Entropy-Based Measures of Multidimensional Health Inequality*

Erik T. Nesson†
Joshua J. Robinson‡

Abstract

We propose a new measure of multidimensional health inequality based on a well-known class of entropy-based inequality measures. We aggregate a set of health attributes into a measure of overall health by minimizing the relative entropic distance between the multivariate distribution of attributes and the distribution of the summary measure. We calculate both general inequality and income-based inequality using generalized entropy, which offers many advantages in comparison to the Gini and condition index family of inequality measures, including accounting for the multidimensional nature of health, allowing flexibility in the degree of complementarity between health attributes, utilizing the information from the entire distribution of health and income rather than income or health ranks, and making normative assumptions transparent and intuitive. We demonstrate the contribution of our inequality measure using data from the Health and Retirement Surveys which contains multiple measures of clinical health, self-reported physical and mental health, and measures of income and wealth.

Keywords: Population Health, Health Inequality, Information Theory, Entropy
JEL Codes: C43, I12, I18

*This work is supported by a grant from the Robert Wood Johnson Foundation. 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 and participants at the 2014 Southern Economic Association Meetings for excellent comments and Daniel Mowan and August Clevenger for excellent research assistance.

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

‡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. However, while a large literature examines the statistical properties of health inequality measurements, some large issues remain. 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. Moreover, these health measurements are almost all univariate, whereas health status is multivariate. The multivariate distribution of health status makes traditional inequality measurements such as the Gini coefficient and concentration index unattractive, as they are univariate measures of inequality.

In this paper, we propose the use of a measure of multidimensional health inequality based on a well-known class of generalized entropy (GE) inequality measures first used in the measurement of multidimensional welfare inequality (Maasoumi, 1986). Our "two step" method first aggregates multiple health attributes into a measure of overall health by minimizing the relative entropic divergence between the multivariate distribution of the attributes and the distribution of the summary measure, thereby minimizing the information lost through aggregation. These ideal summary functions also share a functional form with common utility functions. This allows social preferences regarding the complementarity or substitutability of health attributes to be transparent, changeable, and comparable—a quality that we argue is particularly important when measuring health inequality. Next, we measure inequality by calculating the entropic distance between the shares of the summary measure and the uniform distribution. This measure of inequality allows for differing social preferences regarding inequality aversion at the lower end of the health distribution, is decomposable to within group versus between group inequalities, and is decomposable by the different underlying health attributes under certain assumptions. Finally, the use of the GE family of inequality indices can by justified using the principles of information theory concerning distributional

divergence, without the need for justifying arbitrary assumptions about a hypothetical social welfare function.

We compare our methodology to traditional univariate measures of health inequality and also examine the effects of using different measures of health in the calculation of health inequality. We use data from the 2006 and 2008 waves of the Health and Retirement study, which contain individuals' health as measured by SAH, other self-reported measures of health, many biomarkers of health, and detailed income and demographic information. We examine both the relationship between SAH, other self-reported measures of health, and biomarkers of health, and examine whether these different measures of health consistently estimate the degree of health inequality and the relationship between health inequality and SES. We find a positive relationship between SAH and our other measures of health, although there is a lot of variation between the different measures of health. We find that the level of health inequality, the connection between health inequality and SES, and the contribution of other demographic characteristics to health inequality varies greatly depending on the measure of health used. Specifically, we find that SAH and self-reported mental health suggest a much stronger relationship between SES and health inequality than either using multidimensional health status or other self-reported measures of health. Lastly, we find that there is very little inequality between different demographic groups.

Our paper has two main goals: to demonstrate the value of using a true multidimensional index for health by both examining the methods objective merits and discussing some of the short comings of more commonly used indices and to demonstrate the merits of using the GE family of inequality indices in the context of health inequality. Traditional inequality metrics are typically derived for only a single attribute and cannot account for the multivariate nature of health status. We construct a measure of health using multiple objective indicators of 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). Using multivariate general entropy (MGE), introduced by Maasoumi (1986), we aggregate biomarkers commonly used in the calculation of allostatisis into a summary measure of health. While not applied to the measurement of health, MGE has been widely used to measure economic welfare, inequality, and poverty (e.g. Hirschberg et al., 1991, 2001; Maasoumi and Nickelsburg, 1988; Maasoumi and Jeong, 1985; Lugo, 2007; Maasoumi and Lugo, 2008; Decancq and Lugo, 2013). We then use the MGE health index to construct a measure of multivariate health inequality based on the GE approach also outlined in Maasoumi (1986). Our results suggest that there is less inequality in health overall than suggested by previous research. Moreover, social preferences for inequality aversion do not alter health inequality, suggesting few differences in inequality across the distribution of health. Lastly, we find little evidence of inequality between race, gender, or income groups, indicating that most health inequality lies within these groups.

Furthermore, in addition to measuring inequality in a univariate fashion, many prior studies measure overall health status using subjective SAH evaluated on an ordinal scale. Although SAH is likely correlated with an individual's true unobservable health status, many studies show that SES, culture, or other unrelated variables may also affect an individual's SAH (Crossley and Kennedy, 2002; Johnston et al., 2009; Greene et al., 2014; Zajacova and Dowd, 2011). The implications of these studies are especially concerning for the health inequality literature, where the main research question concerns the relationship between SES and health. We test this by calculating health inequality based on SAH to inequality measures using other, more objective, self-assessed measures of health and multidimensional measures of health derived from MGE. Using both GE inequality and the most common measure of health inequality, the concentration index, we find that SAH-based inequality measurements suggest a much stronger connection to socioeconomic factors relative to other, more objective, self-assessed measures of health and multidimensional measures of health derived from MGE.

The rest of this paper proceeds as follows. Section 2 reviews the relevant literature

measuring health inequality, Section 3 describes our construction of a measure of health and the general entropy inequality methodology and its uses in the measurement of economic welfare and inequality, Section 4 describes the data source, Section 5 summarizes our results, and Section 6 concludes.

2 Background

Most of the literature examining the relationship between SES and health inequality uses the concentration index, which measures the degree of SES-related health inequality. Related to the Lorenz curve and Gini coefficient, the concentration index can be calculated from a graph with SES on the x-axis and cumulative health shares on the y-axis. A 45 degree line represents equality, where every individual receives a uniform allocation of health. The concentration index is defined as twice the area between the line of equality and the line which traces the actual distribution of health and SES (Kakwani, 1977; Wagstaff et al., 1991; O'Donnell et al., 2008).

When using individual-level data, the concentration index can be estimated with a simple linear regression,

$$2\sigma_r^2 \left(\frac{h_i}{\mu}\right) = \beta_0 + \beta_1 r_i + \boldsymbol{x_i}\boldsymbol{\beta_2} + u_i. \tag{1}$$

In this equation, r_i is an individual's fractional SES rank, σ_r^2 is the variance of the fractional rank, h_i is an individual's health level, μ is the mean health status, and $\boldsymbol{x_i}$ and $\boldsymbol{\beta_2}$ represent vectors of other explanatory variables and coefficients (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]

Numerous papers have made contributions and modifications to this original framework. For example, Wagstaff et al. (2003) extends the framework of the concentration index to show how changes in health inequality can be decomposed into changes in the means and

[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.

inequalities of the determinants of health inequality and changes in the size of the effects of the determinants on health inequality. 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).

2.1 Known problems with the concentration index

Despite its ubiquity in the health inequality literature, many potential methodological problems with the concentration index have been identified. Bleichrodt and van Doorslaer (2006) derive the properties of the social welfare function implied if one accepts the concentration index as the "best" measure of SES-related health inequality and find social preferences must satisfy the *the principle of income related health transfers*. This basically means that a SES rank preserving transfer of health from a person of higher SES to a person of lower SES cannot decrease social welfare, which they identified as a contentious assumption. Indeed, Bleichrodt et al. (2012) find little to no evidence that experimental subjects had a preference for decreasing the correlation between health and SES. Another challenge was raised by Fleurbaey and Schokkaert (2009) who point out the concentration index is not necessarily a measure of SES-related inequality because the measurement is likely affected by (potentially sample specific) correlations depending on the variables used. Ignoring the other sources of inequality increases the likelihood of mismeasurement.

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. Wagstaff (2009) challenged Erreygers (2009a) on the no-

tion that his proposed correction was an absolute measure of inequality, not a relative one, and thus answers a completely different question about health inequality. In response, Erreygers (2009b) note that the terms "absolute" and "relative" inequality lose their traditional meaning when dealing with bounded variables. Thus, an inequality measure should be chosen based on it satisfying desirable properties given the underlying measurement scale. He argues that the so called "mirror property"—that inequality indices be invariant to measuring health achievement or shortfall—is critical when variables are bounded. The debate has continued and become more nuanced (Erreygers and Van Ourti, 2011a; Lambert and Zheng, 2011; Wagstaff, 2011; Erreygers and Van Ourti, 2011b). Finally, Kjellsson and Gerdtham (2013) clarified 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 judgement as it is a technical one.

The second issue pertains to the measurement framework for health inequality. To connect SES to health inequality, the general framework compares the ranks of individuals' SES to health levels. But examining the ranks of SES does not account for the difference between SES levels. This becomes a much larger issue when examining health inequality and SES across different geographic regions, as different regions could have very different wealth distributions. For example, the 99th percentile of income in Alabama is $665,177 while the 99th percentile of income in Connecticut is $2,178,817.[2] Concentration index-based measurements of income levels and health examine only an individual's income percentile relative to that specific geography, so they would treat as identical an individual in the top 1% of the income distribution, regardless of that individual's actual level of income. Differences in the income distribution may be inferred by differences in the concentration index, i.e. a larger share of health inequality seems to be related to income in certain geographies, but this is only an indirect connection.

[2]Data taken from `http://www.epi.org/multimedia/unequal-states-interactive`, accessed November 7th, 2014.

The final issue is that the concentration index framework presumes that only inequality related to SES matters. One could argue that significant inequalities exist in health care systems that are unrelated to SES. An inequality in health for race or gender, for example, may persist despite the resources at the individual's disposal. Furthermore, SES-related inequality may not follow the gradient pattern demanded by the concentration index. In many parts of the world, and the United States in particular, there exist significant gaps in health care coverage and availability that may have a larger impact on the lower-middle and middle classes. These very important forms of inequality is not measurable in the Wagstaff framework and can only be accomplish indirectly.

In this paper, we propose a methodology to address the above shortcomings. We develop a continuous, biological measure of health based on the concept of allostatic load, a measure of the physiological burden of adapting to life's demands. As we detail below, previous research has found allostatic load to be a strong predictor of mortality and morbidity. We then develop a multidimensional measure of health inequality which both accounts for the multidimensional nature of health and incorporates different social preferences regarding inequality aversion for individuals with the lowest health status. Lastly, we decompose our measure of health inequality by gender, race, and income to find the within and between group contributions of inequality. The GE decomposition allows for SES groups to be defined based on absolute differences, which incorporates explanatory information that is lost to the SES ranking used in the concentration index.

3 Methods

3.1 On the problem of self-assessed health

Although health is a latent and multidimensional characteristic, many studies use individuals' SAH, often reported on a scale of one to five or one to ten (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 is potentially problematic for a number of reasons. The primary concern that must be address by all researchers that use SAH for the measurement of inequality is how to transform it into cardinal or, preferably in the case of the standard concentration index, ratio scale variable. This inevitably involves a number of potentially arbitrary assumptions to be made about the nature of the relationship between an individuals subjective evaluation of her health and her actual health status. 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 status—a problem commonly referred to as *reporting heterogeneity*. We first discuss three common methods of transforming SAH and the potential problems associated with each. Then we discuss the reliability of SAH and the problems associated with reporting heterogeneity—detailing the methods we will use to assess the extent to which these problems are present in our own sample.

3.1.1 Cardinalizing ordinal self-assessments

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-off 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) 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 short-comings. 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-off points used. An additional problem is that the bounds of the concentration index depend on the mean of the transformed binary variable, and Wagstaff (2005) propose a correction to rescale the

bounds to lie between -1 and 1. However, the value of the concentration index is still mean-dependent in this case, which makes cross-country comparison or comparisons using different values of health problematic. Erreygers (2009a) showed that this issue and others exist for any health variable with a finite upper bound. The the subsequent debate is discussed in detail in Section 2.1. Finally, Ziebarth (2010) shows that dichotomizing SAH may overstate the degree of health inequality by as much as ten fold compared to more objectives measures of health.

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 using the cumulative distribution of the ordinal SAH values to parameterize the log-normal distribution. The obvious short-coming of this method is that health my 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 QALY score to measure health outcomes.

Finally, van Doorslaer and Jones (2003) compare multiple methods of transforming SAH to a more comprehensive health assessment (the Canadian Health Utility Index, or 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.[3] This technique involves first obtaining values for cumulative health stock for each value of the ordered health measure from any survey that contains both a generic health metric and ordinal SAH. The low and high CDF values for each ordinal score are then applied to the data of the population in which the researcher wants to assess health inequality, and predicted health scores are calculated using an interval regression. While appealing in many dimensions, this methods suffers from two large issues: First, Ziebarth (2010) shows that the concentration indices associated with health measures calculated in

[3]It is worth noting that while generic measures, like the HUI, are more comprehensive than a typical 5-point health self assessment, they are measures of health that are still inherently based on a person own subjective evaluations.

the this manner are highly sensitive to the generic health measure chosen and are almost identical to the concentration indices calculated from the generic health measures on which the predicted scores are based, which typically come from a completely different sample. That is, if a researcher used the HUI to predict health scores for her own sample, the resulting inequality scores will most likely reflect the inequality in the Canadian population sampled for the HUI measurement, not necessarily the inequality in her own sample. Secondly, van Doorslaer and Jones (2003) advocate using both demographic data and any indicators of SES available as regressors in the interval regression to add precision to predicted health scores. However, this could theoretically force a correlation between SES and the predicted health scores by construction—at a minimum, this approach is likely to overstate the degree of dependence between SES and health. van Doorslaer and Jones (2003) even note that the interval regression approach tends to produce higher estimates of SES-related inequality than any other measure.

In this study, we 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. We, therefore, believe that the log-normalization transformation is the least likely to unduly influence our results.

3.1.2 Measurement Error & Reporting bias

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 socioeconomic status—systematically self-report their health in different ways. This is referred to as *reporting heterogeneity* or *reporting bias*. With regard to using SAH to evaluate a health concentration index, SES-related reporting bias is particularly problematic. There is a great deal of evidence in the literature that reporting bias is problematic across age and gender, but the evidence for SES-related reporting bias is mixed.

Sometimes researchers test for reporting bias by conditioning on a generic health measure (see footnote 3). Lindeboom and van Doorslaer (2004) find no evidence of reporting bias by income or education, by conditioning SAH on the Canadian HUI as an "objective" health measure. Layes et al. (2012), on the other hand, finds evidence that lower income individuals overstate their health status, also conditioning on the HUI. However, Shmueli (2003) use a multiple indicators–multiple causes (MIMIC) model to estimate the latent health variable and find that SAH and SF-36 (a generic health measure) responses are positively biased by income.[4]

Other researchers choose to condition on a clinical measure of health to test for reporting bias; however, results are still mixed. 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. Finally 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[5] Contrary to this, Ziebarth (2010) uses to grip strength as an objective health measure to show that SAH and SF-12 (a generic health measure) overestimate the degree of health inequality.

Another way to test for reporting bias is to see if the predictive power of SAH on mortality

[4]It is worth noting that Shmueli (2003) uses a different type of SAH, 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").

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

depends on other observable characteristics. However, evidence for SES-related reporting bias is still mixed. 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. Finally, Jürges (2008) finds some evidence income based heterogeneity, but only for women.

The breadth of methodological difficulties associated with the use of SAH as an aggregate measure of health motivates the need for a true multidimensional measure of health. In the next section we introduce our multidimensional health measure based on the medical concept of allostasis and aggregated using MGE methods.

3.2 The measurement of multidimensional health

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 (at least in theory) in a single measure. Thus, even when more objective or clinical measures of health are available, they are often not favored over self-assessed measures of health because a specific measure of objective health is often considered to narrow to capture the broad scope of an individual's health status. This might explain why there are conflicting results in the literature, even when an objective measure of health is used Johnston et al. (2009); Ziebarth (2010). Generic health measures are also problematic because they are inherently based on self reports. Indeed, there is evidence of reporting bias present in several generic measures (Ziebarth, 2010; Shmueli, 2003).

In order to truly overcome this dilemma, clinically measured health must be used to avoid reporting bias, and multiple clinical measures must be collected to avoid defining health too narrowly. This creates another problem in that the data must be aggregated into a single, cardinal (or preferably ratio scale) measure in order to assess inequality. A proposed solution to a similar problem is to measure the "width" rather than the "depth" of the problem. Alkire and Foster (2011) used this idea to 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 adapted for use in health inequality measurement Makdissi et al. (2013) and Makdissi and Yazbeck (2014).

There are two potential problems with this framework, however: the set of attributes used to measure health is likely conveniently chosen and sample dependent and the cut-points are arbitrary. Furthermore, assigning cutoff points for each attribute does not account for any variation in the attribute except for whether the attribute level falls above or below the cutoff. We base our measurement of health on the concept of allostatic load. First developed by McEwen and Stellar (1993), allostatic load is a measure of physiological levels across a number of biological systems relevant to disease risk. Allostatic load includes measures of four main systems: (1) the hypothalamic-pituitary-adrenal axis (HPA 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, such as white blood cell count or C-reactive protein levels (Seeman et al., 2001; Geronimus et al., 2006). In traditional measures of allostatic load, individuals receive one point for each measure for which they fall over a high-risk threshold, and the cumulative number of points is the measure of allostatic load. In addition to the shortcomings of this type of "count" aggregation function mentioned above, linearly summing attributes implicitly assumes that the attributes are perfectly substitutable. This assumption is particular problematic when considering the measurement of health.

We propose a methodology to measure individual and population health by combining

the concepts of MGE and allostatic load. Originally developed by Maasoumi (1986), our methodology chooses a summary measure of population health that minimizes the information loss cause by using a summary measure to represent the multivariate distribution of underlying attributes. Entropy-based aggregation methods are advantageous both for their desirable information-preserving properties and their intuitive link to the economic theories of production and utility, which we will discuss further below. While not applied to the measurement of health, MGE has been widely used to measure economic welfare, inequality, and poverty (e.g. Hirschberg et al., 1991, 2001; Maasoumi and Nickelsburg, 1988; Maasoumi and Jeong, 1985; Lugo, 2007; Maasoumi and Lugo, 2008; Decancq and Lugo, 2013).

Our measure of divergence is *relative entropy*, which measures the information lost by mischaracterizing the probability distribution of a random variable. This measure, also known as Kullback-Leibler distance, can be thought of as the distance between the distributions $p(x)$ and $q(x)$ (Cover and Thomas, 2006).[6] Theil (1967a) is the first to use relative entropy in economic welfare analysis. He uses the entropic distance between a distribution of a population's individual income shares and the uniform distribution of income as a measure of inequality. Several variations on Theil's measures have since been proposed in the literature (see Decancq and Lugo (2013) for a literature review).

3.2.1 Deriving the MGE aggregator

Formally, if X_{if} is the value of some attribute, $f = 1 \dots M$, for an observation, $i = 1 \dots N$, then $X_i = (X_{i1}, X_{i2}, \dots, X_{iM})$ is the row vector of values for all attributes for observation i and $X^f = (X_{1f}, X_{2f}, \dots, X_{Nf})$ if the column vector of values of attribute f for all observations. We regard X^f as the sample distribution of an attribute.[7] The optimal summary function transforms the M-vector of attributes into a single value for each observation, $S_i = h(X_i)$, in such a way as to minimize the generalized multivariate relative entropy

[6]Relative entropy is not technically a true distance measure in the mathematical sense because it fails the triangle inequality and is not symmetric. However, if a true metric is needed there are metric divergence measures in the GE family (e.g. Granger et al., 2004)

[7]The values, X_{if}, are standardized such that their distributions have the same support.

between S_i and the multivariate distribution of X^f's. That is, S_i minimizes the function:

$$\begin{aligned} D_\beta(S, X; \alpha) &= \sum_{f=1}^{M} \alpha_f \left\{ \sum_{i=1}^{N} S_i \left[\left(\frac{S_i}{X_{if}} \right)^\beta - 1 \right] \Big/ \beta(\beta+1) \right\} \\ &= \sum_f \alpha_f \left\{ \sum_i S_i \log(S_i/X_{if}) \right\}, \quad \text{if } \beta = 0 \qquad (2) \\ &= \sum_f \alpha_f \left\{ \sum_i X_{if} \log(X_{if}/S_i) \right\}, \quad \text{if } \beta = -1 \end{aligned}$$

This is essentially a weighted sum of the divergence of S_i from the set of corresponding X_i. Here the α's represent the relative importance (or weight) assigned to each attribute. Maasoumi (1986) shows that the summary functions are

$$S_i \propto \left[\sum_{f=1}^{M} \delta_f X_{if}^{-\beta} \right]^{-1/\beta} \qquad (3a)$$

$$\propto \prod_{f=1}^{M} X_{if}^{\delta_f}, \quad \text{if } \beta = 0 \qquad (3b)$$

$$\propto \sum_{f=1}^{M} \delta_f X_{if}, \quad \text{if } \beta = -1 \qquad (3c)$$

where $\delta_f = \alpha_f / \sum_f \alpha_f$ is each attribute's relative weight. By construction, these summary functions are as close to the original multivariate distribution as possible. The functions in Equation (3) are the weighted harmonic mean (for $\beta \neq 0, -1$), the weighted geometric mean (for $\beta = 0$), and the weighted arithmetic mean for (for $\beta = -1$), respectively.

The value of the summary functions in Equation (3) depends on the choices of β and the vector of relative weights. A major advantage of this framework is the characterization of the parameter β. The general form of S_i ($\beta \neq 0, -1$) has the same functional form as the constant elasticity of substitution utility (or production) function, the second form ($\beta = 0$) is synonymous with a Cobb-Douglas function, and the third form ($\beta = -1$) is a perfect substitutes function. Thus $\sigma = 1/(1+\beta)$ is the constant elasticity of substitution between

attributes across individuals.

With this characterization in mind, one can think of changes in β as altering the degree of substitutability or complementarity between attributes. A β less than zero implies a higher degree of substitutability between attributes, and as $\beta \to -1$, the elasticity of substitution approaches infinity and attributes are combined as though they are perfectly substitutable in the calculation of health. On the other hand, $\beta > 0$ implies a degree of complementarity between attributes. Attributes are considered approaching perfect complements, a preference for all attributes to rise in perfect proportion to one another, as $\sigma \to 0$ and $\beta \to \infty$. This property of the summary function is particularly important because it is often overlooked in favor of the simplicity of an arithmetic mean. However, the characterization above makes it clear that a linear aggregation method is only appropriate when your health attributes are considered to be perfectly substitutable. Put differently, whether or not you believe the assumption of perfectly substitutable attributes is appropriate, you will impart this property to your data by using a linear aggregator—intentionally or unintentionally.

3.3 The measurement of multidimensional health inequality

We utilize a class of general entropy (GE) measures to evaluate inequality in health. Let S_i represent the latent measure of an individual's or group's health status.[8] For the vector $S = \{S_1, S_2, \ldots, S_N\}$ of all individual health statuses, inequality is given by

$$I_\gamma(S) = \sum_{i=1}^{N} p_i \left[\left(S_i^* / p_i \right)^{1+\gamma} - 1 \right] \Big/ \gamma(1+\gamma) \tag{4}$$

Where $S^* = {}^{S_i}/_{\sum_i S_i}$ is the portion of total health granted to individual i, and p_i is individual i's share of the population (typically ${}^1/_N$ for individuals). The parameter γ determines which part of the distribution of health receives the most weight in the inequality calculation. At $\gamma \to 0$, all parts of the distribution are given equal weight. As $\gamma \to -\infty$, more weight is

[8] We discuss how to construct the measure S_i in Section 3.2

given to inequality in the lower portion of the distribution. Likewise, more emphasis can be placed on the upper portion of the distribution by using increasingly positive values of γ.

The GE inequality metric is algebraically equivalent to the Kullback-Liebler divergence between the distribution of health shares the uniform allocation of health. It is often intuitively useful to think of the GE measure as a sum of pairwise divergences from the mean value (of relative shares). Thus, the inequality measure is zero when health is evenly distributed to all individuals in the population and greater than zero otherwise. It is worth noting here that although entropy-based measures of inequality are typically characterized by divergence from the uniform distribution, it would be possible to construct a measure of divergence from any relevant distribution. With regard to inequality in health, examples of other interesting distribution comparisons are the distributions of income, health care expenditures, or age. We discuss these possibilities further in Section 6.

The special cases of $\gamma = 0$ and $\gamma = -1$ converge to Theil's (1967b) first and second measures, respectively, which are given by

$$I_0(S) = \sum_{i=1}^{N} S_i^* \log {}^{S_i^*}\!/_{p_i} \quad \text{for} \quad \gamma = 0 \tag{5a}$$

$$I_{-1} = \sum_{i=1}^{N} p_i \log {}^{p_i}\!/_{S_i^*} \quad \text{for} \quad \gamma = -1 \tag{5b}$$

Furthermore, when $\gamma = 1$ this measure is ordinally equivalent to the coefficient of variation and the Herfindahl index, and for $-\gamma = \epsilon$ the GE family is ordinally equivalent to the Atkinson family of inequality measures with inequality aversion parameter ϵ. Thus, the GE family of inequality measures is flexible enough to use in any application.

When evaluating inequality, it is often desirable to be able to decompose by group in order to examine the potential sources of inequality. Consider any division of the total population of N individuals into G different (mutually exclusive) groups, g_r, with N_r people in each group (for $r = 1, \ldots, G$). We can decompose total inequality, given in Equation (4),

into "between" group and "within" group inequality as follows

$$
\begin{aligned}
I_\gamma(S) = \quad & \textstyle\sum_{r=1}^{G} p_r \left[\left(S_r^* / p_r \right)^{1+\gamma} - 1 \right] \Big/ \gamma(1+\gamma) && \text{(between group } I\text{)} \\
& + \textstyle\sum_{r=1}^{G} p_r^{-\gamma} (S_r^*)^{1+\gamma} \cdot \left\{ \frac{1}{N_r} \left[\left(\frac{S_i / S_r}{1 / N_r} \right)^{1+\gamma} - 1 \right] \Big/ \gamma(1+\gamma) \right\} && \text{(within group } I\text{)}
\end{aligned}
\tag{6}
$$

The first line in Equation (6) is the portion of inequality due to differences in allocations between groups, and the second line is the weighted sum of individual inequality within each group. For any group g_r, the value $p_r = {}^{N_r}/_{N}$ is the portion of the population that belongs to the group, $S_r^* = {}^{S_r}/_{\sum_{i=1}^{N} S_i}$ is the group's population share of S, and $S_r = \sum_{i \in g_r} S_i$ is the total amount of S allocated to the group.[9]

This decomposition adds a valuable assessment tool when the population can be meaningfully subdivided. It allows for more targeted policy prescriptions by revealing information about whether inequality is driven by differences between different groups of people and whether there is greater inequality within some groups relative to others. Regarding the application to health inequality, searching for between and within group inequalities by race, gender, and socioeconomic groups could all have interesting policy implications. We examine the results of these decompositions in Section 5.[10]

4 Data

We use data from the 2006 and 2008 waves of the Health and Retirement Study, a longitudinal data set examining the health and well-being among older individuals in the United

[9] In a seminal paper on the methods of health inequality measurement, Wagstaff et al. (1991) criticizes entropy measures of inequality because they are "insensitive to the socioeconomic dimensions to inequalities in health. What matters in the [entropy measure] is simply how each socioeconomic group's share of the population's health compares with its population share, not how this disparity compares with the socioeconomic group's socioeconomic status." The concerns can be addressed by a combination of decomposing the inequality measure by groups based on SES and by adjusting the parameter of relative inequality aversion, γ, to place more weight on the lower tail of the distribution.

[10] When using the MGE methodology discussed in Section 3.2 to calculate S_i, it is also possible to decompose inequality by each of the attributes. While this is potentially quite interesting in the context of health inequality, it is beyond the scope of the this paper.

States. We excluding respondents over age 85 from our analysis. The HRS includes detailed information on earnings and wealth, measures of self-reported physical and mental health, and demographic information. Importantly for our analysis, the HRS also collected measures of clinical health. We combine information on clinical measures of health from the original HRS data with the RAND HRS data, a cleaned version of the most popular HRS variables.

We focus on total household wealth as our measure of socioeconomic status. Earnings is another popular variable to measure socioeconomic status, but as many individuals in the HRS are retired, we choose to examine total household wealth. Our results are generally robust to using as measures of socioeconomic status household wealth excluding secondary homes, individual earnings, and household earnings. We also collect information on many demographic variables, including gender, age, race (African American or Caucasian), Hispanic ethnicity, marital status (married, widowed, divorced, in a relationship, or never married), and education (less than a high school education, high school degree or equivalent, some college but not a bachelor's degree, a bachelor's degree, or an advanced degree above a bachelor's).

The RAND HRS also includes self-reported measures of health. First, the RAND HRS includes an ordinal self-reported health measure, where the values range from 1 (Excellent) to 5 (Poor). Second, we use an index of daily living activities which includes five activities: bathing, eating, dressing, walking across a room, and getting in or out of bed. This measure sums all dimensions where respondents report at least some difficulty in the respective activity. Third, we use as a measure of self-reported mental health the Center for Epidemiologic Studies Depression (CESD) scale. The CESD measures five negative mental health indicators: whether the respondent reported depression, that everything is an effort, sleep is restless, they felt alone, they felt sad, or could not get going. The CESD then subtracts two positive indicators: whether the respondent felt happy or enjoyed life. Finally, we use a count of medical conditions. We transform each self-reported health measure so that higher values correspond to better health.

During the 2006 and 2008 waves, HRS collected biomarker information from 6517 and 4347 respondents, respectively. HRS collected respondents' total cholesterol, HDL cholesterol, Glycosylated hemoglobin (HbA1c), C-reactive protein, and Cystatin C levels. We additionally measure body mass index (BMI), collected from the RAND HRS. 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). Higher cholesterol levels and lower HDL cholesterol levels are linked to many heart conditions (Stamler et al., 2000; Criqui et al., 1993). Glycosylated hemoglobin (HbA1c) levels are a measure of average blood sugar levels over the past 120 days (Trivelli et al., 1971). C-reactive protein is a marker of inflammation from a number of causes, including infections, inflammatory diseases, injury, and cancers.

Table 1 shows summary statistics for our sample, and Figures 1 and 2 show the distributions of SAH and our health biomarkers. The distribution of SAH appears to be roughly normal, although slightly skewed to the right. Many of the biomarkers have similar distributions, with a few individuals having very high levels of the biomarkers, compared to the mean.

As the biomarkers are measured in different units, we first standardize the values using a procedure adapted from the Human Development Index, $X_{if} = \frac{Y_{if} - \min\{Y_f\}}{\max\{Y_f\} - \min\{Y_f\}}$. Additionally, since an increase in any biomarker except HDL cholesterol is indicative of a decrease in the level of health, we then take the inverse of each biomarker.

5 Results

5.1 Comparison of self-assessed health and biomarkers

Figures 1 through 5 display the raw characteristics of our standard and constructed measures of health. Figure 1 contains histograms of our four measures of self-reported health. Other the the five point SAH measure, these distributions are heavily skewed left, which suggests

that these measures are not sensitive enough to pick up less serious conditions (this could be a good or bad thing, depending on your research question). The sample densities of the six health biomarkers are also heavily skewed (but to the right), which can be seen in Figure 2. Figure 3 shows the relationship between our constructed measure of multidimensional health (using various values of β) and SAH. These plots clearly show that there is little to no correlation between these objective health measures (as proxied by our summary index, S) and SAH, and the lack of relationship is not changed by varying the degree of complementarity or substitutability used to combine the biomarkers.

Figures 4 and 5 show how the choice of substitutability between attributes during the aggregation stage can affect the value of the summary health measure. Figures 4 shows the empirical probability density functions for our summary health measure with each separate panel displaying a different density function for a different value of β. As β increases, it forces a greater degree of complementarity between the attributes when they are combined. In this context, this amounts to a greater emphasis being placed on the worst of the nine health indicators. Given the dramatic amount of change in the density function, it is obvious that a change in the degree of complementarity can dramatically alter the summary health measure. Similarly, Figure 5 shows the scatter plot and correlation between a summary index computed allowing perfect substitutability ($\beta = -1$ which corresponds to an arithmetic mean) and summary indices constructed using greater degrees of complementarity between health attributes. However, despite the noticeable divergence in S_i as γ increases, the correlation never drops below 0.9.

5.2 Health inequality estimated by the concentration index

In order to get a base-line estimate for how much of a difference it makes to use multivariate measures of health to measure health inequality, we first calculate inequality using the Erreygers transformation of the concentration index (Erreygers, 2009a).[11] Table 2 shows

[11]The Erreygers transformation is appropriate here because all of our variables are bounded.

the estimates for SES-related inequality for four different self-reported health measures and our multivariate heath measure calculated for five different values of β. The concentration index, which measures SES-related inequality in each health measure, is the coefficient estimate for "Total Wealth", that is the relative rank of the individuals according to their estimated wealth. The most interesting thing to notice is that the estimated inequality is up to 10 times higher for the self-assessed measures than for any of the multivariate indices. This is consistent with the results of Nesson and Robinson (2016), who find strong evidence of SES-related reporting bias. It is, of course, possible that the self-assessed measures are simply picking up a facet of health that is not reflected in the clinical biomarkers and is particularly sensitive to SES. While formally testing for reporting bias is beyond the scope of this paper, we will note that an unobserved health dimension as a cause for the dramatic differences in inequality measurement between the multidimensional health indices and the self-assessed measures. The pattern holds across all of the self-assessed measures, which presumably represent different dimension of health. Furthermore, the arguably most objective of the self-assessed measures, the index of medical conditions, shows the least potential for reporting bias. Overall, the level of SES-related inequality appears to be rather inconsequential (despite being statistically significant) when health is measured using a combination of objective, clinical health measures.

Table 3 shows a variety of robustness checks for the concentration index results, including using several alternative measures to rank SES and both including and excluding sample weights. There are no notable differences between the results and those discussed above when some measure of wealth is used. However, SES-related inequality drops noticeably across the board—becoming statistically insignificant and approaching zero for the MGE health measures—when a measure of earnings, rather than wealth, is used to rank individuals' SES. This may be due to the well documented sensitivity of the concentration index to the measure of SES chosen for the ranking.

5.3 Health inequality estimated by entropy

We now turn to the results from measuring inequality using GE divergence, which are summarized in Tables 4 and 5. In each table, the columns show how overall inequality changes as we vary the degree of relative inequality aversion. As γ becomes increasingly negative, the measure becomes less tolerant of inequality, and the overall inequality measure tends to increase. For each health measure, an overall level of inequality is given followed beneath by the decomposition results using four different group categorizations. For example, when $\gamma = -0.5$ we can see that the overall inequality associated with SAH is 0.36. Note that these measures of GE inequality are not directly comparable to the inequality measures computed using the concentration index due to the fact that GE inequality is not bound below 1. This could be remedied, though, by dividing the GE measure by its maximum value (most unequal society). For example, in the general case this maximum value is $^{N^\gamma - 1}/_{\gamma(1+\gamma)} \approx 3.96$ for our sample when $\gamma = -0.5$. Thus the normalized GE inequality value is $^{0.36}/_{3.96} \approx 0.091$, which is lower than the comparable concentration index result of 0.176 for SAH.

Below each value of overall inequality is the percentage of total inequality that is caused by differences in health status between households grouped by wealth quintiles and earnings quintiles. These values, which allow us to examine the relationship between SES and health inequality, are 7.41 percent and 8.91 percent, respectively, when $\gamma = -0.5$. The complement of the percentage of *between* group inequality is the corresponding percentage of total inequality due to differences in overall health *within* groups. While differences in health across SES is, by far, the largest contributor to health inequality for the self-assessed measures, more than 90 percent of health inequity is driven by factors other than SES. This surprising result highlights why focusing on SES-related inequality alone may be short-sighted.

The last two results for each health measure show the percentage of the total inequality that is due to differences in health status between men and women and between black and non-black individuals. Again using the example of $\gamma = -0.5$, we can see that gender and race contribute very little to overall inequalities in SAH with values of 0.8 percent and 1.6

percent, respectively.

The setup is the same for Table 5 except there are five sets of rows that correspond to multivariate health indices constructed with different allowances for complementarity. Comparing Tables 4 and 5 we can see that the self-assessed measures of health suggest a greater degree of inequity than the objective health measures, as was the case with the concentration index. Likewise, the self-assessed measures attribute a much higher percentage of inequality to differences between SES groups than the clinical health measures, which, again, may suggest a degree of SES-related reporting bias in the self-assessed measures—SAH in particular. Unlike self-assessed health, differences in health across gender are the largest contributors to overall inequality—measuring somewhere between 4 to 6 percent of overall health inequality.

As expected, the degree of complementarity built into the aggregation function–expressed by the parameter β—has a significant effect on the measurement of inequality. As β increases, the summary measure treats the different attributes as increasingly complementary, which means that an individuals worst health attribute becomes more influential in the calculation of the summary measure. It follows that an aggregator that treats attributes in a more complementary manner has the potential to be more sensitive to health inequalities in any individual health dimension. This is precisely what we see in our results: as β increases the level of inequality increases significantly. For example, when $\gamma = -0.5$ the overall level of inequality septuples as β increases, going from 0.001 when $\beta = -1$ (health attributes treated as perfectly substitutable) to 0.007 when $\gamma = 2$. This difference becomes more (less) pronounces as the inequality measure becomes more (less) averse to inequality (γ becomes more negative).

Overall, the low level of total inequality and lack of any major demographic or SES-related differences in health suggests that health inequality is not a major problem for the HRS sample.

6 Conclusion

In this study, we propose a MGE measure of inequality developed in Maasoumi (1986) to assess health inequality. The MGE methodology offers both theoretical and practical advantages over the more common concentration index framework originally proposed by Wagstaff et al. (1991). Our framework utilizes a true multidimensional index of objective health measures which are aggregated in a manner that minimizes information loss and makes the implicit assumptions of the summary function transparent, intuitive, and changeable. We then measure health inequality using the GE class of measures, which include many popular inequality indices as special cases.The GE measures can be decomposed into between-group *and* within-group inequality. This also allows us to examine how overall inequality is affected by factors other than SES status. We then empirically demonstrate our proposed methodology using a set of biomarkers for health in the 2006 and 2008 waves of the Health and Retirement Study.

Using our multidimensional measure of health and assessing inequality using GE, we find that health inequality is very small overall. There is almost no change to the inequality measure when the parameter of relative inequality aversion is changed, which suggests that inequality in health does not vary significantly across the distribution of health. Moreover, the inequality that does exist does not appear to be due to differences in gender, race, wealth, or earnings. We also find that the degree of cross-dimensional complementarity allowed by the aggregation function can have a significant effect on the measurement of inequality. This result is particularly important given that implicit assumptions about the degree of complementarity/substitutability are often ignored.

Future research could also concentrate on more sophisticated ways to assess how health inequality is related to income or SES. Our current method to connect health inequality to SES is similar to the concentration index, in that it compares the distribution of health to individuals' income or wealth categories (similar to examining the rank of income). Although the MGE inequality measure we utilize is based on the divergence of the distribution

of health from the uniform allocation of health, the framework theoretically allows for a comparison to any distribution of interest. For example, we could examine the divergence in the distribution of health from the distribution of SES, measured as income, wealth, or an MGE-derived multivariate measure of SES. Large inequality values would indicate a lack of common information (more statistical independence) between health and SES, and the measure could be tailored to focus more heavily on the lower tail of the health or SES distribution by adjusting the value of γ. We could then extend the MGE framework to measure the within versus between contributions of race, gender or other demographic groups, to this measure of inequality. This would be similar to the approaches recently employed by Erreygers (2013) and Makdissi and Yazbeck (2016) using Atkinson inequality measures. However, the information theory approach has the potential to make such comparisons more flexible and intuitive and introduces the possibility of using multidimensional measures of both health and SES.

Additionally, we think it would be helpful to fully characterize the axiomatic properties of our GE measure in light of the most resent research and highlight the differences in the implicit choice of social welfare functions from GE inequality and other popular inequality indices. We would also like to formally characterize a normalized GE measure that can be directly compared to other, bounded inequality indices, like Atkinson or Gini class measures.

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.

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.

Cover, T. M. and J. A. Thomas (2006). *Elements of Information Theory*. Hoboken: Wiley-Interscience.

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.

Decancq, K. and M. A. Lugo (2013, January). Weights in Multidimensional Indices of Wellbeing: An Overview. *Econometric Reviews 32*(1), 7–34.

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. (2013). A dual atkinson measure of socioeconomic inequality of health. *Health Economics 22*(4), 466–479.

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).

Granger, C. W., E. Maasoumi, and J. Racine (2004). A dependence metric for possibly nonlinear processes. *Journal of Time Series Analysis 25*(5), 649–669.

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.

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 (2001). Clusters of attributes and well-being in the usa. *Journal of Applied Econometrics 16*(3), 445–460.

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

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 âĂŞ what does self-rated health measure? *Social Science & Medicine 75*(1), 1 – 9.

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.

Lugo, M. A. (2007). Comparing multidimensional indices of inequality: Methods and application. *Research on Economic Inequality 14*, 213–236.

Maasoumi, E. (1986). The measurement and decomposition of multi-dimensional inequality. *Econometrica 54*(4), 991–997.

Maasoumi, E. and J. H. Jeong (1985). The trend and the measurement of world inequality over extended periods of accounting. *Economics Letters 19*(3), 295–301.

Maasoumi, E. and M. A. Lugo (2008). The information basis of multivariate poverty assessments. *Quantitative Approaches to Multidimensional Poverty Measurement*, 1–29.

Maasoumi, E. and G. Nickelsburg (1988). Multivariate measures of well-being and an analysis of inequality in the michigan data. *Journal of Business & Economic Statistics 6*(3), 327–334.

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 (2016). Avoiding blindness to health status in health achievement and health inequality measurement. *Social Science & Medicine 171*, 39 – 47.

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.

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.

Nesson, E. T. and J. J. Robinson (2016). The measurement of health and the connection between health inequality and income. *Unpublished Working Paper*.

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.

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

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.

Theil, H. (1967a). *Economics and Information Theory*. Chicago: Rand McNally and Company.

Theil, H. (1967b). *Economics and information theory*. North-Holland.

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. (2005). The bounds of the concentration index when the variable of interest is binary, with an application to immunization inequality. *Health Economics 14*(4), 429–432.

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.

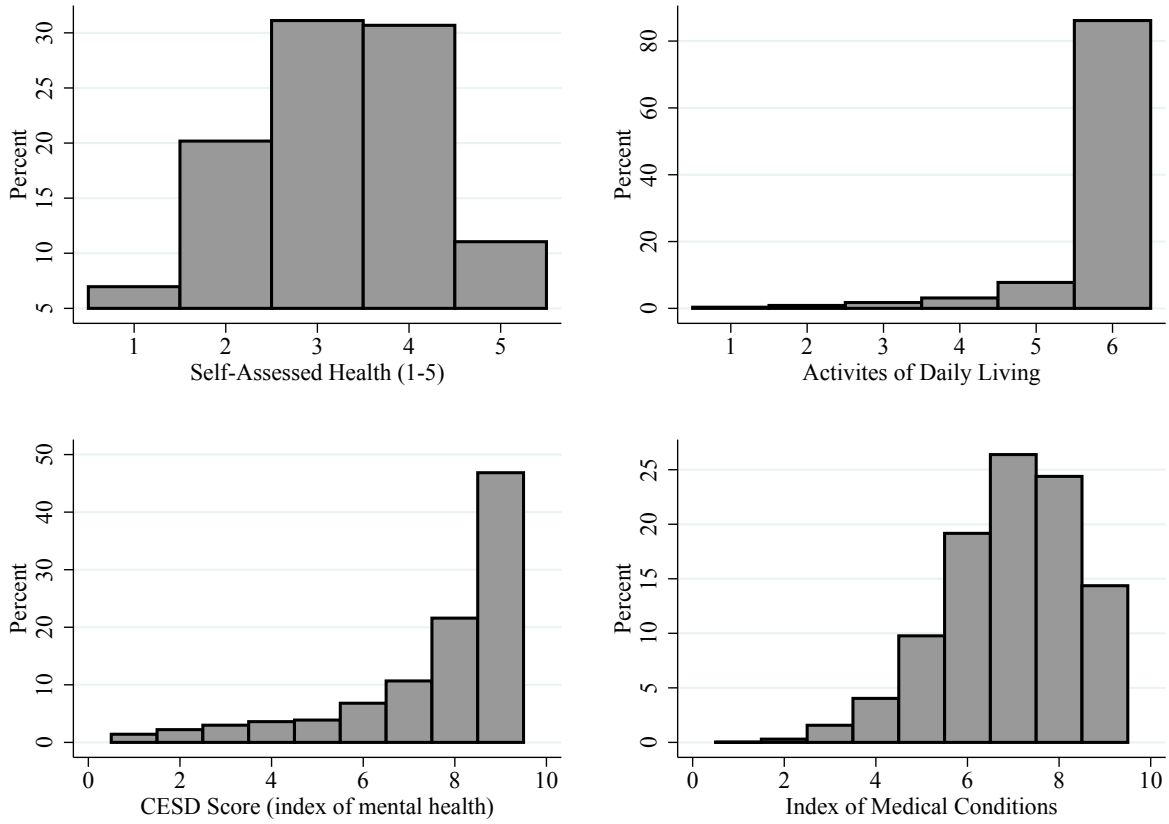
Wagstaff, A., E. Van Doorslaer, and N. Watanabe (2003). On decomposing the causes of health sector inequalities with an application to malnutrition inequalities in vietnam. *Journal of Econometrics 112*(1), 207–223.

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.

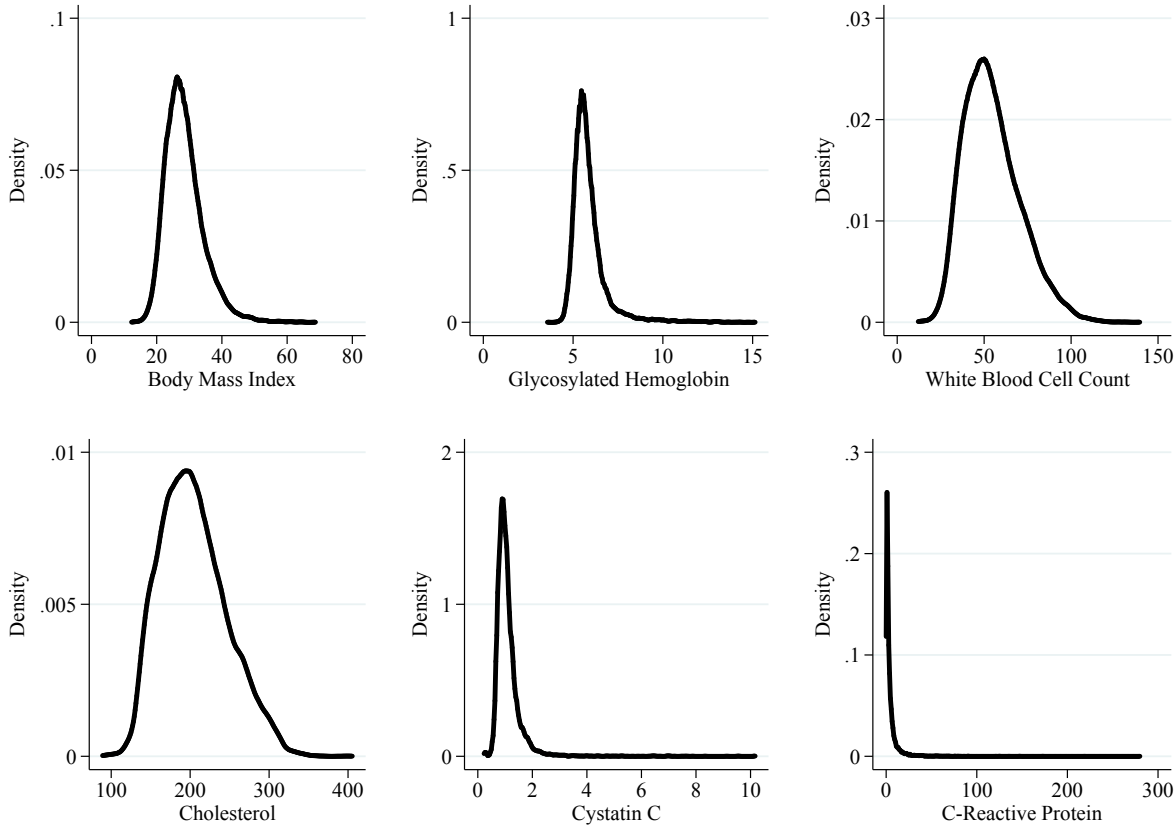
Ziebarth, N. (2010). Measurement of health, health inequality, and reporting heterogeneity. *Social Science & Medicine 71*(1), 116–124.

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



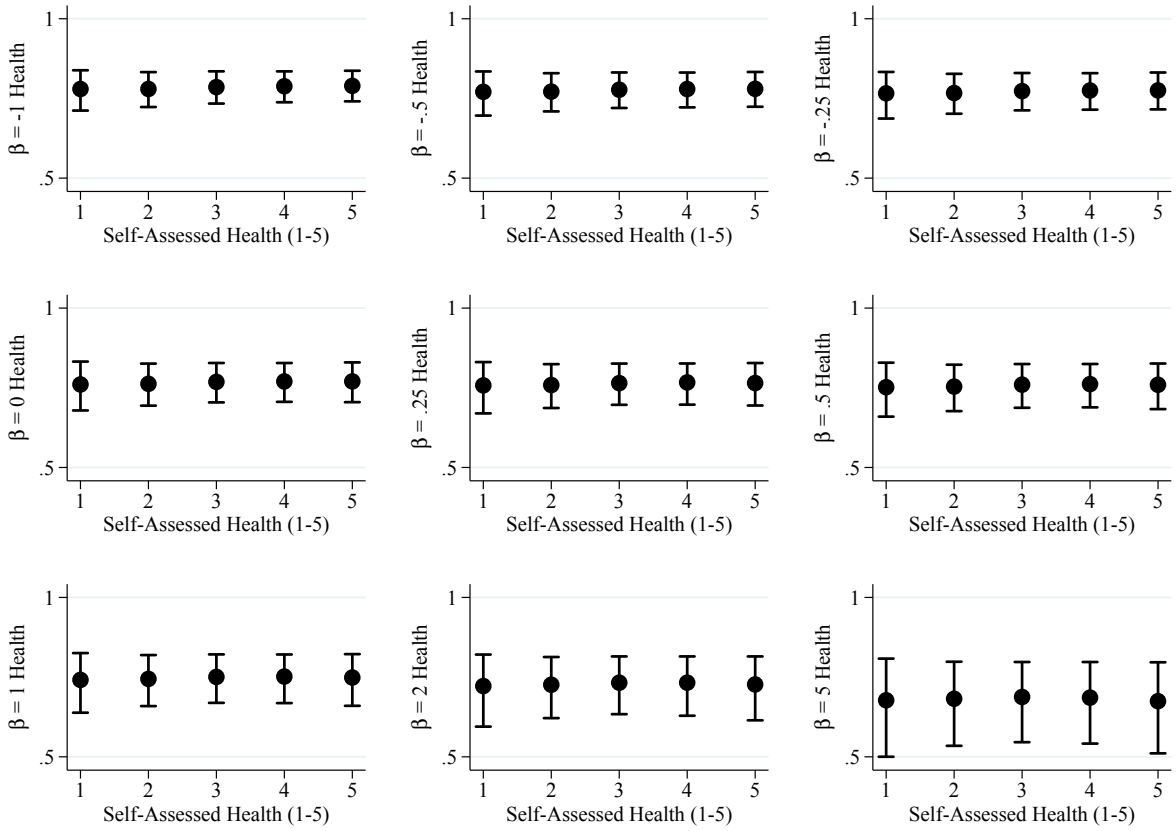
Notes: Data from the 2006 and 2008 waves of the Health and Retirement Survey, including respondents under age 85. Higher values for SAH correspond to better self-assessed health.

Figure 2: Densities of Health Biomarkers



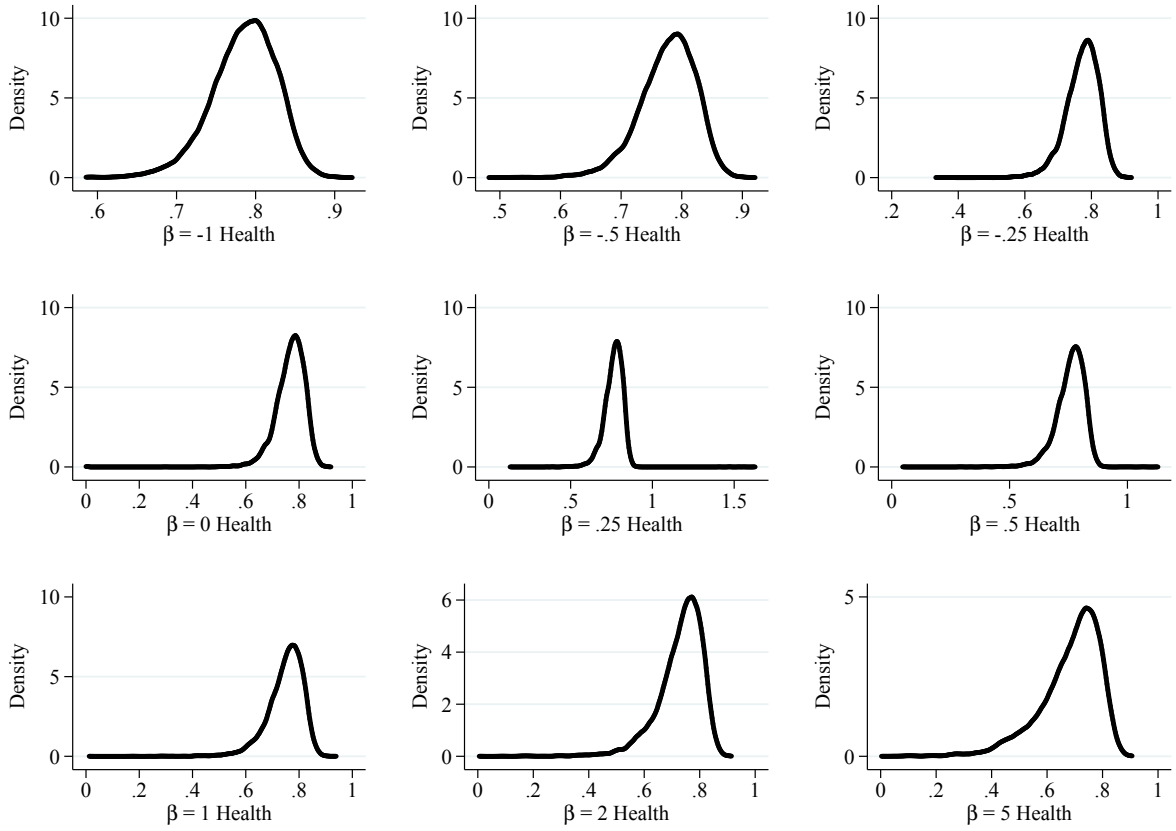
Notes: Data from the 2006 and 2008 waves of the Health and Retirement Survey, including respondents under age 85.

Figure 3: Mean Health Level, 10th Quantile, and 90th Quantile by Self-Assessed Health Status



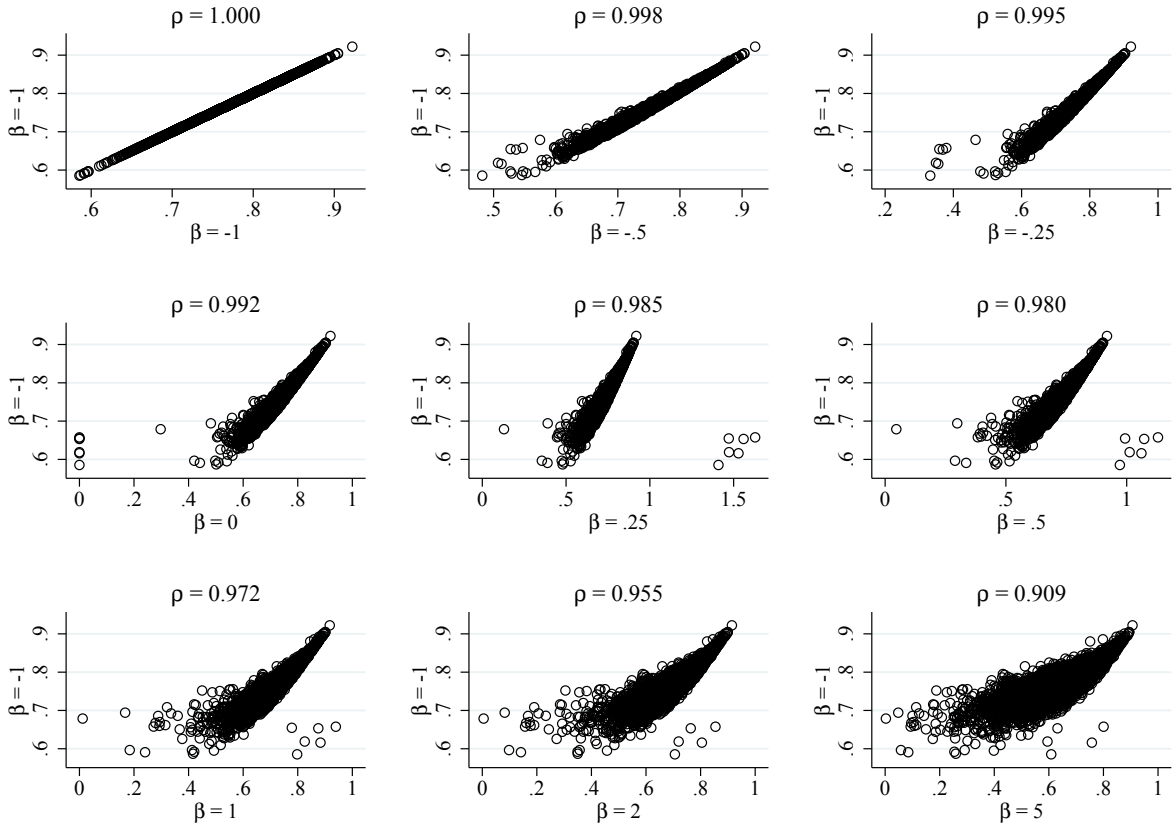
Notes: Data from the 2006 and 2008 waves of the Health and Retirement Survey, including respondents under age 85. Higher values for SAH correspond to better self-assessed health.

Figure 4: Densities of Entropy Health Levels Under Different Values of β



Notes: Data from the 2006 and 2008 waves of the Health and Retirement Survey, including respondents under age 85.

Figure 5: Scatterplots of Health Under Different Values of β



Notes: Data from the 2006 and 2008 waves of the Health and Retirement Survey, including respondents under age 85.

Table 1: Summary Statistics

Variable	Mean	Std.Dev	Min	Max
Body Mass Index	28.558	5.964	12.300	68.700
Glycosylated Hemoglobin	5.828	0.992	3.570	15.140
White Blood Cell Count	54.684	16.319	12.110	139.520
Cholesterol	206.165	42.020	89.040	405.410
Cystatin C	1.069	0.484	0.230	10.170
C-Reactive Protein	4.517	8.172	0.020	280.000
Female	0.553	0.497	0.000	1.000
Black	0.086	0.280	0.000	1.000
Hispanic	0.074	0.262	0.000	1.000
Married	0.661	0.473	0.000	1.000
Widowed	0.136	0.343	0.000	1.000
Divorced/Separated	0.139	0.346	0.000	1.000
Never Married	0.036	0.186	0.000	1.000
Partner	0.027	0.164	0.000	1.000
Age	65.152	8.667	52.000	85.000
Graduate Degree	0.109	0.312	0.000	1.000
College Degree	0.143	0.350	0.000	1.000
Some College	0.057	0.231	0.000	1.000
High School Diploma	0.534	0.499	0.000	1.000
Less than High School	0.157	0.364	0.000	1.000
N	9156			

Notes: Data from the 2006 and 2008 waves of the Health and Retirement Survey, including respondents under age 85.

Table 2: Condition Index Results Using Household Wealth

	Self-Reported Health				MGE Clinical Health				
	Self-Assessed Health	Mental Health	Activities of Daily Living	Medical Conditions	β = -1	β = -0.5	β = 0	β = 1	β = 2
Total Wealth	0.176***	0.193***	0.122***	0.106***	0.017***	0.013***	0.007***	0.007***	0.007**
	(0.010)	(0.012)	(0.010)	(0.005)	(0.004)	(0.003)	(0.002)	(0.002)	(0.003)
Age	-0.009**	0.022***	0.006	-0.020***	-0.002	-0.002	-0.000	-0.000	0.000
	(0.004)	(0.006)	(0.004)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Age/100 Squared	0.005	-0.016***	-0.006*	0.012***	0.003**	0.002*	0.001	0.001	0.000
	(0.003)	(0.004)	(0.003)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Female	0.007	-0.028***	-0.007	-0.005*	-0.038***	-0.033***	-0.018***	-0.022***	-0.027***
	(0.005)	(0.007)	(0.005)	(0.003)	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)
Black	-0.030***	-0.008	-0.014*	-0.003	-0.021***	-0.016***	-0.008***	-0.008***	-0.008***
	(0.007)	(0.010)	(0.009)	(0.005)	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)
Hispanic	-0.036***	-0.054***	-0.029***	0.023***	-0.007*	-0.006*	-0.003	-0.004	-0.004
	(0.009)	(0.012)	(0.010)	(0.005)	(0.004)	(0.003)	(0.002)	(0.003)	(0.003)
Married	-0.007	0.037**	-0.003	-0.016*	-0.003	-0.003	-0.002	-0.001	-0.001
	(0.014)	(0.019)	(0.013)	(0.008)	(0.006)	(0.005)	(0.003)	(0.003)	(0.004)
Widowed	-0.010	-0.026	-0.018	-0.017*	-0.001	-0.000	0.000	0.001	0.002
	(0.015)	(0.020)	(0.015)	(0.009)	(0.006)	(0.005)	(0.003)	(0.004)	(0.005)
Divorced/Separated	-0.006	-0.011	-0.011	-0.015	-0.004	-0.004	-0.003	-0.002	-0.003
	(0.015)	(0.021)	(0.015)	(0.009)	(0.006)	(0.006)	(0.003)	(0.004)	(0.005)
Partner	-0.015	0.013	-0.019	-0.013	-0.003	-0.003	-0.001	-0.002	-0.002
	(0.020)	(0.027)	(0.020)	(0.013)	(0.008)	(0.007)	(0.004)	(0.005)	(0.006)
Adj. R^2	0.092	0.083	0.052	0.135	0.076	0.077	0.071	0.069	0.068
Num Obs	9155	9155	9155	9155	9155	9155	9155	9155	9155

Notes: Data from the 2006 and 2008 waves of the Health and Retirement Survey, including respondents under age 85. Each regression shows . coefficients from a concentration index model, and robust standard errors are shown in parenthesis. The first four columns measure health using self-reported health measures, and the remaining five columns measure MGE Clinical Health. All health measures are transformed in the method suggested by Erreygers (2009a). In addition to the coefficients shown, all models include survey wave fixed effects. All models include sample weights, and stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%

Table 3: Condition Index Results: Robustness Checks

	Self-Reported Health				MGE Clinical Health				
	Self-Assessed Health	Mental Health	Activities of Daily Living	Medical Conditions	β = -1	β = -0.5	β = 0	β = 1	β = 2
Total Wealth (Weighted)	0.176***	0.193***	0.122***	0.106***	0.017***	0.013***	0.007***	0.007***	0.007**
	(0.010)	(0.012)	(0.010)	(0.005)	(0.004)	(0.003)	(0.002)	(0.002)	(0.003)
Total Wealth (Unweighted)	0.157***	0.188***	0.116***	0.099***	0.015***	0.012***	0.006***	0.005***	0.005**
	(0.007)	(0.010)	(0.008)	(0.005)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
Total Wealth Excl. 2nd Homes (Weighted)	0.170***	0.198***	0.123***	0.101***	0.018***	0.014***	0.007***	0.007***	0.008***
	(0.010)	(0.012)	(0.010)	(0.005)	(0.004)	(0.003)	(0.002)	(0.002)	(0.003)
Total Wealth Excl. 2nd Homes (Unweighted)	0.155***	0.197***	0.120***	0.096***	0.015***	0.011***	0.005***	0.005***	0.005**
	(0.007)	(0.010)	(0.008)	(0.005)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
Total Household Earnings (Weighted)	0.194***	0.228***	0.153***	0.103***	0.006	0.004	0.002	0.003	0.003
	(0.010)	(0.014)	(0.011)	(0.006)	(0.004)	(0.004)	(0.002)	(0.003)	(0.003)
Total Household Earnings (Unweighted)	0.169***	0.205***	0.133***	0.091***	0.004	0.003	0.001	0.002	0.002
	(0.008)	(0.011)	(0.008)	(0.005)	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)
Individual Earnings (Weighted)	0.075***	0.114***	0.097***	0.058***	0.003	0.003	0.002	0.003	0.003
	(0.010)	(0.013)	(0.010)	(0.006)	(0.004)	(0.004)	(0.002)	(0.003)	(0.003)
Individual Earnings (Unweighted)	0.073***	0.090***	0.080***	0.056***	0.003	0.003	0.002	0.002	0.002
	(0.007)	(0.010)	(0.007)	(0.004)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)

Notes: Data from the 2006 and 2008 waves of the Health and Retirement Survey, including respondents under age 85. Each regression shows coefficients from a concentration index model, and robust standard errors are shown in parenthesis. Each row shows results using a different measure of income or wealth and either using sample weights or not. The first four columns measure health using self-reported health measures, and the remaining five columns measure MGE Clinical Health. All health measures are transformed in the method suggested by Erreygers (2009a). In addition to the coefficients shown, all models include survey wave fixed effects. Stars denote statistical significance levels: *: 10%, **: 5%, and ***: 1%

Table 4: MGE Health Inequality: Self-Reported Health

		Values for Gamma			
		γ=-2	γ=-1	γ=-0.5	γ=0
Self-Assessed Health (1-5)	All	0.610 (0.010) ***	0.391 (0.004) ***	0.360 (0.004) ***	0.357 (0.003) ***
	Household Wealth Quintiles (%)	4.690 (0.412) ***	6.950 (0.556) ***	7.411 (0.598) ***	7.339 (0.593) ***
	Household Earnings Quintiles (%)	5.626 (0.413) ***	8.336 (0.621) ***	8.911 (0.619) ***	8.864 (0.626) ***
	Female (%)	0.005 (0.012)	0.007 (0.022)	0.008 (0.024)	0.008 (0.030)
	Black (%)	1.097 (0.197) ***	1.556 (0.255) ***	1.616 (0.264) ***	1.557 (0.256) ***
CESD Score (index of mental health)	All	0.511 (0.009) ***	0.286 (0.003) ***	0.241 (0.003) ***	0.214 (0.003) ***
	Household Wealth Quintiles (%)	2.859 (0.258) ***	4.831 (0.399) ***	5.600 (0.455) ***	6.159 (0.477) ***
	Household Earnings Quintiles (%)	3.019 (0.252) ***	5.072 (0.435) ***	5.865 (0.452) ***	6.436 (0.510) ***
	Female (%)	0.301 (0.065) ***	0.540 (0.116) ***	0.644 (0.136) ***	0.726 (0.160) ***
	Black (%)	0.282 (0.072) ***	0.484 (0.124) ***	0.563 (0.139) ***	0.620 (0.162) ***
Activites of Daily Living	All	0.276 (0.008) ***	0.124 (0.003) ***	0.092 (0.002) ***	0.073 (0.002) ***
	Household Wealth Quintiles (%)	0.722 (0.098) ***	1.578 (0.217) ***	2.092 (0.254) ***	2.630 (0.345) ***
	Household Earnings Quintiles (%)	1.041 (0.117) ***	2.266 (0.247) ***	2.997 (0.314) ***	3.762 (0.397) ***
	Female (%)	0.020 (0.012)	0.045 (0.028)	0.060 (0.036)	0.077 (0.048)
	Black (%)	0.085 (0.031) **	0.186 (0.068) **	0.247 (0.088) **	0.311 (0.108) **
Index of Medical Conditions	All	0.030 (0.001) ***	0.026 (0.001) ***	0.024 (0.000) ***	0.023 (0.000) ***
	Household Wealth Quintiles (%)	2.954 (0.300) ***	3.441 (0.354) ***	3.639 (0.364) ***	3.808 (0.396) ***
	Household Earnings Quintiles (%)	4.686 (0.373) ***	5.484 (0.435) ***	5.813 (0.447) ***	6.099 (0.452) ***
	Female (%)	0.012 (0.022)	0.015 (0.026)	0.015 (0.025)	0.016 (0.030)
	Black (%)	0.380 (0.114) ***	0.439 (0.120) ***	0.463 (0.140) ***	0.482 (0.148) **

Notes: Data from the 2006 and 2008 waves of the Health and Retirement Survey, including respondents under age 85. Each set of results show MGE inequality metrics for a particular self-reported health measure and value for γ. Standard errors in parentheses are calculated using 500 bootstrap iterations. Below the overall inequality measure are decompositions of the percent of the inequality that is due to differences health attributes between groups, as defined by that particular variable (the complement of this percentage is the portion of total inequality that is due to within-group variation in health).

Table 5: MGE Health Inequality: MGE Clinical Health Levels

		Values for Gamma			
		γ=-2	γ=-1	γ=-0.5	γ=0
Beta = -1	All	0.001 (0.000) ***	0.001 (0.000) ***	0.001 (0.000) ***	0.001 (0.000) ***
	Household Wealth Quintiles (%)	1.159 (0.219) ***	1.175 (0.227) ***	1.183 (0.231) ***	1.190 (0.220) ***
	Household Earnings Quintiles (%)	0.340 (0.124) **	0.345 (0.121) **	0.347 (0.123) **	0.349 (0.120) **
	Female (%)	5.794 (0.454) ***	5.890 (0.490) ***	5.934 (0.467) ***	5.976 (0.471) ***
	Black (%)	1.486 (0.279) ***	1.501 (0.255) ***	1.507 (0.274) ***	1.513 (0.275) ***
Beta = -.5	All	0.002 (0.000) ***	0.002 (0.000) ***	0.002 (0.000) ***	0.002 (0.000) ***
	Household Wealth Quintiles (%)	0.981 (0.197) ***	1.004 (0.205) ***	1.014 (0.222) ***	1.024 (0.217) ***
	Household Earnings Quintiles (%)	0.255 (0.106) *	0.260 (0.108) *	0.263 (0.114) *	0.266 (0.115) *
	Female (%)	5.713 (0.433) ***	5.859 (0.450) ***	5.927 (0.481) ***	5.991 (0.456) ***
	Black (%)	1.279 (0.246) ***	1.303 (0.245) ***	1.313 (0.257) ***	1.322 (0.270) ***
Beta = 0	All	0.002 (0.000) ***	0.002 (0.000) ***	0.004 (0.000) ***	0.003 (0.000) ***
	Household Wealth Quintiles (%)	1.009 (0.217) ***	1.226 (0.274) ***	0.620 (0.143) ***	0.754 (0.171) ***
	Household Earnings Quintiles (%)	0.237 (0.117) *	0.288 (0.146) *	0.145 (0.073) *	0.177 (0.092)
	Female (%)	6.524 (0.520) ***	7.951 (0.647) ***	4.026 (0.317) ***	4.902 (0.377) ***
	Black (%)	1.375 (0.293) ***	1.662 (0.343) ***	0.838 (0.191) ***	1.017 (0.229) ***
Beta = 1	All	0.008 (0.003) **	0.005 (0.000) ***	0.005 (0.000) ***	0.004 (0.000) ***
	Household Wealth Quintiles (%)	0.348 (0.089) ***	0.579 (0.149) ***	0.623 (0.160) ***	0.652 (0.165) ***
	Household Earnings Quintiles (%)	0.071 (0.041)	0.117 (0.072)	0.126 (0.077)	0.132 (0.075)
	Female (%)	2.730 (0.222) ***	4.564 (0.392) ***	4.914 (0.410) ***	5.158 (0.424) ***
	Black (%)	0.384 (0.106) ***	0.637 (0.161) ***	0.683 (0.185) ***	0.714 (0.184) ***
Beta = 2	All	0.017 (0.008) *	0.008 (0.001) ***	0.007 (0.000) ***	0.007 (0.000) ***
	Household Wealth Quintiles (%)	0.216 (0.058) ***	0.447 (0.125) ***	0.491 (0.146) ***	0.522 (0.140) ***
	Household Earnings Quintiles (%)	0.039 (0.029)	0.081 (0.061)	0.089 (0.057)	0.095 (0.063)
	Female (%)	2.008 (0.172) ***	4.173 (0.353) ***	4.590 (0.403) ***	4.894 (0.425) ***
	Black (%)	0.225 (0.064) ***	0.462 (0.141) **	0.506 (0.131) ***	0.537 (0.156) ***

Notes: Data from the 2006 and 2008 waves of the Health and Retirement Survey, including respondents under age 85. Each set of results show MGE inequality metrics for a particular value of β and γ. Standard errors in parentheses are calculated using 500 bootstrap iterations. Below the overall inequality measure are decompositions of the percent of the inequality that lies between a particular variable (as compared to the percent that lies within the particular variables).