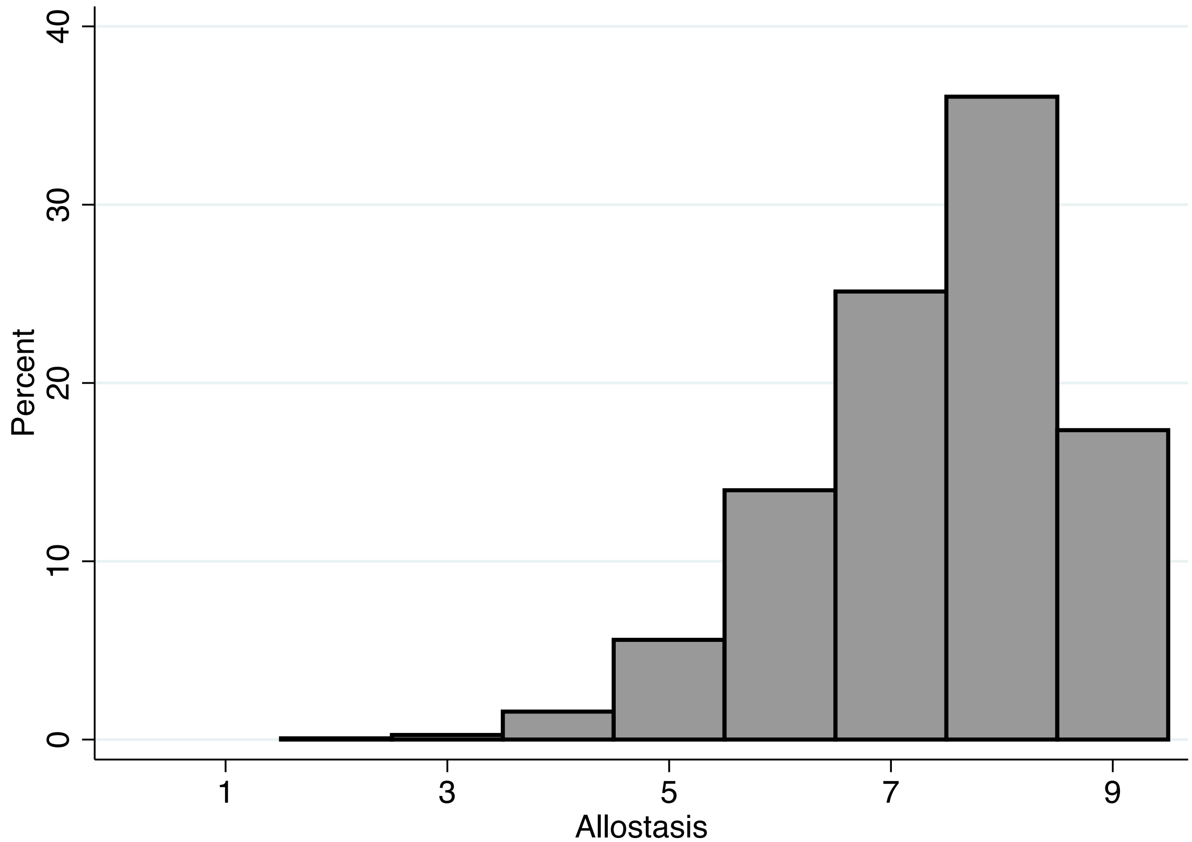
Notes: Data from the 2005-2006, 2007-2008 and 2009-2010 waves of NHANES including adults over age 20. Higher values for self-assessed health correspond to better self-assessed health.

Figure 2: Densities of Health Biomarkers



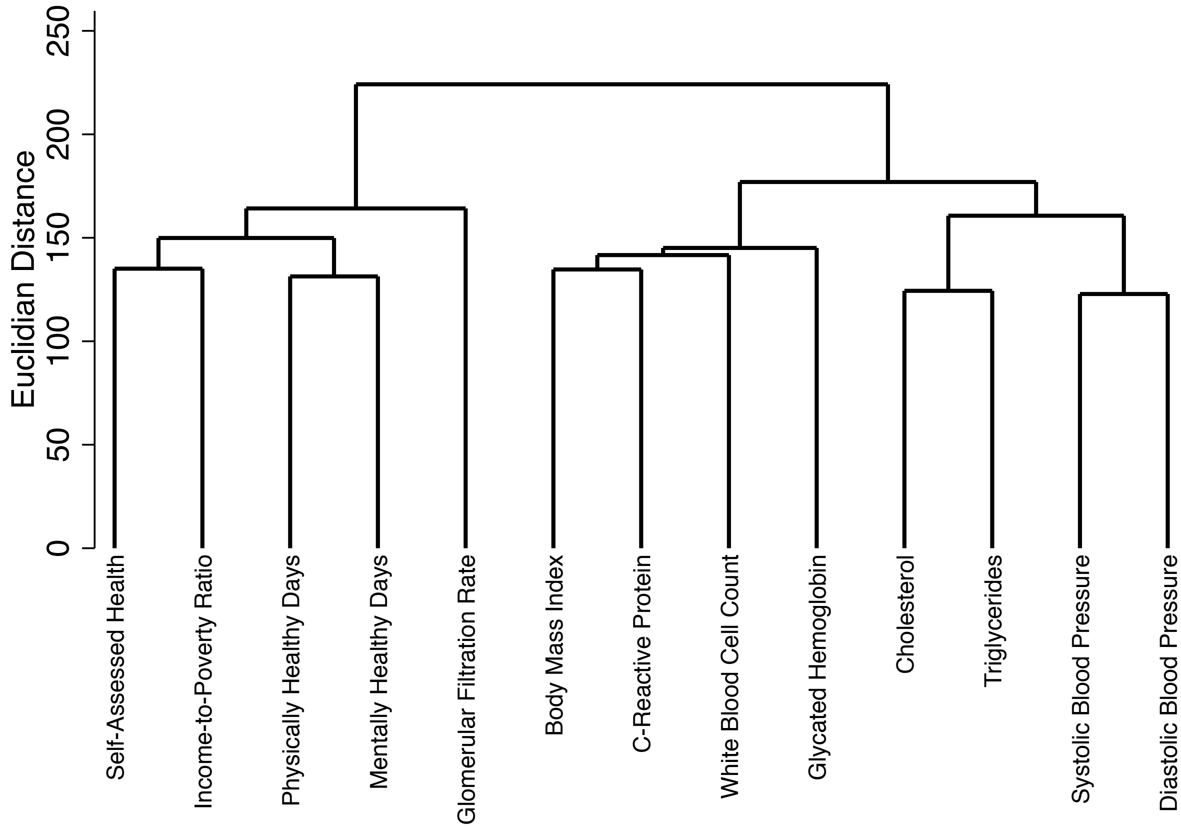
Notes: Data from the 2005-2006, 2007-2008 and 2009-2010 waves of NHANES including adults over age 20.

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.

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	≥ 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	≥ 3.5K & ≤ 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	7.525	1.176	2.000	9.000
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	-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
$-2 \cdot (\lambda^R - \sum_m \lambda_m^U)$	308.6	225.8	274.2	67.27	76.83
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
$-2 \cdot (\lambda^R - \sum_m \lambda_m^C)$	98.68	130.1	181.8	10.85	56.74
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					
$-2 \cdot (\sum_m \lambda_m^C - \sum_m \lambda_m^U)$	209.9	95.76	92.31	56.42	59.12
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 Reporting Bias by Sub

		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
$-2 \cdot (\lambda^R - \sum_m \lambda_m^U)$	208.4	134.6	61.7	71.2	108.7	130.3	103.5	94.71
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
$-2 \cdot (\lambda^R - \sum_m \lambda_m^C)$	87.6	43.3	18.8	37.6	45.6	46.3	51.84	45.57
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								
$-2 \cdot (\sum_m \lambda_m^C - \sum_m \lambda_m^U)$	120.8	91.3	42.9	33.6	63.1	84.0	51.67	49.14
Deg. of Freedom	39	33	33	33	36	36	33	33
χ^2 test p-value	**0.000**	**0.000**	0.12	0.44	**0.038**	**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 5: Regression Results of Self-Assessed Health on Income and Allostasis

	Self-Assessed Health	Physically Healthy Days	Mentally Healthy Days
Allostasis	0.185 (0.013) ***	0.356 (0.083) ***	0.249 (0.083) ***
Allostasis 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%

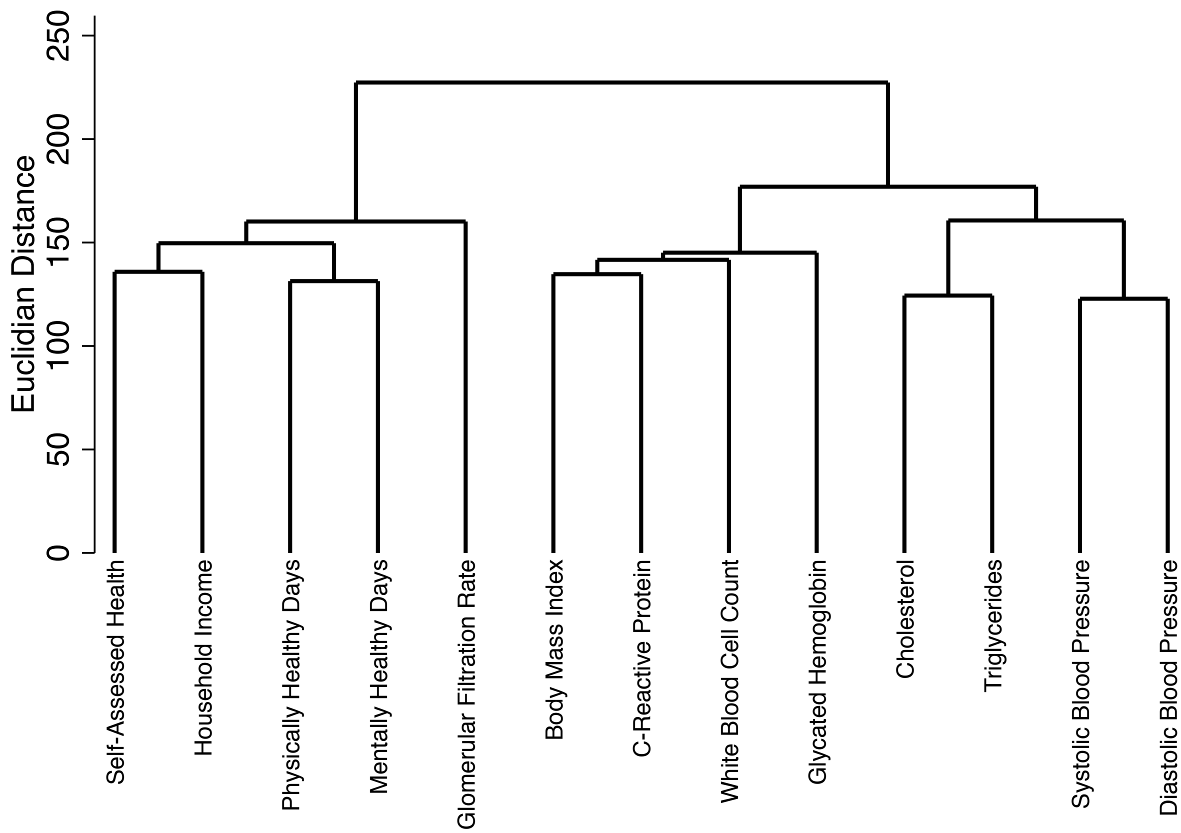
Table 6: Condition Index Results Using Income-to-Poverty Ratio

	Self-Reported Health			
	Self-Assessed Health	Physically Healthy Days	Mentally Healthy Days	Allostasis
Income-to-Poverty Ratio	0.059***	0.017***	0.020***	0.006***
	(0.006)	(0.002)	(0.002)	(0.001)
Age/100	-0.059***	-0.026***	0.025***	-0.046***
	(0.010)	(0.003)	(0.003)	(0.002)
Age/100 Squared	0.226***	0.088***	0.149***	0.071***
	(0.053)	(0.018)	(0.018)	(0.009)
Female	-0.004	-0.004***	-0.010***	0.005***
	(0.003)	(0.001)	(0.001)	(0.001)
Black	-0.022***	0.002	0.002	-0.001**
	(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	-0.004***
	(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***	0.006***
	(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	0.134
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%

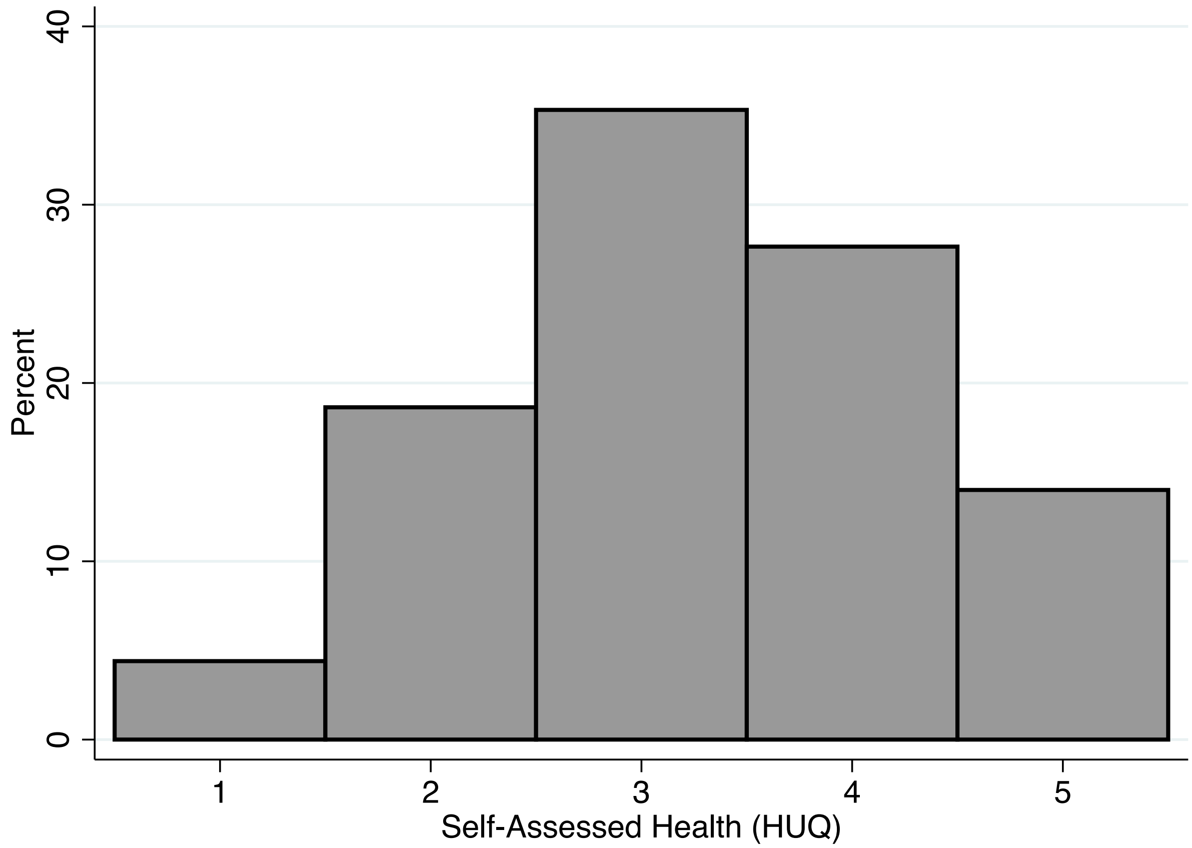
Appendix Tables

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: 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 (0.014) ***	-0.016 (0.015)	0.144 (0.015) ***	0.192 (0.015) ***	-0.025 (0.023)	0.115 (0.020) ***	0.138 (0.021) ***	0.105 (0.017) ***	0.053 (0.016) ***
Biomarker x IPR	0.063 (0.009) ***	-0.007 (0.010)	0.054 (0.009) ***	0.053 (0.010) ***	-0.032 (0.015) **	0.033 (0.011) ***	0.025 (0.016)	0.040 (0.011) ***	0.020 (0.011) *
Income-to-Poverty Ratio	0.128 (0.038) ***	0.095 (0.039) **	0.097 (0.039) **	0.090 (0.039) **	0.092 (0.039) **	0.091 (0.039) **	0.095 (0.039) **	0.101 (0.039) ***	0.099 (0.039) **
Income-to-Poverty Ratio Squared	-0.007 (0.009)	-0.004 (0.009)	-0.005 (0.009)	-0.005 (0.009)	-0.005 (0.009)	-0.003 (0.009)	-0.004 (0.009)	-0.005 (0.009)	-0.003 (0.009)
Age/100	-0.505 (0.103) ***	-0.648 (0.105) ***	-0.586 (0.104) ***	-0.324 (0.108) ***	-0.729 (0.143) ***	-0.601 (0.105) ***	-0.702 (0.105) ***	-0.410 (0.112) ***	-0.654 (0.105) ***
Age/100 x IPR	0.181 (0.070) ***	0.118 (0.071) *	0.134 (0.071) *	0.212 (0.074) ***	0.003 (0.093)	0.131 (0.071) *	0.122 (0.072) *	0.200 (0.076) ***	0.111 (0.072)
Age/100 Squared	1.243 (0.553) **	2.843 (0.562) ***	2.104 (0.553) ***	2.299 (0.555) ***	2.699 (0.554) ***	2.586 (0.551) ***	2.621 (0.550) ***	2.839 (0.550) ***	2.151 (0.578) ***
Age/100 Squared x IPR	-1.584 (0.373) ***	-1.132 (0.380) ***	-1.396 (0.375) ***	-1.205 (0.376) ***	-1.203 (0.376) ***	-1.190 (0.374) ***	-1.173 (0.375) ***	-1.111 (0.374) ***	-1.396 (0.393) ***
Female	-0.085 (0.060)	-0.148 (0.060) **	-0.152 (0.060) **	-0.148 (0.060) **	-0.152 (0.060) **	-0.149 (0.061) **	-0.151 (0.060) **	-0.143 (0.060) **	-0.144 (0.061) **
Female x IPR	0.005 (0.019)	0.032 (0.020)	0.017 (0.020)	0.028 (0.020)	0.034 (0.020) *	0.038 (0.020) *	0.033 (0.020) *	0.021 (0.020)	0.026 (0.020)
Black	0.009 (0.070)	-0.023 (0.071)	-0.036 (0.072)	-0.011 (0.071)	-0.039 (0.072)	-0.019 (0.071)	-0.057 (0.074)	-0.019 (0.071)	-0.019 (0.071)
Black x IPR	-0.063 (0.023) ***	-0.081 (0.024) ***	-0.097 (0.024) ***	-0.066 (0.024) ***	-0.074 (0.024) ***	-0.079 (0.024) ***	-0.089 (0.025) ***	-0.075 (0.024) ***	-0.082 (0.024) ***
Hispanic	-0.200 (0.066) ***	-0.211 (0.068) ***	-0.221 (0.068) ***	-0.210 (0.068) ***	-0.224 (0.069) ***	-0.219 (0.068) ***	-0.224 (0.068) ***	-0.207 (0.068) ***	-0.211 (0.068) ***
Hispanic x IPR	-0.002 (0.024)	-0.010 (0.024)	-0.003 (0.024)	0.006 (0.025)	-0.004 (0.025)	-0.004 (0.024)	-0.008 (0.024)	-0.011 (0.024)	-0.010 (0.024)
Married	0.035 (0.091)	-0.020 (0.093)	-0.010 (0.092)	-0.008 (0.092)	-0.018 (0.093)	-0.012 (0.092)	-0.006 (0.092)	-0.015 (0.092)	-0.017 (0.092)
Married x IPR	-0.020 (0.032)	0.006 (0.033)	0.004 (0.033)	-0.002 (0.033)	0.006 (0.033)	0.002 (0.033)	0.001 (0.033)	0.001 (0.033)	0.005 (0.033)
Widowed	0.025 (0.128)	0.003 (0.131)	0.004 (0.131)	0.007 (0.131)	0.009 (0.131)	0.006 (0.131)	0.021 (0.131)	0.022 (0.131)	0.008 (0.131)
Widowed x IPR	-0.038 (0.049)	-0.025 (0.050)	-0.019 (0.050)	-0.027 (0.050)	-0.028 (0.050)	-0.026 (0.050)	-0.029 (0.050)	-0.035 (0.050)	-0.027 (0.050)
Divorced	-0.039 (0.111)	-0.059 (0.114)	-0.064 (0.113)	-0.053 (0.114)	-0.060 (0.114)	-0.059 (0.114)	-0.032 (0.114)	-0.060 (0.114)	-0.062 (0.114)
Divorced x IPR	-0.009 (0.042)	0.010 (0.044)	0.015 (0.043)	0.005 (0.043)	0.012 (0.044)	0.013 (0.043)	0.010 (0.043)	0.009 (0.043)	0.012 (0.043)
College Degree	0.760 (0.125) ***	0.719 (0.129) ***	0.725 (0.128) ***	0.695 (0.128) ***	0.723 (0.129) ***	0.718 (0.128) ***	0.699 (0.128) ***	0.727 (0.129) ***	0.725 (0.129) ***
College Degree x IPR	-0.025 (0.037)	0.008 (0.038)	-0.002 (0.038)	0.004 (0.038)	0.005 (0.038)	0.004 (0.038)	0.003 (0.038)	-0.001 (0.038)	0.005 (0.038)
Some College	0.316 (0.079) ***	0.301 (0.080) ***	0.284 (0.080) ***	0.278 (0.079) ***	0.308 (0.081) ***	0.310 (0.080) ***	0.291 (0.080) ***	0.295 (0.081) ***	0.301 (0.080) ***
Some College x IPR	0.007 (0.030)	0.008 (0.031)	0.016 (0.031)	0.014 (0.030)	0.005 (0.031)	0.005 (0.031)	0.007 (0.030)	0.011 (0.031)	0.010 (0.031)
High School Diploma	0.213 (0.072) ***	0.204 (0.073) ***	0.201 (0.073) ***	0.192 (0.072) ***	0.210 (0.073) ***	0.202 (0.072) ***	0.203 (0.072) ***	0.202 (0.073) ***	0.208 (0.073) ***
High School Diploma x IPR	-0.003 (0.030)	-0.002 (0.031)	-0.002 (0.031)	-0.001 (0.030)	-0.006 (0.031)	-0.000 (0.031)	-0.003 (0.031)	-0.001 (0.031)	-0.004 (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 A2: Regression Results of Self-Reported Health on Income and Allostasis
No Sample Weights

	Self-Assessed Health	Physically Healthy Days	Mentally Healthy Days
Allostasis	0.162 (0.011) ***	0.321 (0.070) ***	0.177 (0.066) ***
Allostasis x IPR	0.041 (0.007) ***	0.005 (0.044)	-0.022 (0.041)
Income-to-Poverty Ratio	0.111 (0.031) ***	0.587 (0.182) ***	0.999 (0.183) ***
Income-to-Poverty Ratio Squared	-0.003 (0.007)	-0.186 (0.043) ***	-0.165 (0.043) ***
Age/100	-0.369 (0.089) ***	-4.130 (0.503) ***	4.730 (0.484) ***
Age/100 x IPR	0.152 (0.059) **	0.748 (0.330) **	0.550 (0.313) *
Age/100 Squared	1.919 (0.464) ***	18.050 (2.747) ***	22.412 (2.546) ***
Age/100 Squared x IPR	-1.313 (0.312) ***	-8.620 (1.762) ***	-8.706 (1.596) ***
Female	-0.131 (0.045) ***	-1.214 (0.314) ***	-2.747 (0.311) ***
Female x IPR	0.005 (0.016)	0.164 (0.092) *	0.345 (0.091) ***
Black	0.003 (0.062)	1.092 (0.420) ***	1.691 (0.433) ***
Black x IPR	-0.078 (0.022) ***	-0.257 (0.127) **	-0.432 (0.134) ***
Hispanic	-0.175 (0.052) ***	1.433 (0.379) ***	3.301 (0.380) ***
Hispanic x IPR	-0.015 (0.021)	-0.545 (0.128) ***	-0.837 (0.124) ***
Married	0.033 (0.070)	0.220 (0.413)	0.928 (0.470) **
Married x IPR	0.005 (0.026)	-0.013 (0.133)	-0.192 (0.151)
Widowed	0.007 (0.101)	0.669 (0.785)	1.546 (0.725) **
Widowed x IPR	0.017 (0.042)	-0.362 (0.287)	-0.664 (0.250) ***
Divorced	-0.016 (0.083)	-0.406 (0.575)	-0.057 (0.612)
Divorced x IPR	0.013 (0.033)	-0.043 (0.189)	-0.082 (0.198)
College Degree	0.667 (0.101) ***	1.298 (0.544) **	2.301 (0.521) ***
College Degree x IPR	-0.002 (0.031)	0.057 (0.174)	-0.383 (0.161) **
Some College	0.357 (0.064) ***	0.219 (0.435)	0.154 (0.446)
Some College x IPR	-0.006 (0.025)	0.170 (0.161)	-0.045 (0.153)
High School Diploma	0.165 (0.057) ***	0.701 (0.431)	0.825 (0.423) *
High School Diploma x IPR	0.015 (0.025)	-0.009 (0.170)	-0.230 (0.155)
Adj. R^2	0.118	0.043	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 A3: 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	0.185 (0.013) ***	0.158 (0.012) ***	0.039 (0.004) ***	0.034 (0.004) ***
Allostasis x IPR	0.044 (0.008) ***	0.031 (0.008) ***	-0.001 (0.002)	-0.003 (0.002)
Income-to-Poverty Ratio	0.100 (0.038) ***	0.128 (0.036) ***	0.049 (0.010) ***	0.058 (0.010) ***
Income-to-Poverty Ratio Squared	-0.006 (0.009)	-0.025 (0.008) ***	-0.010 (0.002) ***	-0.012 (0.002) ***
Age/100	-0.240 (0.107) **	-0.597 (0.100) ***	-0.172 (0.027) ***	-0.226 (0.027) ***
Age/100 x IPR	0.195 (0.072) ***	0.184 (0.066) ***	0.070 (0.016) ***	0.068 (0.016) ***
Age/100 Squared	2.104 (0.548) ***	2.375 (0.529) ***	0.697 (0.145) ***	0.709 (0.145) ***
Age/100 Squared x IPR	-1.240 (0.371) ***	-1.122 (0.340) ***	-0.400 (0.087) ***	-0.580 (0.087) ***
Female	-0.132 (0.059) **	-0.104 (0.055) *	-0.033 (0.017) *	-0.016 (0.017)
Female x IPR	0.009 (0.020)	0.006 (0.018)	0.008 (0.004) *	0.003 (0.005)
Black	-0.031 (0.070)	0.131 (0.067) **	-0.049 (0.022) **	-0.043 (0.022) *
Black x IPR	-0.074 (0.023) ***	-0.103 (0.022) ***	-0.005 (0.007)	-0.004 (0.007)
Hispanic	-0.215 (0.067) ***	-0.102 (0.063)	-0.105 (0.021) ***	-0.107 (0.021) ***
Hispanic x IPR	0.001 (0.024)	-0.041 (0.022) *	0.012 (0.006) *	0.016 (0.006) **
Married	0.008 (0.092)	0.087 (0.080)	0.022 (0.022)	0.031 (0.022)
Married x IPR	-0.004 (0.033)	-0.008 (0.029)	0.000 (0.007)	-0.005 (0.007)
Widowed	0.029 (0.130)	0.045 (0.119)	0.045 (0.042)	0.043 (0.041)
Widowed x IPR	-0.024 (0.050)	0.011 (0.049)	-0.012 (0.013)	-0.010 (0.013)
Divorced	-0.061 (0.112)	0.090 (0.101)	-0.007 (0.032)	0.036 (0.031)
Divorced x IPR	0.014 (0.042)	-0.014 (0.037)	0.001 (0.010)	-0.010 (0.009)
College Degree	0.703 (0.127) ***	0.548 (0.110) ***	0.190 (0.029) ***	0.224 (0.028) ***
College Degree x IPR	-0.005 (0.038)	0.022 (0.036)	-0.015 (0.009)	-0.022 (0.009) **
Some College	0.285 (0.080) ***	0.294 (0.078) ***	0.142 (0.024) ***	0.126 (0.024) ***
Some College x IPR	0.014 (0.031)	0.019 (0.031)	-0.010 (0.009)	-0.005 (0.009)
High School Diploma	0.204 (0.072) ***	0.203 (0.071) ***	0.125 (0.025) ***	0.121 (0.025) ***
High School Diploma x IPR	-0.003 (0.031)	-0.009 (0.031)	-0.017 (0.009) *	-0.016 (0.009) *
Adj. R^2	0.111	0.115	0.126	0.130
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 A4: 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	0.142***	0.233***	0.186***	0.156***	0.218***	0.166***	0.326***	0.147***	0.152***	0.163***	0.231***	0.156***
	(0.019)	(0.017)	(0.014)	(0.026)	(0.021)	(0.016)	(0.033)	(0.013)	(0.013)	(0.023)	(0.020)	(0.021)
Allostatic Load x IPR	0.047***	0.042***	0.045***	0.021	0.049***	0.041***	0.034	0.030***	0.022**	0.021	0.062***	0.047***
	(0.012)	(0.011)	(0.009)	(0.017)	(0.013)	(0.010)	(0.024)	(0.009)	(0.010)	(0.015)	(0.014)	(0.014)
Income-to-Poverty Ratio	0.144***	0.068	0.105**	0.014	0.102*	0.127*	0.156*	0.076*	0.120**	0.095	0.160**	0.020
	(0.053)	(0.052)	(0.042)	(0.066)	(0.053)	(0.069)	(0.086)	(0.042)	(0.048)	(0.077)	(0.065)	(0.059)
Income-Poverty Ratio Squared	-0.002	-0.009	-0.007	0.012	-0.005	0.001	0.041	-0.016*	-0.016	-0.028*	-0.019	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	-0.549***	0.156	-0.180	-0.974***	-2.986*	0.724	0.164	-0.372***	-0.412***	-0.162	-0.489***	-0.102
	(0.149)	(0.154)	(0.116)	(0.220)	(1.564)	(0.510)	(0.262)	(0.116)	(0.116)	(0.200)	(0.185)	(0.170)
Age/100 x IPR	0.246**	0.178*	0.195**	0.074	0.200	-0.129	0.277	0.127	0.097	0.150	0.188	0.218*
	(0.098)	(0.106)	(0.077)	(0.148)	(0.999)	(0.333)	(0.194)	(0.084)	(0.093)	(0.133)	(0.122)	(0.121)
Age/100 Squared	2.327***	1.917**	2.174***	0.449	-6.292	-0.779	-0.111	2.860***	2.841***	2.275**	1.165	2.582***
	(0.790)	(0.752)	(0.594)	(1.168)	(4.598)	(1.717)	(1.524)	(0.572)	(0.571)	(0.949)	(1.006)	(0.896)
Age/100 Squared x IPR	-1.266**	-1.132**	-1.281***	-1.143	-0.993	-0.517	-0.382	-1.171***	-1.052**	-1.065*	-1.759***	-0.862
	(0.528)	(0.519)	(0.398)	(0.848)	(2.946)	(1.185)	(1.056)	(0.416)	(0.476)	(0.629)	(0.656)	(0.641)
Black	0.102	-0.127			0.039	-0.127	-0.439*	0.000	0.021	-0.100	0.051	-0.039
	(0.111)	(0.090)			(0.109)	(0.082)	(0.251)	(0.074)	(0.075)	(0.126)	(0.123)	(0.120)
Black x IPR	-0.071**	-0.084***			-0.083**	-0.061**	0.001	-0.069***	-0.105***	-0.048	-0.098**	-0.082**
	(0.036)	(0.030)			(0.036)	(0.029)	(0.065)	(0.026)	(0.029)	(0.040)	(0.041)	(0.041)
Hispanic	-0.237**	-0.195**	-0.193***		-0.274***	-0.099	-0.316	-0.182**	-0.216***	-0.276*	-0.177*	-0.240**
	(0.097)	(0.092)	(0.069)		(0.085)	(0.110)	(0.259)	(0.071)	(0.073)	(0.162)	(0.101)	(0.103)
Hispanic x IPR	0.017	-0.018	-0.003		0.001	-0.007	0.023	-0.010	0.008	0.023	-0.033	0.030
	(0.036)	(0.031)	(0.024)		(0.031)	(0.038)	(0.065)	(0.027)	(0.032)	(0.053)	(0.035)	(0.039)
Married	0.181	-0.197	-0.019	0.081	-0.016	0.024	0.009	-0.041	-0.012	-0.109	-0.016	0.110
	(0.128)	(0.130)	(0.107)	(0.154)	(0.113)	(0.161)	(0.310)	(0.096)	(0.098)	(0.178)	(0.157)	(0.146)
Married x IPR	-0.048	0.047	-0.004	0.028	0.009	-0.030	-0.041	0.038	0.010	-0.008	-0.023	0.026
	(0.045)	(0.047)	(0.036)	(0.055)	(0.040)	(0.063)	(0.080)	(0.038)	(0.042)	(0.060)	(0.058)	(0.052)
Widowed	-0.047	-0.143	0.014	-0.030	-0.467	-0.002	-0.282	0.022	0.006	-0.171	-0.106	0.300
	(0.202)	(0.171)	(0.151)	(0.203)	(0.398)	(0.180)	(0.571)	(0.130)	(0.140)	(0.247)	(0.224)	(0.210)
Widowed x IPR	-0.046	0.020	-0.035	0.157*	0.045	-0.033	0.021	-0.003	-0.006	-0.003	0.032	-0.080
	(0.075)	(0.066)	(0.055)	(0.091)	(0.097)	(0.073)	(0.155)	(0.052)	(0.062)	(0.086)	(0.097)	(0.075)
Divorced	0.035	-0.182	-0.078	0.020	-0.032	-0.065	-0.132	-0.083	-0.052	-0.081	-0.198	0.059
	(0.166)	(0.151)	(0.132)	(0.180)	(0.156)	(0.175)	(0.399)	(0.117)	(0.119)	(0.220)	(0.188)	(0.183)
Divorced x IPR	-0.004	0.038	0.016	0.018	-0.028	0.015	0.000	0.046	0.012	-0.053	0.041	0.062
	(0.062)	(0.057)	(0.048)	(0.063)	(0.059)	(0.070)	(0.106)	(0.048)	(0.053)	(0.076)	(0.075)	(0.070)
College Degree	0.839***	0.570***	0.758***	0.340	0.834***	0.571***			0.959***	0.436*	0.802***	0.813***
	(0.183)	(0.176)	(0.139)	(0.242)	(0.177)	(0.180)			(0.141)	(0.259)	(0.228)	(0.187)
College Degree x IPR	-0.032	0.017	-0.013	0.021	-0.069	0.042			-0.129***	0.091	-0.070	-0.019
	(0.054)	(0.052)	(0.041)	(0.074)	(0.058)	(0.051)			(0.049)	(0.074)	(0.066)	(0.058)
Some College	0.346***	0.242**	0.277***	0.413**	0.333***	0.311***		0.272***	0.338***	0.267	0.228*	0.359***
	(0.116)	(0.109)	(0.089)	(0.165)	(0.110)	(0.116)		(0.080)	(0.087)	(0.164)	(0.134)	(0.123)
Some College x IPR	-0.012	0.033	0.020	-0.068	-0.033	0.031		0.020	-0.024	0.026	0.020	-0.002
	(0.044)	(0.041)	(0.033)	(0.060)	(0.048)	(0.040)		(0.031)	(0.038)	(0.059)	(0.052)	(0.048)
High School Diploma	0.242**	0.181*	0.259***	-0.132	0.218**	0.190**		0.195***	0.155*	0.268*	0.247**	0.111
	(0.112)	(0.093)	(0.081)	(0.138)	(0.110)	(0.093)		(0.072)	(0.079)	(0.143)	(0.117)	(0.120)
High School Diploma x IPR	-0.004	-0.015	-0.017	0.090	-0.013	0.006		0.001	0.018	-0.014	-0.051	0.057
	(0.046)	(0.040)	(0.033)	(0.059)	(0.050)	(0.038)		(0.031)	(0.039)	(0.056)	(0.051)	(0.052)
Female			-0.097	-0.275**	-0.219***	-0.022	-0.286	-0.122**	-0.138**	-0.047	-0.138	-0.206**
			(0.066)	(0.118)	(0.085)	(0.082)	(0.235)	(0.061)	(0.065)	(0.116)	(0.100)	(0.095)
Female x IPR			0.006	-0.012	0.013	-0.007	0.032	0.012	0.016	-0.011	-0.002	0.042
			(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	0.140	0.111	0.077	0.097	0.125	0.059	0.074	0.085	0.100	0.131	0.110
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 A5: 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	0.004 (0.001) ***	0.034 (0.003) ***	0.065 (0.005) ***	0.030 (0.003) ***
Allostasis x IPR	0.000 (0.000)	0.006 (0.002) ***	0.010 (0.003) ***	0.006 (0.002) ***
Income-to-Poverty Ratio	0.001 (0.002)	0.025 (0.008) ***	0.040 (0.015) ***	0.010 (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	-0.024 (0.007) ***	-0.143 (0.024) ***	-0.111 (0.041) ***	-0.004 (0.024)
Age/100 x IPR	0.005 (0.006)	0.022 (0.018)	0.033 (0.027)	0.033 (0.015) **
Age/100 Squared	0.082 (0.033) **	0.478 (0.128) ***	0.995 (0.215) ***	0.269 (0.121) **
Age/100 Squared x IPR	-0.066 (0.045)	-0.470 (0.121) ***	-0.465 (0.139) ***	-0.134 (0.081) *
Female	0.007 (0.000)	-0.002 (0.000)	-0.001 (0.000)	-0.036 (0.000)
Female x IPR	0.002 (0.001)	0.008 (0.004) *	0.011 (0.007)	-0.001 (0.004)
Black	-0.003 (0.000)	0.073 (0.000)	-0.028 (0.000)	-0.033 (0.000)
Black x IPR	-0.001 (0.001)	-0.008 (0.007)	-0.022 (0.009) **	-0.012 (0.005) **
Hispanic	0.007 (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	-0.007 (0.000)	-0.028 (0.000)	0.004 (0.000)	-0.015 (0.000)
Married x IPR	0.001 (0.002)	0.003 (0.007)	0.004 (0.011)	-0.003 (0.007)
Widowed	-0.022 (0.000)	-0.009 (0.000)	0.003 (0.000)	-0.035 (0.000)
Widowed x IPR	0.000 (0.002)	-0.006 (0.011)	0.014 (0.021)	-0.009 (0.010)
Divorced	-0.001 (0.000)	0.004 (0.000)	-0.002 (0.000)	0.003 (0.000)
Divorced x IPR	0.000 (0.002)	0.002 (0.009)	0.002 (0.015)	0.002 (0.008)
College Degree	-0.039 (0.000)	-0.141 (0.000)	-0.007 (0.000)	0.120 (0.000)
College Degree x IPR	-0.002 (0.003)	-0.014 (0.009)	-0.006 (0.013)	0.003 (0.010)
Some College	-0.020 (0.000)	-0.101 (0.000)	0.006 (0.000)	0.047 (0.000)
Some College x IPR	0.000 (0.002)	-0.010 (0.008)	-0.003 (0.012)	0.006 (0.007)
High School Diploma	-0.009 (0.000)	-0.054 (0.000)	0.005 (0.000)	0.037 (0.000)
High School Diploma x IPR	0.000 (0.002)	-0.013 (0.007) *	-0.008 (0.013)	0.004 (0.008)
Pseudo R^2	0.080			
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 A6: Regression Results of Self-Assessed Health on Income and Allostasis: Robustness Tests

	Controlling for Physically and Mentally Healthy Days	Education Interactions Only	Income and Education Interactions
Allostasis	0.172***	0.084***	0.102***
	(0.012)	(0.023)	(0.024)
IPR x Allostasis	0.043***		0.023**
	(0.008)		(0.009)
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)		
Allostasis x BA Degree		0.255***	0.202***
		(0.040)	(0.045)
Allostasis x Some College		0.128***	0.096***
		(0.032)	(0.033)
Allostasis x HS Degree		0.072**	0.051
		(0.035)	(0.035)
Income-to-Poverty Ratio	0.074**	0.100***	0.087**
	(0.038)	(0.023)	(0.040)
Income-to-Poverty Ratio Squared	0.001		-0.004
	(0.009)		(0.009)
College Degree	0.643***	-1.278***	-0.772*
	(0.123)	(0.398)	(0.436)
College Degree x IPR	-0.003	-0.005	-0.005
	(0.037)	(0.037)	(0.039)
Some College	0.290***	-0.961***	-0.664**
	(0.077)	(0.312)	(0.318)
Some College x IPR	0.008	0.001	0.011
	(0.030)	(0.030)	(0.031)
High School Diploma	0.184***	-0.273	-0.078
	(0.071)	(0.340)	(0.343)
High School Diploma x IPR	0.001	-0.003	-0.002
	(0.030)	(0.031)	(0.031)
Adj. R^2	0.143	0.112	0.113
Num Obs	11751	11751	11751

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 A7: Condition 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	0.031***	0.004***	0.002**	0.046***	0.013***	0.003***	-0.001
	(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***	-0.108***	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***	0.551***	0.070***	0.050***	0.459***	0.041**	-0.023***	0.297***
	(0.014)	(0.013)	(0.046)	(0.008)	(0.010)	(0.116)	(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	0.025***	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	0.070	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 A8: Condition Index Results
Other Income Measures, Education, and Sample Weights

	Self-Reported Health			
	Self-Assessed Health	Physically Healthy Days	Mentally Healthy Days	Allostasis
Income-to-Poverty Ratio Incl. Educ. (Weighted)	0.059***	0.017***	0.020***	0.006***
	(0.006)	(0.002)	(0.002)	(0.001)
Income-to-Poverty Ratio Incl. Educ. (Unweighted)	0.072***	0.018***	0.020***	0.006***
	(0.006)	(0.002)	(0.002)	(0.001)
Family Income (Category) Incl. Educ. (Weighted)	0.052***	0.018***	0.019***	0.005***
	(0.006)	(0.002)	(0.002)	(0.001)
Family Income (Category) Incl. Educ. (Unweighted)	0.063***	0.019***	0.020***	0.005***
	(0.006)	(0.002)	(0.002)	(0.001)
Household Income (Category) Incl. Educ. (Weighted)	0.053***	0.017***	0.019***	0.005***
	(0.006)	(0.002)	(0.002)	(0.001)
Household Income (Category) Incl. Educ. (Unweighted)	0.064***	0.019***	0.020***	0.006***
	(0.006)	(0.002)	(0.002)	(0.001)
Income-to-Poverty Ratio No Educ. (Weighted)	0.094***	0.020***	0.024***	0.009***
	(0.006)	(0.002)	(0.002)	(0.001)
Income-to-Poverty Ratio No Educ. (Unweighted)	0.114***	0.023***	0.024***	0.009***
	(0.005)	(0.002)	(0.002)	(0.001)
Family Income (Category) No Educ. (Weighted)	0.086***	0.021***	0.023***	0.008***
	(0.006)	(0.002)	(0.002)	(0.001)
Family Income (Category) No Educ. (Unweighted)	0.103***	0.023***	0.023***	0.008***
	(0.005)	(0.002)	(0.002)	(0.001)
Household Income (Category) No Educ. (Weighted)	0.087***	0.020***	0.023***	0.008***
	(0.006)	(0.002)	(0.002)	(0.001)
Household Income (Category) No Educ. (Unweighted)	0.104***	0.023***	0.024***	0.009***
	(0.005)	(0.002)	(0.002)	(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%.

Table A9: Erreygers Index Results Using Income-to-Poverty Ratio

	Self-Reported Health			
	Self-Assessed Health	Physically Healthy Days	Mentally Healthy Days	Allostasis
Income-to-Poverty Ratio	0.075***	0.060***	0.068***	0.024***
	(0.008)	(0.008)	(0.008)	(0.005)
Age/100	-0.076***	-0.092***	0.088***	-0.196***
	(0.013)	(0.011)	(0.011)	(0.007)
Age/100 Squared	0.290***	0.313***	0.520***	0.305***
	(0.067)	(0.065)	(0.064)	(0.040)
Female	-0.005	-0.014***	-0.036***	0.022***
	(0.004)	(0.003)	(0.004)	(0.002)
Black	-0.028***	0.006	0.007	-0.006**
	(0.004)	(0.004)	(0.005)	(0.003)
Hispanic	-0.028***	0.005	0.025***	-0.010***
	(0.005)	(0.004)	(0.005)	(0.003)
Married	0.000	0.005	0.009	-0.001
	(0.006)	(0.005)	(0.006)	(0.003)
Widowed	-0.009	-0.011	-0.004	-0.018***
	(0.009)	(0.010)	(0.010)	(0.006)
Divorced	-0.006	-0.017**	-0.015*	-0.006
	(0.008)	(0.007)	(0.008)	(0.005)
College Degree	0.090***	0.024***	0.026***	0.027***
	(0.006)	(0.006)	(0.006)	(0.004)
Some College	0.038***	0.007	0.002	0.001
	(0.005)	(0.005)	(0.006)	(0.003)
High School Diploma	0.023***	0.007	0.001	0.001
	(0.005)	(0.006)	(0.006)	(0.003)
Adj. R^2	0.084	0.035	0.051	0.134
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. 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 A10: Condition 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)	0.075***	0.098***	0.136***	0.143***	0.206***	0.221***
	(0.008)	(0.009)	(0.010)	(0.010)	(0.000)	(0.000)
Income-to-Poverty Ratio Incl. Educ. (Unweighted)	0.082***	0.105***	0.156***	0.169***	0.215***	0.236***
	(0.007)	(0.007)	(0.010)	(0.010)	(0.000)	(0.000)
Family Income (Category) Incl. Educ. (Weighted)	0.067***	0.089***	0.119***	0.128***	0.183***	0.200***
	(0.008)	(0.009)	(0.011)	(0.010)	(0.000)	(0.000)
Family Income (Category) Incl. Educ. (Unweighted)	0.071***	0.095***	0.137***	0.151***	0.192***	0.209***
	(0.006)	(0.007)	(0.010)	(0.010)	(0.000)	(0.000)
Household Income (Category) Incl. Educ. (Weighted)	0.069***	0.084***	0.120***	0.129***	0.192***	0.205***
	(0.008)	(0.009)	(0.010)	(0.010)	(0.000)	(0.000)
Household Income (Category) Incl. Educ. (Unweighted)	0.073***	0.093***	0.137***	0.153***	0.199***	0.215***
	(0.006)	(0.007)	(0.010)	(0.010)	(0.000)	(0.000)
Income-to-Poverty Ratio No Educ. (Weighted)	0.121***	0.149***	0.185***	0.194***	0.310***	0.314***
	(0.007)	(0.008)	(0.009)	(0.009)	(0.000)	(0.000)
Income-to-Poverty Ratio No Educ. (Unweighted)	0.129***	0.159***	0.228***	0.240***	0.331***	0.338***
	(0.006)	(0.007)	(0.009)	(0.009)	(0.000)	(0.000)
Family Income (Category) No Educ. (Weighted)	0.111***	0.138***	0.169***	0.178***	0.289***	0.304***
	(0.008)	(0.008)	(0.010)	(0.010)	(0.000)	(0.000)
Family Income (Category) No Educ. (Unweighted)	0.116***	0.145***	0.206***	0.219***	0.301***	0.310***
	(0.006)	(0.007)	(0.009)	(0.009)	(0.000)	(0.000)
Household Income (Category) No Educ. (Weighted)	0.111***	0.133***	0.168***	0.178***	0.296***	0.308***
	(0.007)	(0.008)	(0.010)	(0.010)	(0.000)	(0.000)
Household Income (Category) No Educ. (Unweighted)	0.118***	0.144***	0.205***	0.220***	0.309***	0.318***
	(0.006)	(0.007)	(0.009)	(0.009)	(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; and (4) Untransformed SAH from the Hospital Utilization and Access to Care Questionnaire. Robust standard errors are shown in parenthesis, and stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%.