

Costa-i-Font, Joan; Cowell, Frank

**Working Paper**

## The Measurement of Health Inequalities: Does Status Matter?

CESifo Working Paper, No. 6117

**Provided in Cooperation with:**

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

*Suggested Citation:* Costa-i-Font, Joan; Cowell, Frank (2016) : The Measurement of Health Inequalities: Does Status Matter?, CESifo Working Paper, No. 6117, Center for Economic Studies and ifo Institute (CESifo), Munich

This Version is available at:

<https://hdl.handle.net/10419/147371>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



## The Measurement of Health Inequalities: Does Status Matter?

Joan Costa-Font  
Frank A. Cowell

CESIFO WORKING PAPER NO. 6117

CATEGORY 3: SOCIAL PROTECTION

OCTOBER 2016

*An electronic version of the paper may be downloaded*

- from the SSRN website: [www.SSRN.com](http://www.SSRN.com)
- from the RePEc website: [www.RePEc.org](http://www.RePEc.org)
- from the CESifo website: [www.CESifo-group.org/wp](http://www.CESifo-group.org/wp)

ISSN 2364-1428

**CESifo**

Center for Economic Studies & Ifo Institute

# The Measurement of Health Inequalities: Does Status Matter?

## Abstract

The measurement of health inequalities usually involves either estimating the concentration of health outcomes using an income-based measure of status or applying conventional inequality-measurement tools to a health variable that is non-continuous or, in many cases, categorical. However, these approaches are problematic as they ignore less restrictive approaches to status. The approach in this paper is based on measuring inequality conditional on an individual's position in the distribution of health outcomes: this enables us to deal consistently with categorical data. We examine several status concepts to examine self-assessed health inequality using the sample of world countries contained in the World Health Survey. We also perform correlation and regression analysis on the determinants of inequality estimates assuming an arbitrary cardinalisation. Our findings indicate major heterogeneity in health inequality estimates depending on the status approach, distributional-sensitivity parameter and measure adopted. We find evidence that pure health inequalities vary with median health status alongside measures of government quality.

JEL-Codes: D630, H230, I180.

Keywords: health inequality, categorical data, entropy measures, health surveys, upward status, downward status.

*Joan Costa-Font*  
*London School of Economics*  
*Houghton Street*  
*United Kingdom - WC2A 2AE London*  
*j.costa-font@lse.ac.uk*

*Frank A. Cowell*  
*London School of Economics*  
*Houghton Street*  
*United Kingdom - WC2A 2AE London*  
*F.Cowell@lse.ac.uk*

October 1, 2016

We are grateful to Xuezhong Shi and to Alejandra Hernandez for research assistance and to STICERD for financial assistance.

# 1 Introduction

Measuring health inequality presents a challenge quite different from the standard problem of measuring income or wealth inequality. The challenge principally lies in the measurement of health itself: health cannot be assumed to be directly and unambiguously observable and it may not make sense to treat it as a continuous variable. As a consequence one has to use indirect methods that may involve elicitation of a person's self-assessed health status or explicit modelling using observables that are thought to be related to health. Such indirect methods can be problematic. So the purpose of this paper is to examine the main practical approaches to inequality measurement in the health context and the extent to which different assumptions about health status affect inequality comparisons.

Why are indirect approaches to health measurement typically problematic? The first reason is because of the assumptions that have to be adopted in modelling health: if health status is taken as a latent variable, with what observables is it correlated?<sup>1</sup> There is evidently room for several alternative answers: some research suggests that SAH correlates with mortality, some with hospital records (Heien 2015, Idler and Benyamini 1997). The second – and perhaps more fundamental – reason that such approaches are problematic is that health cannot be taken as a monetary-equivalent measure and that, in many health models, it should be treated as an ordinal or categorical variable rather than a continuous variable. That being so, standard methods of inequality analysis and standard properties of inequality indexes do not apply (Van Doorslaer and Jones 2003). So, how is one to measure inequality?

This paper addresses the main theoretical and practical difficulties presented by the measurability problem of health-status inequality and, in doing so, examines the problems of working with self-assessed health (SAH) indicators, the use of alternative approaches to the measurement health inequality and the information content of different concepts of status. We compare our approach to the case of inequality analysis based on a standard but arbitrary cardinalisation of health using standard inequality indices.

The results from this paper go towards the identification of a more appropriately based definition of health status and of health-inequality measures. We provide researchers with a simple means of testing alternative ways of

---

<sup>1</sup>On the correlation between SAH and objective measures of health status see Bound (1991).

measuring inequalities of non-cardinal outcomes that may have significant policy implications. This is particularly important when one takes account of the fact that measures of health inequality are used to rank health systems, and increasingly measures of well-being are used by the World Health Organisation and other government bodies to evaluate institutions and public policies. Specifically, we undertake the following:

1. We provide an alternative estimation of (pure) health inequalities using a measure of status that is not imposed through an arbitrary cardinalisation strategy, but based on two perspectives of the distribution of health status. We focus on concepts of status derived from the distribution of health outcomes to compute inequality measurements. This is an application of the cardinalisation method employed in the Cowell-Flachaire (2014) status-inequality approach. We compare these estimates to those that would emerge from simple direct cardinalisations of health status.
2. We provide the first multi-country estimate of health inequalities using the status-inequality approach.
3. We then examine the determinants of the two different perspectives of status-inequality (mentioned in 1 above) across countries taking into account determinants such as per capita income, average education demographics as well as a collection of indicators for political development. This allows one to examine a number of important questions: do poor countries exhibit higher health-inequality irrespective of status? Are specific demographics more likely to engender health inequalities? What is the role of institutional variables, and specifically measures of democratic development (which can explain policy sensitivity to health needs)? We control for cross-country fixed effects and regional consistency in inequality patterns,

The paper is organised as follows. Section 2 contains some necessary theoretical background, section 3 introduces the data set and explains our empirical strategy, section 4 contains our results and section 5 concludes.

## 2 Health-inequality measurement – principles and practice

For a coherent approach to the measurement of health inequality we need two basic concepts and a methodology for measuring or estimating values of these concepts and then aggregating the values..

### 2.1 Basic concepts: health and status

If we were able to treat “health” like “wealth” then a person’s health could be taken as a continuous variable that is, in itself, an objectively measurable and observable measure of a person’s status. For a broadly-defined interpretation of health this is unrealistic. One could try to use proxies – since the contribution of Idler and Benyamini (1997), categorical self-assessed measures of health are taken as acceptable proxies for individuals – but perception and observation might not necessarily match Sen (2002). One could also focus on the inequality of individual components or aspects of health that can be objectively measured (just as particular components of wealth or income are interesting subjects of inequality analysis), but this is necessarily of limited interest and applicability.

Since there is no standard off-the-shelf measure of health status that is going to be generally suitable for inequality analysis, we need to be clear about two steps: (1) how to model health  $h_i$  for each individual  $i = 1, \dots, n$  and, (2) given  $\{h_1, \dots, h_n\}$ , how to model the status variable  $s_i$  that is to be used in inequality computation.

#### 2.1.1 Individual health, $h_i$

Given the difficulties in observing a broadly defined indicator of individual  $h_i$  there are two main ways forward. First one might try estimate a health production function assuming the following kind of relationship:

$$h_i = \Phi(\mathbf{X}_i) + \varepsilon_i, \tag{1}$$

where  $\Phi$  is the production function,  $\mathbf{X}_i$  represents a vector of determinants of health (such as income, demographics, institutions) and  $\varepsilon_i$  a random component. Depending on the nature of the assumed specification of  $\Phi$ ,  $h_i$  could be continuous or discrete variable. Specifying and estimating such a

function is challenging because health status is a latent variable that cannot be observed.

The second approach is to model  $h_i$  thus:

$$h_i \in \{c', c'', c''', \dots\}, \quad (2)$$

where  $c', c'', c'''$  represent different health categories. Many national and international surveys contain information on measures of SAH in categorical form; the categories may or may not have a natural ordering.

So, in principle, an individual's health  $h_i$  can take the form of a censored variable, an interval variable, or an ordered categorical variable, depending on the underlying assumptions about how to conceptualise it and model it.

### 2.1.2 Individual status, $s_i$

How one models an individual's status depends in part on the way  $h_i$  is modelled. If one follows the production-function approach, it might be possible to use the  $h_i$  value as an indicator of health status, just as it pops out of equation (1), or a transformation of it. If one is using a specification such as that of equation (2) then a number of problems immediately present themselves: how to order the members of the set  $\{c', c'', c''', \dots\}$ , how to calibrate the "distance" between members of the set of categories and so on. One might, alternatively, incorporate in the concept of status information about health and some other personal characteristic, such as income.

Whether one starts with equation (2) or (1), or perhaps something else, the analysis involves three main components: (1) extraction of suitable categorical variables on which to construct health and status indices; (2) computation of cardinal imputations, status measures and associated inequality indices and rankings; (3) an analysis of cross-country inequality comparisons. These three things are addressed in the following subsections.

## 2.2 Concentration curves and indices

The concentration curve and concentration indices (Costa-Font and Hernández-Quevedo 2012, Koolman and van Doorslaer 2004) form arguably the most popular approach in the health-inequality literature. At the core of this approach is the idea that a person's status should be based not only health but also on other personal or social information, as discussed in section 2.1:

this supplementary information is personal income or other similar indicators that are believed to co-vary with health. Individual status is usually taken to mean a person’s position in the income hierarchy,<sup>2</sup> but the use of income-rank as a status variable is a matter of choice (why not consumption or wealth?) and its use can be problematic insofar as income is measured with limited precision. However, pursuing this approach make it difficult to identify whether policies really affect *health* inequality, or simply affect determinants of health such as the distribution of material conditions (which in turn give raise to a fairer distribution of health status). Furthermore the approach can appear ad hoc rather than rigorously founded on economic theory.<sup>3</sup>

Studies that use income as a measure of social hierarchy on self-assessed measures of health are problematic, in several ways. First, they tend to underestimate income-related inequalities in health (Dowd and Todd 2011). Second, the distribution should perhaps be adjusted to eliminate factors such as age or gender which could be considered to distort the picture of inequality. Third, even if such adjustments are successfully made, the approach typically ignores the contribution of essentially “avoidable” determinants of health, or even potentially “ethically legitimate” differences in health resulting from preventive effort and choice (Le Grand 1987).

## 2.3 Inequality of cardinal indicators

An alternative approach is to apply conventional inequality tools to specific cardinal indicators of health such as life expectancy, or measures of hypertension. Some of the socio-economic difference in health might not be pure socio-economic inequalities, but may be determined by lifestyles (Fleurbaey and Schokkaert 2009, 2012; other inequalities – such as those resulting from poor health production in younger ages, or biologically driven gender differences in health – are, arguably, not avoidable (Wagstaff et al. 1991). Should one attempt to remove all avoidable components from the analysis and fo-

---

<sup>2</sup>See, for example, Marmot (2005), Wagstaff and van Doorslaer (2000). This was the position adopted in the early literature, with few exceptions (such as Le Grand 1987); recent research has argued that socioeconomic background is only one dimension of health inequality.

<sup>3</sup>An attempt has been made to justify the approach on the basis of the principle of income-related health transfers (Bleichrodt and van Doorslaer 2006, Fleurbaey 2006, Fleurbaey and Schokkaert 2012, page 1012), the plausibility of which is questionable.



cus solely on the remaining health inequalities? To do so makes an implicit claim of what is an illegitimate inequality, an approach that might seem to be tendentious. Instead one might adopt a value-free choice.

The further difficulty with the cardinal-indicators approach is that some of the most important indicators of health status are categorical variables that do not have a natural cardinalisation. Clearly one could try to circumvent this by imputing some artificial index of individual health status as a function of the categories (Fonseca and Jones 2003). For example, in some cases the imputation is achieved through subjective evaluation by individuals (for example on a Likert scale) and in some cases by making use of quality of life indices (for example, EuroQuol-EQ5).<sup>4</sup> There are various types of cardinalisation methods that have been proposed in the literature – such as imputation of quality of life scales (for example, values from Visual Analogue Scales), interval regression and so on – but there is insufficient discussion of the economic rationale for these methods or the practical implications of using one method rather than another and it has been shown that cardinalisation can be an important source of bias.<sup>5</sup> As a consequence, one should treat some of the existing literature with some caution.<sup>6</sup> A more promising approach involves a proposed cardinalisation of SAH status using an imputation of the values of Health-Related Quality of Life (HRQoL): one obtains the value of the cut-off point of each response to the ordinal question in order to estimate the determinants of SAH using an interval-regression approach (assuming income-related health inequality – Fonseca and Jones

---

<sup>4</sup> The same procedure can be applied to entities that do not have a natural ordering, such as vectors of attributes or endowments; one uses a utility function to force an ordering of the data. This is similar to one of the standard theoretical approaches to the measurement of multidimensional inequality – one computes the “utility” of factors and then computes inequality of utility, where the utility function is an appropriate aggregator (Maasoumi 1986, Tsui 1995). However the approach faces serious objections such as the arbitrariness of the cardinalisation and of aggregation. Even if the resulting well-being index appears reasonable over a wide subset of categories one might still be concerned about the way extreme values are represented in the index and their consequences for inequality comparisons.

<sup>5</sup> In a paper running a meta-regression of health inequality studies in the economics literature Costa-Font and Hernández-Quevedo (2013) find that the main reason for estimate heterogeneity is the way studies cardinalise health status.

<sup>6</sup>Some of the empirical attempts to examine such problems are in Costa-Font and Cowell (2013) where they use World Health Survey data to examine alternative pragmatic methods for measuring health-inequality and examining regional and country patterns of inequality orderings.

2003, Van Doorslaer and Jones 2003). Let us consider further the underlying distribution-based concept as a basis for the evaluation of status.

## 2.4 Alternative approaches using categorical data

Here we consider approaches that use categorical data directly, without trying to impose *a priori* a particular cardinalisation. The theoretical literature on the problem of making inequality comparisons when the underlying equalisation is ordinal has mainly resulted in a number of rather limited propositions that are difficult to interpret or apply.<sup>7</sup> However, recent work on the analysis of distributions of categorical variables has shown how natural interpretations of individual *status* can be used to provide a robust approach to the inequality-measurement problem in this context without resort to arbitrary cardinalisation of ordinal concepts (Cowell and Flachaire 2014). The status concept is similar to concepts used in poverty and relative deprivation and in recent approaches to the inequality of opportunity (de Barros et al. 2008).

Status interpreted as an individual's position in the health distribution is important in understanding several relationships in the economics of health. For example Costa-Font and Costa-Font (2009) and Hausman et al. (2002) show that effect of income on SAH depends in part on the individual's position in the health distribution: this finding is potentially important in understanding the persistence of health inequalities over time, and more specifically suggest that the their effect depends on individual position within a given health distribution. In the present context the status approach gives rise to an alternative way of making inequality comparisons; it also gives rise to a set of inequality indices that incorporate conventional distributional views such as degree of inequality aversion and that can be applied to commonly-used measures of individual well-being.

The Cowell and Flachaire (2014) approach tackles the problem by separating out carefully the two tricky components of inequality measurement

---

<sup>7</sup>It involves a reworking of traditional inequality-ranking approaches focusing on first-order dominance criteria (Abul Naga and Yalcin 2008, Allison and Foster 2004, Zheng 2011). It is commonly suggested that the median could be used as an equality concept corresponding to the use of the mean in conventional inequality analysis, although it has been noted that comparing distributions with different medians raises special issues (Abul Naga and Yalcin 2010). But the approach runs into difficulty if quantiles are not well-defined, as may happen in the case of categorical variables – see Cowell and Flachaire 2014).

mentioned in the introduction, the equalisand and the aggregation method. Each of these is underpinned by an axiomatic argument that goes based on first principles. The resulting Cowell and Flachaire method amounts to an aggregation of the discrepancies between each person's actual status and some status reference point. In such an approach clearly a lot rests on the precise definition of status. In the case of applications where the equalisand has a natural cardinalisation (income or wealth for example) then it makes sense to define status as just income or wealth. However, where only ordinal information is available – as with categorical data on health status – then we have to do more. Suppose that information is purely categorical, in that we only know how many people are in each category  $k = 1, 2, \dots, K$ , but that the categories can be arranged in increasing order of their desirability. Then a simple argument shows that, if there are  $n_k$  persons in category  $k = 1, 2, 3, \dots, K$ , the status of person  $i$  who is currently in category  $k(i)$  must be a function of either  $\sum_{\ell=1}^{k(i)} n_\ell$  or  $\sum_{\ell=k(i)}^K n_\ell$ . The first of these is a “downward looking” concept and the second is its “upward looking” counterpart. It may be appropriate to normalise by the size of the total population  $n := \sum_1^K n_k$  so that person  $i$ 's status is given by either the downward-looking version

$$s_i = \frac{1}{n} \sum_{\ell=1}^{k(i)} n_\ell, \quad (3)$$

or by the “upward-looking” counterpart of (3):

$$s'_i = \frac{1}{n} \sum_{\ell=k(i)}^K n_\ell; \quad (4)$$

On either definition status must lie between zero and one. If there were perfect equality (everyone in the same category) then it is clear that both (3) and (4) take the value 1. It turns out that this, the maximum-status value, is the only thing that makes sense as the reference point.

The inequality-measurement problem then amounts to aggregating the information in the vector  $\mathbf{s} := (s_1, s_2, \dots, s_n)$  in relation to the equality vector  $(1, 1, \dots, 1)$ . On the basis of a small number of elementary axioms Cowell and Flachaire show that inequality must take the form of an index in the

following family, indexed by  $\alpha$ :

$$I_\alpha(\mathbf{s}) = \begin{cases} \frac{1}{\alpha(\alpha-1)} \left[ \frac{1}{n} \sum_{i=1}^n s_i^\alpha - 1 \right], & \text{if } \alpha \neq 0, \\ -\frac{1}{n} \sum_{i=1}^n \log s_i, & \text{if } \alpha = 0. \end{cases} \quad (5)$$

where  $\alpha < 1$  is a parameter indicating the desired sensitivity of the index to a particular part of the income distribution: for low values of  $\alpha$  the index  $I_\alpha(\mathbf{s})$  is particularly sensitive to values of  $s_i$  close to zero. If  $\mathbf{s}$  is given by (3) ( $i = 1, \dots, n$ ) then we have an index of ordinal inequality based on downward-looking status; if we replace  $\mathbf{s}$  with  $\mathbf{s}'$  given by (4) then we have ordinal inequality defined on the corresponding upward-looking status concept.

So we have a *family* of indices that is suitable for making comparisons of inequality in terms of health status. In addition, members of the family can be adjusted by different health-inequality aversion parameters in a flexible way as other inequality indices. In what follows we shall suggest a way of using this to make health-inequality comparisons internationally.

Clearly equation (5) has a form similar to the well-known Generalised Entropy class of inequality indices (Cowell 1980, Shorrocks 1980)

$$G_\alpha(\mathbf{s}) = \begin{cases} \frac{1}{\alpha(\alpha-1)} \left[ \frac{1}{n} \sum_{i=1}^n \left[ \frac{s_i}{\mu(\mathbf{s})} \right]^\alpha - 1 \right], & \text{if } \alpha \neq 0, 1 \\ -\frac{1}{n} \sum_{i=1}^n \log \frac{s_i}{\mu(\mathbf{s})}, & \text{if } \alpha = 0, \\ \frac{1}{n} \sum_{i=1}^n \frac{s_i}{\mu(\mathbf{s})} \log \frac{s_i}{\mu(\mathbf{s})}, & 1, \end{cases} \quad (6)$$

where  $\mu(\mathbf{s})$  is the mean of the vector  $\mathbf{s}$ . Whereas  $I_\alpha(\mathbf{s})$  has the reference point 1 the GE index  $G_\alpha(\mathbf{s})$  has the reference point  $\mu(\mathbf{s})$  and, obviously, this only makes sense where status is cardinal, in other words, if status defined in such a way that it is meaningful to add the status values together. With ordinal data one could impose an arbitrary cardinalisation and, in section 4 we will try out the performance of  $G_\alpha(\mathbf{s})$  for two such arbitrary cardinalisations and compare them with the theoretically appropriate  $I_\alpha(\mathbf{s})$ .

## 3 Data and Methods

### 3.1 Data

In view of these points it is clear that the underlying data and the health-status indicator derived from it could be of the following forms:

**Continuous, censored.** In some circumstances, health status can be measured using a censored continuous variable (for example when visual analogue scales are employed). However, there are still problems related to focal responses so that certain points in a scale are more common than others (De Boer et al. 2004).

**Ordinal.** Given the categorical nature of SAH, it may be reasonable to take the ordering of question responses as naturally given and to employ techniques designed for ordered variables. For example, an ordered probit model, could be used capture the degree of intensity of health or ill health, without explicitly cardinalising the health concept.

Our approach requires quantitative analysis of internationally comparable data that contain measures of health status. Accordingly the main data source to be used is the World Health Survey which contains data from seventy countries; it collects comparable multidimensional micro-data on income, employment education and health. There are two reasons for the choice of this data base: first, its great advantage for comparative work; second, its standardised world-wide structure can assist in examining cross country patterns across heterogeneous world regions that exhibits different levels of economic and social development.

The World Health Survey (WHS) is a general population survey, developed by WHO to address the need for reliable information and to cater to the increased attention to the role of health in economic and human development. Other alternative options (ISSP, Gallup etc) were not as rich in terms of controls and measurement. Indeed, the survey contains data from randomly selected adults (i.e. older than 18 years of age) who reside in seventy-one countries who implemented household face-to-face surveys, computer assisted telephone interview, or computer-assisted personal interview in 2002. Sample sizes range from 1,000 to 10,000.

Our measure of health status is the standard measure of SAH status

widely used in the literature; this is a categorical measure of health is based on the responses to the question “how would you rate your health today?” and yields a personal evaluation of overall health with potential responses in five categories ranging from “very good” to “very bad”.<sup>8</sup> As a measure, it suffers from cultural adaptation problems that make cross-country comparison challenging, but it appears to be an adequate measure for computing within-country inequalities.

## 3.2 Cardinalisation

For categorical data a simple way to process the data is to rank the values underlying the latent variable health. But the “distance” between categories is unknown and an arbitrary scale may not be informative; there is no theoretically sound consensus strategy to measure such a latent variable from categorical responses

### 3.2.1 “Natural” cardinalisation?

In some cases it is possible to employ existing quality-of-life measures of health status that are available in health surveys to impute a cardinal value to the categorical responses to the SAH questions, for example the imputation of values from the Health-Related Quality of Life scales as in Van Doorslaer and Jones (2003), Fonseca and Jones (2003).

Another way forward is to obtain a linear index based on scaling the ordered variable to obtain a normalised health index (Cutler and Richardson 1997). However, this still requires arbitrary assumptions on the value and distribution of a person’s health status. Makdissi and Yazbeck (2014) address the question of the categorical measurement of health variables by using a ratio-scale transformation that modifies the information provided and focuses on the ‘breadth’ rather than the ‘depth’ of the health-indicator information. However, they lose some important information on the distribution of the health variable and they focus on income-related health inequalities which involves important and questionable assumptions

So, instead, some papers interpret SAH status as a individual’s categorisation into an interval, which can be ascertained by finding a link between self-assessed measures of health and some health utility indices. For example, Van Doorslaer and Jones (2003) use the equivalent cardinal value of the

---

<sup>8</sup>The detailed values are given in Table 5 in the Appendix

cut-off point of each response to the ordinal question was obtained so as to estimate the cardinal value of SAH using an interval-regression approach.

### 3.2.2 Regression approach

Both ordered and interval regressions models can be used to transform a categorical outcome into a continuous variable based on the parameters of the regression. So, if the health variable allows an unambiguous ordering, then a logit or probit regression model will take into account the structure of the data.<sup>9</sup> By assuming an order the probability of respondents' classifying themselves on a specific scale can modelled in the standard fashion. However, even where this is an improvement with respect to binary measures of health for the purposes of measuring health inequality, it is still difficult to interpret the meaning of a change in the order between scales of SAH status.

However, the transformation is dependent in the covariates of the regression and on the arbitrary nature of different variable categories. The strategy we pursue here addresses this latter point and provides an alternative cardinalisation method, that we argue is more suitable to measure inequalities in health.

### 3.2.3 Pure health inequality?

Instead of trying to use the structure of the data to produce a cardinalisation of health status, Allison and Foster (2004) develop a stochastic dominance approach to 'pure health inequalities', which is not limited by the extent of assumptions implicit in partial inequality measurements of income-related health inequality. However, the range of results that are available from this approach is rather narrow and so it is likely to be limited.

In this paper we use the Cowell and Flachaire (2014) methodology discussed in section 2.4 to undertake international comparisons of inequality of SAH status.

---

<sup>9</sup>See for example the logistic regression technique used by Kunst and Mackenbach (1994).

### 3.3 Inequality comparisons

Our approach in this paper is to measure the inequality of SAH status using the WHS international data set and the robust Cowell and Flachaire (2014) approach that takes account of the categorical nature of the data and the problems of making comparisons between countries. This involves the following steps:

1. We estimate self-assessed health inequality using the Cowell and Flachaire (2014) class of measures for several values of the sensitivity parameter  $\alpha$  ranging from  $-2$  (effectively negative infinity) to  $+0.99$  (arbitrarily close to the upper bound of  $\alpha$ ). We do this both for both downward-looking status  $\mathbf{s}$  and upward-looking status  $\mathbf{s}'$ .
2. We compare these measures of health inequality with those that would emerge from conventional inequalities using an arbitrary cardinalisation. In fact we take two different such cardinalisations. The first simply numbers the five health categories from low to high as  $(1, \dots, 5)$  so that, if there are  $(n_1, \dots, n_5)$  observations in each of the categories, the status vector is given by.

$$\mathbf{s}^\uparrow := \left( \underbrace{1, \dots, 1}_{n_1}, \underbrace{2, \dots, 2}_{n_2}, \underbrace{3, \dots, 3}_{n_3}, \underbrace{4, \dots, 4}_{n_4}, \underbrace{5, \dots, 5}_{n_5} \right). \quad (7)$$

To capture the idea of inequality of ill health we also look at the “inverse” case where the the same five health categories are labelled  $(5, \dots, 1)$ .<sup>10</sup>

$$\mathbf{s}^\downarrow := \left( \underbrace{5, \dots, 5}_{n_1}, \underbrace{4, \dots, 4}_{n_2}, \underbrace{3, \dots, 3}_{n_3}, \underbrace{2, \dots, 2}_{n_4}, \underbrace{1, \dots, 1}_{n_5} \right) \quad (8)$$

We then compute  $G_\alpha(\mathbf{s}^\uparrow)$  and  $G_\alpha(\mathbf{s}^\downarrow)$ , using the same values of  $\alpha$  as for the ordinal inequality statistics computed in step 1

---

<sup>10</sup>The different status measures here address the so-called: ‘mirror problem’ discussed by Clarke et al. (2002) who find that concentration indexes for SAH show inconsistent results when ‘health’ or ‘ill-health’ is as a dependent variable.



3. We use rank-correlation analysis to examine the association of country inequality orderings under the alternative definitions of status (3), (4), (7) and (8) for different values of the inequality sensitivity parameter. In other words we look at the correlations between pairs from  $\{I_\alpha(\mathbf{s}), I_\alpha(\mathbf{s}'), G_\alpha(\mathbf{s}^\uparrow), G_\alpha(\mathbf{s}^\downarrow)\}$  for a number of values of  $\alpha$ .
4. We regress  $I_\alpha(\mathbf{s})$  (downward) and  $I_\alpha(\mathbf{s}')$  (upward) on a number of explanatory variables in order to get some insight on the factors associated with high health inequality. Focusing on inequality avoids problems that may arise from systematic response bias between countries.<sup>11</sup> We carry out a similar analysis using the simple cardinalisation  $G_\alpha(\mathbf{s}^\uparrow)$  and the inverse cardinalisation  $G_\alpha(\mathbf{s}^\downarrow)$ .
5. Furthermore we examine possible patterns of health inequality by looking at the way in which (i)  $I_\alpha(\mathbf{s})$  for each country varies and (ii) the way country orderings change as the parameter  $\alpha$  varies.

### 3.4 Inequality regression Analysis

What factors underlie the SAH-inequality rankings for different specifications of status variable? we can use standard regression analysis to address this question, assuming a linear relationship for the variables that may potentially influence health inequalities by country. Given the small number of observations we limit the number of variables to avoid running into limited degrees of freedom. Our dependent variable is a country-specific inequality index  $I_\alpha$ ,  $G_\alpha$  for each country  $i$  and our independent variables are the country-specific income ( $Y_i$ ), median SAH ( $H_i$ ),<sup>12</sup> indicators of institutional performance (such as rule of law, corruption, government effectiveness, democracy),  $Z_i$  and a number of controls for country-specific characteristics (such as the proportion of old age population, female population ratio),  $X_i$ . The estimated equation is as follows:

---

<sup>11</sup>Comparing median categories across countries is regarded as uninformative given that some countries habitually over-report. The term “moderate health” means different things across countries because people’s expectations are different. Some progress has been made in using anchoring vignettes that is increasingly used to correct for this type of bias see Kapteyn et al. (2007) and Rice et al. (2012).

<sup>12</sup>The variables  $Y_i$  and  $H_i$  allow for the possibility of there being a Kuznets curve in income or in health.

$$\{I_\alpha, G_\alpha\} = \gamma_0 + \gamma_1 Y_i + \gamma_2 H_i + \gamma_3 X_i + \gamma_4 Z_i + e_i$$

where  $e_i$  is a random error. We hypothesize whether the four measures of inequality are driven by the same country-specific determinants. Namely, would an improvement in government effectiveness or income reduce health inequalities? Would it do so in the same way irrespective of the measure of inequality or the value of  $\alpha$ ? Is there a non linear effect of income (Kuznets curve) on income or health? Are changes in the composition of the countries population driving the changes in health inequality? The following section reports the specific variables employed in the regression analysis.

## 4 Results

Table 1 provides the summary statistics for 70 country-level observations, organised as follows:

**Dependent variables: Ordinal inequality indices.** The first eight rows give the sample statistics of  $I_\alpha(\cdot)$  the Cowell-Flachaire (2014) for inequality of SAH, for both downward-looking status  $\mathbf{s}$  and upward-looking status  $\mathbf{s}'$ . This is for  $\alpha \in \{-2, -1, 0, 0.99\}$ .

**Dependent variables: Cardinal inequality indices.** The next eight rows give the sample statistics of  $G_\alpha(\cdot)$ , the Generalised Entropy inequality index using the simple cardinalisation of SAH  $\mathbf{s}^\uparrow$ , and the inverse cardinalisation  $\mathbf{s}^\downarrow$ . Again this is for  $\alpha \in \{-2, -1, 0, 0.99\}$ .

**Independent variables.** From the literature we might expect the following to have an effect on health inequality:

- Income levels defined as GDP per capita at the country level;
- the median category of health status: may help to disentangle a potential Kuznets curve effect;
- the proportion of females (fempop\_2005) and the proportion of the elderly (pop65\_2005) in each country: one might expect that health inequality is higher where there are more elderly;

- six measures of institutional characteristics produced by the World Bank (corrupt\_2005, rule\_2005, reg\_2005, goveff\_2005, voiceacc\_2005, ope\_2005): one might expect lower quality to increase inequality, the intuition being that the less efficient the state is for general purposes, the worse it will be at channelling funds to reduce inequalities;
- regional dummies

## Visual and Graphical Analysis

To illustrate whether the different status concepts produce different results in terms of inequality rankings Figure 1 shows the geographical distribution of inequality for the four concepts of status (downward-looking ordinal  $\mathbf{s}$  upward-looking ordinal  $\mathbf{s}'$ , simple cardinal  $\mathbf{s}^\uparrow$ , inverse cardinal  $\mathbf{s}^\downarrow$ ). It employs a central value of the sensitivity parameter,  $\alpha = 0$ . It is immediately clear that the different status concepts reveal quite different inequality patterns: see, for example, the switch between the relative position of India and Russia in inequality rankings as one switches from downward-looking status  $\mathbf{s}$  to upward-looking status  $\mathbf{s}'$ . Figure 2 shows how the distribution of inequality changes with the sensitivity parameter. It is evident that changing the value of  $\alpha$  changes the inequality ranking of the countries, but with no clear patterns: although Russia's inequality ranking continually increases with  $\alpha$ , India achieves the highest inequality ranking at  $\alpha = 0$ , and Brazil at  $\alpha = -1$ .

## Correlation Analysis

An obvious way to check for the overall effect of different status concepts on comparisons of SAH inequality is to examine the extent to which different status concepts produce similar inequality orderings across the 70 countries in the sample. This can be done by computing correlation coefficients for inequality estimates for each possible pair of status concepts; the estimates are, of course, contingent on a particular value of the sensitivity parameter  $\alpha$ . Table 2 provides the Spearman correlation coefficients for inequality rankings using pairwise comparisons of the four different status concepts; this is done separately for each of the following cases  $\alpha = -2, -1, 0, 0.99$ . We find that for low levels of the sensitivity index ( $\alpha = -2$ ) the upward-looking status

| Variable                           | Definition                        | Mean      | Std. Err. |
|------------------------------------|-----------------------------------|-----------|-----------|
| <b>Ordinal inequality indices</b>  |                                   |           |           |
| down_i2                            | $I_{-2}(\mathbf{s})$              | 34.475    | 7.292884  |
| down_i1                            | $I_{-1}(\mathbf{s})$              | 1.184714  | 0.017345  |
| down_i0                            | $I_0(\mathbf{s})$                 | 0.601429  | 0.004337  |
| down_i099                          | $I_{0.99}(\mathbf{s})$            | 34.00371  | 0.237403  |
| up_i2                              | $I_{-2}(\mathbf{s}')$             | 1.123429  | 0.10497   |
| up_i1                              | $I_{-1}(\mathbf{s}')$             | 0.533857  | 0.016055  |
| up_i0                              | $I_0(\mathbf{s}')$                | 0.530571  | 0.005766  |
| up_i099                            | $I_{0.99}(\mathbf{s}')$           | 33.984    | 0.237727  |
| <b>Cardinal inequality indices</b> |                                   |           |           |
| g2                                 | $G_{-2}(\mathbf{s}^\uparrow)$     | 0.059505  | 0.0031    |
| g1                                 | $G_{-1}(\mathbf{s}^\uparrow)$     | 0.044311  | 0.002068  |
| g0                                 | $G_0(\mathbf{s}^\uparrow)$        | 0.036522  | 0.00158   |
| g099                               | $G_{0.99}(\mathbf{s}^\uparrow)$   | 0.032383  | 0.001356  |
| gI2                                | $G_{-2}(\mathbf{s}^\downarrow)$   | 0.127     | 0.0031    |
| gI1                                | $G_{-1}(\mathbf{s}^\downarrow)$   | 0.1018    | 0.0026    |
| gI0                                | $G_0(\mathbf{s}^\downarrow)$      | 0.0881    | 0.0024    |
| gI099                              | $G_{0.99}(\mathbf{s}^\downarrow)$ | 0.0821    | 0.0024    |
| <b>Independent variables</b>       |                                   |           |           |
| gdp_pc2005                         | Gross Domestic Product            | 14037.18  | 379.64    |
| median_cat~y                       | Median Health Status              | 2.1286    | 0.009348  |
| fempop_2005                        | % Female population               | 50.506    | 0.06265   |
| pop65_2005                         | % population over 65              | 9.076714  | 0.143807  |
| corrupt_2005                       | Control of Corruption             | 0.1261    | 0.02663   |
| rule_2005                          | Rule of Law                       | 0.102714  | 0.025796  |
| reg_2005                           | Regulatory Quality                | 0.169857  | 0.024874  |
| goveff_2005                        | Government effectiveness          | 0.196143  | 0.026391  |
| voiceacc_2005                      | Voice and Accountability          | 0.129143  | 0.02603   |
| ope_2005                           | Openness                          | 77.2271   | 0.51429   |
| east_asia~c                        | East Asia Region                  | 0.1       | 0.007479  |
| europe_cen~a                       | Europe and Central Asia           | 0.414286  | 0.012281  |
| latin_amer~n                       | Latin-American region             | 0.1       | 0.007479  |
| middle_eas~a                       | Middle Eastern Region             | 0.057143  | 0.005787  |
| south_asia                         | South Asian Region                | 0.071429  | 0.006421  |
| subsaharan~a                       | Sub Saharan Africa                | 0.2571429 | 0.0108959 |

Table 1: Summary Statistics

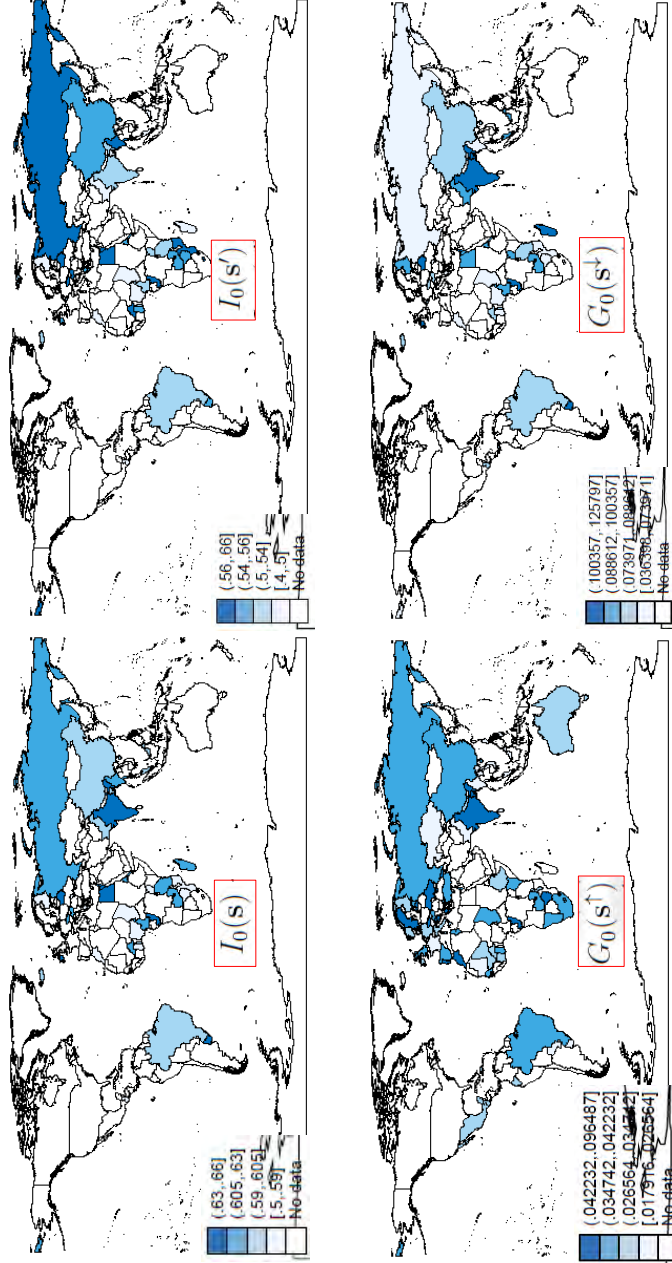


Figure 1: Country distribution of inequality: four different status concepts ( $\alpha = 0$ )

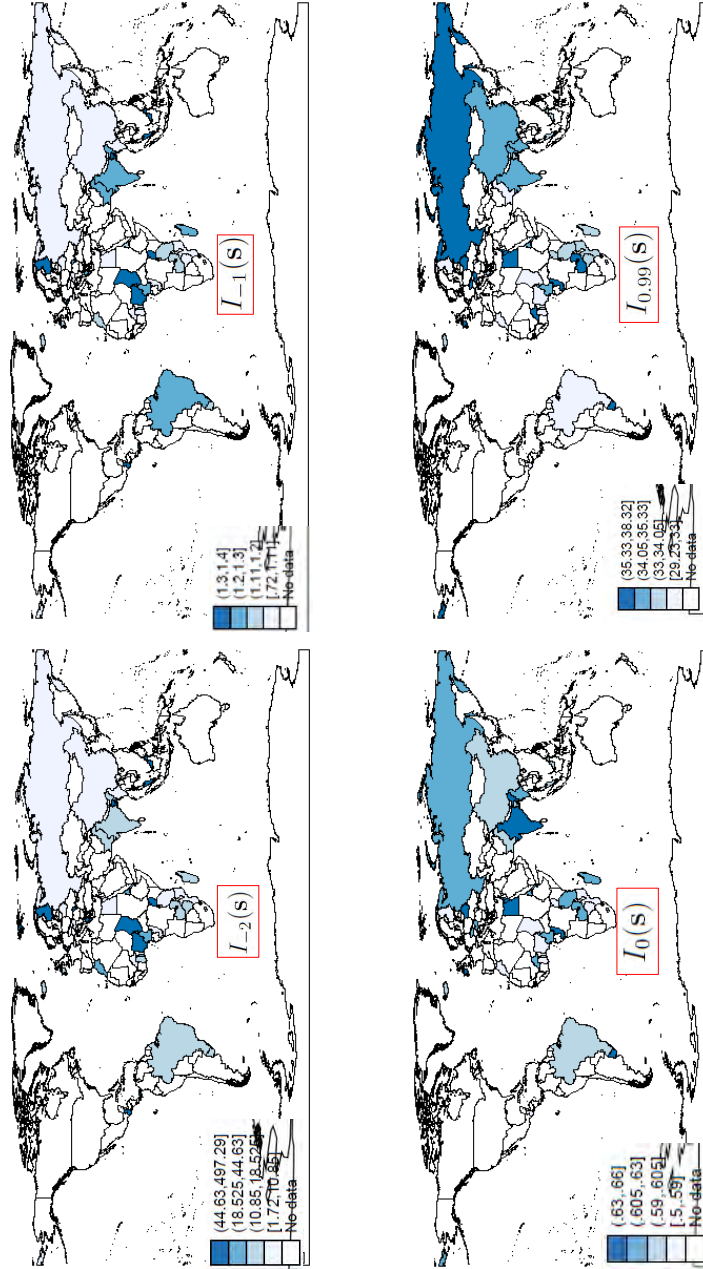


Figure 2: Sensitivity of country distribution of inequality to  $\alpha$ : downward looking status

Table 2: Pairwise correlation for inequality using different status concepts

| $\alpha = -2$                   |                      |                       | $\alpha = -1$                   |                      |                       |
|---------------------------------|----------------------|-----------------------|---------------------------------|----------------------|-----------------------|
|                                 | $I_{-2}(\mathbf{s})$ | $I_{-2}(\mathbf{s}')$ |                                 | $I_{-1}(\mathbf{s})$ | $I_{-1}(\mathbf{s}')$ |
| $I_{-2}(\mathbf{s})$            | 1                    | -0.1274               | $I_{-1}(\mathbf{s})$            | 1                    | -0.8385*              |
| $G_{-2}(\mathbf{s}^\uparrow)$   | 0.2841               | -0.4419*              | $G_{-1}(\mathbf{s}^\uparrow)$   | 0.4785*              | -0.8582*              |
| $G_{-2}(\mathbf{s}^\downarrow)$ | -0.5893*             | 0.3056*               | $G_{-1}(\mathbf{s}^\downarrow)$ | -0.6283*             | 0.1543                |

| $\alpha = 0$                 |                   |                    | $\alpha = 0.99$                   |                        |                         |
|------------------------------|-------------------|--------------------|-----------------------------------|------------------------|-------------------------|
|                              | $I_0(\mathbf{s})$ | $I_0(\mathbf{s}')$ |                                   | $I_{0.99}(\mathbf{s})$ | $I_{0.99}(\mathbf{s}')$ |
| $I_0(\mathbf{s})$            | 1                 | 0.363*             | $I_{0.99}(\mathbf{s})$            | 1                      | 0.7395*                 |
| $G_0(\mathbf{s}^\uparrow)$   | 0.7695            | 0.6389*            | $G_{0.99}(\mathbf{s}^\uparrow)$   | 0.7972*                | 0.798*                  |
| $G_0(\mathbf{s}^\downarrow)$ | -0.2605*          | 0.7389*            | $G_{0.99}(\mathbf{s}^\downarrow)$ | 0.3202*                | 0.3286*                 |

Note: \* significant at the 5% level.

inequality rankings  $I_\alpha(\mathbf{s}')$  correlates negatively with  $G_\alpha(\mathbf{s}^\uparrow)$  but positively with  $G_\alpha(\mathbf{s}^\downarrow)$  (the Generalised Entropy inequality index using, respectively, the simple cardinalisation and the inverse cardinalisation of SAH); but there is no significant correlation of  $I_\alpha(\mathbf{s}')$  with  $I_\alpha(\mathbf{s})$ . Consistently, the ranking using the downward- looking status concept  $I_\alpha(\mathbf{s})$  is negatively correlated with  $G_\alpha(\mathbf{s}^\downarrow)$ . The negative correlations of  $(I_\alpha(\mathbf{s}'), G_\alpha(\mathbf{s}^\uparrow))$  and of  $(I_\alpha(\mathbf{s}), G_\alpha(\mathbf{s}^\downarrow))$  become larger when  $\alpha = -1$ . By contrast, when inequality is evaluated at  $\alpha = 0.99$ , there is a positive correlation between each pairwise combination of inequality orderings: only in this extreme case do we find evidence of a similar pattern of inequality across countries, whatever the status concept.

## Results from Regression Analysis

Now let us examine the apparent role of income, overall health level, social institutions performance and country characteristics in explaining the pattern of health inequalities for the different health- status concepts. In Table 3 we find some evidence that GDP, government effectiveness and the share of female in the population increase health inequalities when a downward-

Table 3: Determinants of Inequality Down (OLS)

| VARIABLES       | (1)<br>down_i2        | (2)<br>down_i1          | (3)<br>down_i0          | (4)<br>down_i099          |
|-----------------|-----------------------|-------------------------|-------------------------|---------------------------|
| median_category | -26.45<br>(19.47)     | -0.251***<br>(0.0359)   | -0.0029<br>(0.0127)     | 0.916<br>(0.658)          |
| gdp_pc2005_01   | 0.00185*<br>(0.00103) | 3.37e-06*<br>(1.89e-06) | -1.06e-06<br>(6.72e-07) | -7.19e-05**<br>(3.47e-05) |
| fempop_2005     | 8.458**<br>(4.193)    | 0.0178**<br>(0.00772)   | -0.00539*<br>(0.00274)  | -0.296**<br>(0.142)       |
| pop65_2005      | 1.493<br>(2.552)      | -0.0064<br>(0.00470)    | 0.0003<br>(0.00167)     | 0.0217<br>(0.0862)        |
| ope_2005        | -0.433<br>(0.411)     | -0.0007<br>(0.000756)   | 0.00035<br>(0.000268)   | 0.0321**<br>(0.0139)      |
| voiceacc_2005   | -12.87<br>(17.91)     | -0.0042<br>(0.0330)     | 0.0119<br>(0.0117)      | 0.334<br>(0.605)          |
| polstab_2005    | 11.32<br>(11.56)      | -0.0102<br>(0.0213)     | -0.0055<br>(0.00756)    | -0.0742<br>(0.391)        |
| goveff_2005     | 91.47**<br>(35.30)    | 0.165**<br>(0.0650)     | -0.0470**<br>(0.0231)   | -2.647**<br>(1.193)       |
| reg_2005        | -125.1***<br>(32.18)  | -0.133**<br>(0.0593)    | 0.0127<br>(0.0210)      | 1.285<br>(1.088)          |
| rule_2005       | -16.77<br>(31.30)     | -0.0237<br>(0.0576)     | 0.0276<br>(0.0205)      | 0.499<br>(1.058)          |
| corrupt_2005    | 11.74<br>(26.68)      | -0.0077<br>(0.0491)     | 0.00363<br>(0.0174)     | 0.832<br>(0.902)          |
| healthexp_2005  | -8.392<br>(5.933)     | -0.0067<br>(0.0109)     | 0.00635<br>(0.00388)    | 0.328<br>(0.200)          |
| Constant        | -278.4<br>(219.8)     | 0.924**<br>(0.405)      | 0.824***<br>(0.144)     | 43.20***<br>(7.427)       |
| Observations    | 69                    | 69                      | 69                      | 69                        |
| R-squared       | 0.391                 | 0.634                   | 0.262                   | 0.336                     |

Standard errors in parentheses: \*\*\* p&lt;0.01, \*\* p&lt;0.05, \* p&lt;0.1



Table 4: Determinants of Inequality Up (OLS)

| VARIABLES       | (1)<br>up_i2            | (2)<br>up_i1            | (3)<br>up_i0              | (4)<br>up_i099            |
|-----------------|-------------------------|-------------------------|---------------------------|---------------------------|
| median_category | 1.605***<br>(0.180)     | 0.231***<br>(0.0313)    | 0.0570***<br>(0.0140)     | 0.934<br>(0.658)          |
| gdp_pc2005_01   | -4.73e-06<br>(9.49e-06) | -2.00e-06<br>(1.66e-06) | -1.50e-06**<br>(7.41e-07) | -7.22e-05**<br>(3.47e-05) |
| fempop_2005     | -0.0722*<br>(0.0387)    | -0.0107<br>(0.00675)    | -0.00609**<br>(0.00302)   | -0.296**<br>(0.142)       |
| pop65_2005      | 0.0402*<br>(0.0236)     | 0.00636<br>(0.00411)    | 0.00209<br>(0.00184)      | 0.0225<br>(0.0862)        |
| ope_2005        | -0.00131<br>(0.00379)   | 0.000593<br>(0.000661)  | 0.000652**<br>(0.000296)  | 0.0323**<br>(0.0139)      |
| voiceacc_2005   | -0.108<br>(0.165)       | -0.0151<br>(0.0289)     | 0.00132<br>(0.0129)       | 0.331<br>(0.605)          |
| polstab_2005    | 0.178<br>(0.107)        | 0.0225<br>(0.0186)      | 0.00393<br>(0.00833)      | -0.0694<br>(0.391)        |
| goveff_2005     | -0.613*<br>(0.326)      | -0.0801<br>(0.0569)     | -0.0440*<br>(0.0254)      | -2.639**<br>(1.193)       |
| reg_2005        | 0.887***<br>(0.297)     | 0.119**<br>(0.0518)     | 0.0365<br>(0.0232)        | 1.292<br>(1.088)          |
| rule_2005       | 0.0200<br>(0.289)       | -0.0338<br>(0.0504)     | -0.0146<br>(0.0226)       | 0.476<br>(1.058)          |
| corrupt_2005    | -0.224<br>(0.246)       | 0.00430<br>(0.0430)     | 0.0218<br>(0.0192)        | 0.842<br>(0.902)          |
| healthexp_2005  | -0.0810<br>(0.0548)     | -0.00655<br>(0.00955)   | 0.00281<br>(0.00428)      | 0.327<br>(0.200)          |
| Constant        | 1.697<br>(2.029)        | 0.552<br>(0.354)        | 0.650***<br>(0.158)       | 43.16***<br>(7.428)       |
| Observations    | 69                      | 69                      | 69                        | 69                        |
| R-squared       | 0.750                   | 0.674                   | 0.487                     | 0.338                     |

Standard errors in parentheses: \*\*\* p&lt;0.01, \*\* p&lt;0.05, \* p&lt;0.1

looking ordinal health-status concept is used and the distributional sensitivity parameter is “bottom-sensitive” ( $\alpha = -2$  and  $\alpha = -1$ ). However, such coefficients reverse sign when  $\alpha$  takes values of 0 and 0.99. We find for  $\alpha = -1$  some evidence of reduction of inequalities after a certain level of health.

In contrast to the previous tables we find now more extended evidence in Table 4 that except for values of  $\alpha$  close to 0.99, median health correlated with health inequalities. Now we find that regulation positively correlates with higher inequality at low levels of  $\alpha$  ( $-2, -1$ ) but it is not significant otherwise. Consistently with previous result, government effectiveness is associated with less inequalities at values of  $\alpha$  close to 0.99. As before, for values of  $\alpha$  of 0.99 we find that female share and GDP reduced health inequalities.

When we use the GE measure using the simple cardinalisation – in Table 5 we find that, irrespective of  $\alpha$ , median income increases inequality which is a finding we only found for up-status inequality when  $\alpha$  was -1. Now, government effectiveness, reduced inequalities in all values of  $\alpha$ . Overall  $R^2$  values are large given the small number of observations. Interestingly, when we examine the determinants of GE on the “inverse” cardinalisation (Table 6) we find that median health reduces health inequality on a similar magnitude irrespective of  $\alpha$  with values ranging from 0.01 to 0.023 suggesting that doubling health status would reduce inequality by between 1 and 2.3%.

## 5 Conclusion

In the case of wealth inequality, getting better estimates is, to some extent, largely a function of getting better data; but in the case of health inequality, more is involved. Even if one has very good, carefully collected data on self-assessed health, almost always one has to deal with the fact that the data will be categorical in nature and require special treatment in order to make reliable inequality comparisons. Here we have followed the Cowell and Flachaire (2014) status-inequality approach that defines a family of inequality indices indexed by a sensitivity parameter  $\alpha$ . The status concept could be downward or upward-looking, and we employ an arbitrary cardinalisation to measure results from generalised entropy indices.

In this paper we find that, for low values of the sensitivity parameter  $\alpha$  (where the index is most sensitive to the bottom of the distribution) status

Table 5: Determinants of Inequality GE (OLS)

| VARIABLES       | (1)<br>g_i2             | (2)<br>g_i1             | (3)<br>g_i0             | (4)<br>g_i099           |
|-----------------|-------------------------|-------------------------|-------------------------|-------------------------|
| median_category | 0.0356***<br>(0.00841)  | 0.0244***<br>(0.00544)  | 0.0188***<br>(0.00406)  | 0.0162***<br>(0.00342)  |
| gdp_pc2005_01   | -6.28e-07<br>(4.44e-07) | -4.20e-07<br>(2.87e-07) | -3.17e-07<br>(2.14e-07) | -2.65e-07<br>(1.81e-07) |
| fempop_2005     | -0.00268<br>(0.00181)   | -0.00177<br>(0.00117)   | -0.00132<br>(0.000874)  | -0.00108<br>(0.000736)  |
| pop65_2005      | 0.000490<br>(0.00110)   | 0.000324<br>(0.000713)  | 0.000247<br>(0.000532)  | 0.000212<br>(0.000448)  |
| ope_2005        | 4.71e-05<br>(0.000177)  | 4.49e-05<br>(0.000115)  | 4.19e-05<br>(8.56e-05)  | 3.90e-05<br>(7.21e-05)  |
| voiceacc_2005   | 0.00258<br>(0.00774)    | 0.000505<br>(0.00501)   | -0.000426<br>(0.00374)  | -0.000986<br>(0.00315)  |
| polstab_2005    | -0.000428<br>(0.00499)  | 9.88e-05<br>(0.00323)   | 0.000320<br>(0.00241)   | 0.000449<br>(0.00203)   |
| goveff_2005     | -0.0305*<br>(0.0152)    | -0.0225**<br>(0.00987)  | -0.0183**<br>(0.00736)  | -0.0163**<br>(0.00620)  |
| reg_2005        | 0.00912<br>(0.0139)     | 0.00760<br>(0.00900)    | 0.00689<br>(0.00671)    | 0.00669<br>(0.00565)    |
| rule_2005       | 0.0202<br>(0.0135)      | 0.0127<br>(0.00875)     | 0.00877<br>(0.00653)    | 0.00671<br>(0.00550)    |
| corrupt_2005    | -2.29e-05<br>(0.0115)   | 0.00199<br>(0.00746)    | 0.00286<br>(0.00556)    | 0.00336<br>(0.00469)    |
| healthexp_2005  | 0.00336<br>(0.00256)    | 0.00218<br>(0.00166)    | 0.00158<br>(0.00124)    | 0.00128<br>(0.00104)    |
| Constant        | 0.0990<br>(0.0949)      | 0.0681<br>(0.0614)      | 0.0526<br>(0.0458)      | 0.0437<br>(0.0386)      |
| Observations    | 69                      | 69                      | 69                      | 69                      |
| R-squared       | 0.371                   | 0.408                   | 0.435                   | 0.456                   |

Standard errors in parentheses: \*\*\* p&lt;0.01, \*\* p&lt;0.05, \* p&lt;0.1

Table 6: Determinants of Inequality Inverse GE (OLS)

| VARIABLES       | (1)<br>g_i2             | (2)<br>g_i1             | (3)<br>g_i0             | (4)<br>g_i099           |
|-----------------|-------------------------|-------------------------|-------------------------|-------------------------|
| median_category | -0.0191**<br>(0.00910)  | -0.0215***<br>(0.00718) | -0.0222***<br>(0.00645) | -0.0233***<br>(0.00646) |
| gdp_pc2005_01   | -5.13e-07<br>(4.80e-07) | -2.82e-07<br>(3.79e-07) | -1.51e-07<br>(3.41e-07) | -6.48e-08<br>(3.41e-07) |
| fempop_2005     | -0.00152<br>(0.00196)   | -0.000887<br>(0.00155)  | -0.000532<br>(0.00139)  | -0.000311<br>(0.00139)  |
| pop65_2005      | -0.000268<br>(0.00119)  | -0.000462<br>(0.000941) | -0.000582<br>(0.000846) | -0.000696<br>(0.000847) |
| ope_2005        | 0.000180<br>(0.000192)  | 8.82e-05<br>(0.000151)  | 2.99e-05<br>(0.000136)  | -1.28e-05<br>(0.000136) |
| voiceacc_2005   | 0.00338<br>(0.00837)    | 0.00335<br>(0.00661)    | 0.00320<br>(0.00594)    | 0.00321<br>(0.00594)    |
| polstab_2005    | -0.00178<br>(0.00540)   | -0.00235<br>(0.00426)   | -0.00276<br>(0.00383)   | -0.00323<br>(0.00384)   |
| goveff_2005     | -0.0215<br>(0.0165)     | -0.0147<br>(0.0130)     | -0.0119<br>(0.0117)     | -0.0112<br>(0.0117)     |
| reg_2005        | -0.00317<br>(0.0150)    | -0.00512<br>(0.0119)    | -0.00584<br>(0.0107)    | -0.00634<br>(0.0107)    |
| rule_2005       | 0.00777<br>(0.0146)     | 0.0101<br>(0.0115)      | 0.0124<br>(0.0104)      | 0.0152<br>(0.0104)      |
| corrupt_2005    | 0.0136<br>(0.0125)      | 0.00715<br>(0.00984)    | 0.00330<br>(0.00884)    | 0.000630<br>(0.00885)   |
| healthexp_2005  | 0.00342<br>(0.00277)    | 0.00285<br>(0.00219)    | 0.00260<br>(0.00197)    | 0.00258<br>(0.00197)    |
| Constant        | 0.219**<br>(0.103)      | 0.176**<br>(0.0811)     | 0.151**<br>(0.0728)     | 0.140*<br>(0.0729)      |
| Observations    | 69                      | 69                      | 69                      | 69                      |
| R-squared       | 0.278                   | 0.341                   | 0.382                   | 0.405                   |

Standard errors in parentheses: \*\*\* p&lt;0.01, \*\* p&lt;0.05, \* p&lt;0.1

does matter to a point that we find no, or even negative, correlation between up and down versions of status. In contrast, for zero or positive values of  $\alpha$  the association becomes positive and large. Similarly, when we compare different measures of status with the corresponding cardinal inequality index (the generalised entropy measure,  $G_\alpha$ ), using the arbitrary 1-to-5 cardinalisation, we find that for negative values of  $\alpha$  there is a negative correlation between up-status (ordinal)  $I_\alpha$  and  $G_\alpha$ ; this flips and becomes positive for  $\alpha \geq 0$ . We find a similar story if we use the inverse – 5-to-1 – cardinalisation for health status; the negative correlation between  $I_\alpha$  and (cardinal)  $G_\alpha$  that is observed for  $\alpha < 0$  becomes positive for high values of the sensitivity parameter  $\alpha$ .

Regression analysis indicates a number of different determinants of inequality measures. Specifically, we find that median health status only increases as  $\alpha$  becomes close to unity. In that case, government effectiveness reduces health inequality. Government effectiveness reduces inequality if a down-status measure is employed, for positive value of  $\alpha$  and the opposite applies for negative values of  $\alpha$ .

Our findings in the Appendix indicate no evidence of a Kuznets curve on health and on GDP, and no association with health expenditure.

The paper has important policy implications. First, our findings suggest that government attempts to reduce health inequalities need to pay specific attention to the nature of the data, and they need to specify the sensitivity to inequality in different parts of the distribution (the parameter  $\alpha$ ) in accordance with the values of a specific society. Second, we find evidence of heterogeneous determinants of different inequality measures. Our results suggest robust evidence that health inequalities are sensitive to some measures of institutional performance (e.g., government effectiveness). However, these results need to be taken with caution given the small number of observations.

## References

- Abul Naga, R. H. and T. Yalcin (2008). Inequality measurement for ordered response health data. *Journal of Health Economics* 27, 1614–1625.
- Abul Naga, R. H. and T. Yalcin (2010). Median independent inequality orderings. Technical report, University of Aberdeen Business School.
- Allison, R. A. and J. E. Foster (2004). Measuring health inequality using qualitative data. *Journal of Health Economics* 23, 505–552.
- Bleichrodt, H. and E. van Doorslaer (2006). A welfare economics foundation for health inequality measurement. *Journal of Health Economics* 25, 945–957.
- Bound, J. (1991). Self-reported versus objective measures of health in retirement models. 26, 106–138.
- Clarke, P. M., G. U. G., M. Johannesson, K. Bingefors, and L. Smith (2002). On the measurement of relative and absolute income-related health inequality. *Social Science and Medicine* 55, 1923–1928.
- Costa-Font, J. and F. A. Cowell (2013). Measuring health inequality with categorical data: some regional patterns. *Research on Economic Inequality* 21, 53–76.
- Costa-Font, J. and C. Hernández-Quevedo (2012). Measuring inequalities in health: What do we know? what do we need to know? *Health Policy* 106, 195–206.
- Costa-Font, J. and C. Hernández-Quevedo (2013). Inequalities in self-reported health: a meta-regression analysis. LSE Health working paper series in health policy and economics 32/2013, LSE Health and Social Care, London School of Economics and Political Science, London, UK.
- Costa-Font, M. and J. Costa-Font (2009). Heterogeneous "adaptation" and "income effects" across self-reported health distribution? *The Journal of Socio-Economics* 38, 574–580.
- Cowell, F. A. (1980). On the structure of additive inequality measures. *Review of Economic Studies* 47, 521–531.

- Cowell, F. A. and E. Flachaire (2014). Inequality with ordinal data. Public Economics Programme Discussion Paper 16, London School of Economics, <http://darp.lse.ac.uk/pdf/IneqOrdinal.pdf>.
- Cutler, D. and E. Richardson (1997). Measuring the health of the United States population. *Brookings Papers on Economic Activity* 2, 217–271.
- de Barros, R. P., F. Ferreira, J. Chanduvi, and J. Vega (2008). *Measuring Inequality of Opportunities in Latin America and the Caribbean*. Palgrave Macmillan.
- De Boer, A. G. E. M., J. J. B. Van Lanschot, P. F. M. Stalmeier, J. W. Van Sandick, J. B. F. Hulscher, J. C. J. M. De Haes, and M. A. G. Sprangers (2004). Is a single-item visual analogue scale as valid, reliable and responsive as multi-item scales in measuring quality of life? 13, 311–320.
- Dowd, J. B. and M. Todd (2011). Does self-reported health bias the measurement of health inequalities in us adults? evidence using anchoring vignettes from the Health and Retirement Study.
- Fleurbaey, M. (2006). Health, equity and social welfare. *Annales d'Économie et de Statistique* 83/84, 21–59.
- Fleurbaey, M. and E. Schokkaert (2009). Unfair inequalities in health and health care. *Journal of Health Economics* 28(1), 73–90.
- Fleurbaey, M. and E. Schokkaert (2012). Health and health care. In M. Pauly, T. McGuire, and P. Pita-Barros (Eds.), *Handbook of Health Economics*, Volume 2. North-Holland.
- Fonseca, L. and A. M. Jones (2003). Inequalities in self-assessed health in the health survey of England. Equity Working Paper 11, University of York.
- Hausman, D. M., Y. Asada, and T. Hedemann (2002). Health inequalities and why they matter. *Health Care Anal* 10, 177–191.
- Heien, T. N. (2015). The relationship between self-rated health and hospital records. *Health Economics in press*.
- Idler, E. L. and Y. Benyamini (1997). Self-rated health and mortality: A review of twenty-seven community studies. *Journal of Health and Social Behavior* 38, 21–37.

- Kapteyn, A., J. P. Smith, and A. Van Soest (2007). Vignettes and self-reports of work disability in the United States and the Netherlands. *The American Economic Review* 97(1), 461–473.
- Koolman, X. and E. van Doorslaer (2004). On the interpretation of a concentration index of inequality. *Health Economics* 13(7), 649–656.
- Kunst, A. E. and J. P. Mackenbach (1994). International variation in the size of mortality differences associated with occupational status. *International Journal of Epidemiology* 23, 1–9.
- Le Grand, J. (1987). Inequalities in health: some international comparisons. *European Economic Review* 31, 182–191.
- Maasoumi, E. (1986). The measurement and decomposition of multidimensional inequality. *Econometrica* 54, 991–997.
- Makdissi, P. and M. Yazbeck (2014). Measuring socioeconomic health inequalities in presence of multiple categorical information. *Journal of Health Economics* 34, 84–95.
- Marmot, M. (2005). Social determinants of health inequalities. *Lancet* 365(9464), 1099–104.
- Rice, N., S. Robone, and P. C. Smith (2012). Vignettes and health systems responsiveness in cross-country comparative analyses. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 175(2), 337–369.
- Sen, A. (2002). Health: perception vs. observation. *British Medical Journal* 324, 860–861.
- Shorrocks, A. F. (1980). The class of additively decomposable inequality measures. *Econometrica* 48, 613–625.
- Tsui, K.-Y. (1995). Multidimensional generalizations of the relative and absolute inequality indices: the Atkinson-Kolm-Sen approach. *Journal of Economic Theory* 67, 251–265.
- 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, 61–87.
- Wagstaff, A., P. Paci, and E. van Doorslaer (1991). On the measurement of inequalities in health. *Social Science and Medicine* 33, 545–557.



- Wagstaff, A. and E. van Doorslaer (2000). Equity in health care financing and delivery. In A. J. Culyer and J. P. Newhouse (Eds.), *Handbook of Health Economics*, pp. 1803–1806. North Holland.
- Zheng, B. (2011). A new approach to measure socioeconomic inequality in health. *Journal Of Economic Inequality* 9, 555–577.

## Appendix

The data are taken from the the World Health Organization’s *World Health Survey*, which is described here:

<http://www.who.int/healthinfo/survey/instruments/en/index.html>

In particular, for our categorical variable, we use the responses to one specific question in this survey’s collection of individual questions about overall health.

“Q2000: In general, how would you rate your health today? The respondent should answer according to how he/she considers his/her health to be and give his/her best estimate. Both physical and mental health must be taken into consideration.”

Table 5 reports the answers to this question across 70 countries. The titles of Columns (1) to (5) give the categories used in the survey: we follow the natural order taking (1) as the best category and (5) as the worst. The total number of respondents is in column (6): if in any row this total exceeds the sum of columns (1) to (5), the difference is attributable to non-response.

Table 7: Responses in WHS to self-rated health question

|    |                | (1)              | (2)         | (3)             | (4)        | (5)             | (6)          |
|----|----------------|------------------|-------------|-----------------|------------|-----------------|--------------|
|    | <i>Country</i> | <i>Very good</i> | <i>Good</i> | <i>Moderate</i> | <i>Bad</i> | <i>Very Bad</i> | <i>Total</i> |
| 1  | Australia      | 487              | 775         | 446             | 74         | 11              | 1793         |
| 2  | Austria        | 423              | 390         | 200             | 36         | 4               | 1053         |
| 3  | Bangladesh     | 494              | 1949        | 2132            | 741        | 228             | 5544         |
| 4  | Belgium        | 252              | 487         | 197             | 48         | 11              | 995          |
| 5  | Bosnia         | 271              | 328         | 279             | 127        | 23              | 1028         |
| 6  | Brazil         | 715              | 1934        | 1881            | 348        | 119             | 4997         |
| 7  | Burkina Faso   | 1254             | 2104        | 1137            | 288        | 36              | 4819         |
| 8  | Chad           | 889              | 1767        | 1371            | 549        | 37              | 4613         |
| 9  | China          | 982              | 1485        | 1215            | 277        | 34              | 3993         |
| 10 | Comoros        | 312              | 631         | 523             | 261        | 30              | 1757         |
| 11 | Congo          | 693              | 550         | 693             | 252        | 33              | 2221         |
| 12 | Côte d'Ivoire  | 661              | 1215        | 955             | 266        | 21              | 3118         |
| 13 | Croatia        | 200              | 302         | 312             | 132        | 43              | 989          |
| 14 | Czech          | 160              | 350         | 311             | 90         | 19              | 930          |
| 15 | Denmark        | 320              | 472         | 166             | 40         | 4               | 1002         |
| 16 | Dominican      | 722              | 1806        | 1560            | 397        | 34              | 4519         |
| 17 | Ecuador        | 650              | 1945        | 1569            | 378        | 53              | 4595         |
| 18 | Estonia        | 70               | 293         | 499             | 134        | 16              | 1012         |
| 19 | Ethiopia       | 2138             | 1549        | 972             | 220        | 47              | 4926         |
| 20 | Finland        | 158              | 395         | 391             | 64         | 3               | 1011         |
| 21 | France         | 255              | 525         | 192             | 34         | 2               | 1008         |
| 22 | Georgia        | 265              | 778         | 1111            | 476        | 125             | 2755         |
| 23 | Germany        | 229              | 582         | 343             | 85         | 13              | 1252         |
| 24 | Ghana          | 1379             | 1433        | 830             | 234        | 46              | 3922         |
| 25 | Greece         | 347              | 325         | 246             | 63         | 19              | 1000         |
| 26 | Guatemala      | 730              | 1790        | 1747            | 472        | 24              | 4763         |
| 27 | Hungary        | 139              | 579         | 503             | 155        | 34              | 1410         |
| 28 | India          | 2159             | 3577        | 2616            | 1311       | 202             | 9865         |
| 29 | Ireland        | 366              | 257         | 101             | 29         | 5               | 758          |
| 30 | Israel         | 519              | 405         | 234             | 40         | 23              | 1221         |
| 31 | Italy          | 182              | 449         | 305             | 46         | 15              | 997          |
| 32 | Kazakhstan     | 265              | 1894        | 2088            | 231        | 18              | 4496         |
| 33 | Kenya          | 1115             | 1798        | 1144            | 309        | 40              | 4406         |
| 34 | Lao            | 1787             | 2005        | 906             | 168        | 17              | 4883         |

---

| <b>...continued</b> |                |                  |             |                 |            |                 |              |
|---------------------|----------------|------------------|-------------|-----------------|------------|-----------------|--------------|
|                     | <i>Country</i> | <i>Very good</i> | <i>Good</i> | <i>Moderate</i> | <i>Bad</i> | <i>Very Bad</i> | <i>Total</i> |
| 35                  | Latvia         | 35               | 244         | 390             | 155        | 32              | 856          |
| 36                  | Luxembourg     | 164              | 346         | 155             | 33         | 2               | 700          |
| 37                  | Malawi         | 2855             | 1334        | 838             | 231        | 33              | 5291         |
| 38                  | Malaysia       | 1194             | 3495        | 1111            | 204        | 12              | 6016         |
| 39                  | Mali           | 1334             | 1526        | 895             | 266        | 11              | 4032         |
| 40                  | Mauritania     | 941              | 1672        | 1024            | 154        | 9               | 3800         |
| 41                  | Mauritius      | 850              | 1677        | 827             | 427        | 104             | 3885         |
| 42                  | Mexico         | 7193             | 18112       | 11221           | 2002       | 218             | 38746        |
| 43                  | Morocco        | 598              | 1454        | 1754            | 821        | 372             | 4999         |
| 44                  | Myanmar        | 1215             | 3412        | 1100            | 157        | 2               | 5886         |
| 45                  | Namibia        | 1622             | 1249        | 863             | 204        | 45              | 3983         |
| 46                  | Nepal          | 1455             | 3908        | 2505            | 767        | 53              | 8688         |
| 47                  | Netherlands    | 189              | 640         | 214             | 41         | 2               | 1086         |
| 48                  | Norway         | 314              | 456         | 140             | 46         | 13              | 969          |
| 49                  | Pakistan       | 1770             | 2996        | 1315            | 263        | 25              | 6369         |
| 50                  | Paraguay       | 1700             | 1920        | 1370            | 133        | 16              | 5139         |
| 51                  | Philippines    | 817              | 5127        | 3759            | 354        | 19              | 10076        |
| 52                  | Portugal       | 62               | 342         | 390             | 180        | 55              | 1029         |
| 53                  | Russia         | 261              | 1102        | 2192            | 770        | 91              | 4416         |
| 54                  | Senegal        | 646              | 1028        | 984             | 217        | 24              | 2899         |
| 55                  | Slovakia       | 400              | 798         | 506             | 94         | 17              | 1815         |
| 56                  | Slovenia       | 90               | 238         | 193             | 53         | 9               | 583          |
| 57                  | South Africa   | 837              | 865         | 467             | 129        | 42              | 2340         |
| 58                  | Spain          | 1051             | 2984        | 1689            | 502        | 117             | 6343         |
| 59                  | Sri Lanka      | 1844             | 3019        | 1535            | 298        | 22              | 6718         |
| 60                  | Swaziland      | 198              | 451         | 508             | 676        | 236             | 2069         |
| 61                  | Sweden         | 262              | 354         | 235             | 133        | 14              | 998          |
| 62                  | Tunisia        | 1236             | 1850        | 1476            | 411        | 58              | 5031         |
| 63                  | Turkey         | 1301             | 4809        | 4035            | 869        | 189             | 11203        |
| 64                  | UAE            | 536              | 472         | 146             | 18         | 6               | 1178         |
| 65                  | UK             | 318              | 498         | 278             | 82         | 17              | 1193         |
| 66                  | Ukraine        | 129              | 659         | 1364            | 594        | 103             | 2849         |
| 67                  | Uruguay        | 725              | 1632        | 547             | 63         | 9               | 2976         |
| 68                  | Vietnam        | 398              | 1368        | 1489            | 225        | 10              | 3490         |
| 69                  | Zambia         | 1436             | 1292        | 816             | 228        | 39              | 3811         |
| 70                  | Zimbabwe       | 837              | 1263        | 1501            | 385        | 65              | 4051         |

Table 8: Inequality and health expenditure ( $\alpha = 0$ )

| VARIABLES      | (1)<br>down_i0       | (2)<br>up_i0          | (3)<br>g0               |
|----------------|----------------------|-----------------------|-------------------------|
| healthexp_2005 | 0.00136<br>(0.00184) | -0.00388<br>(0.00241) | -0.000327<br>(0.000671) |
| Constant       | 0.593***<br>(0.0128) | 0.556***<br>(0.0167)  | 0.0387***<br>(0.00466)  |
| Observations   | 70                   | 70                    | 70                      |
| R-squared      | 0.008                | 0.037                 | 0.003                   |

Standard errors in parentheses: \*\*\* p<0.01, \*\*  
p<0.05, \* p<0.1

| Table 9: Kuznets curve on GDP                                     |                         |                         |                         |
|---|-------------------------|-------------------------|-------------------------|
| VARIABLES   | (1)<br>down_i0          | (2)<br>up_i0            | (3)<br>g0               |
| gdp_pc2005_01   | -3.96e-07<br>(7.40e-07) | -1.03e-06<br>(9.77e-07) | -1.90e-07<br>(2.28e-07) |
| gdp2  | 0<br>(0)                | 0<br>(0)                | 0<br>(0)                |
| Constant  | 0.607***<br>(0.00703)   | 0.541***<br>(0.00928)   | 0.0387***<br>(0.00265)  |
| Observations  | 70                      | 70                      | 70                      |
| R-squared   | 0.025                   | 0.038                   | 0.023                   |
| Standard errors in parentheses: *** p<0.01, **<br>p<0.05, * p<0.1 |                         |                         |                         |

| Table 10: Kuznets curve on health                                 |                      |                      |                      |
|---|----------------------|----------------------|----------------------|
| VARIABLES   | (1)<br>down_i0       | (2)<br>up_i0         | (3)<br>g0            |
| median_category   | -0.00365<br>(0.0952) | -0.00365<br>(0.0952) | -0.0216<br>(0.0296)  |
| media2  | -0.00121<br>(0.0198) | -0.00121<br>(0.0198) | 0.00826<br>(0.00616) |
| Constant  | 0.615***<br>(0.111)  | 0.615***<br>(0.111)  | 0.0440<br>(0.0347)   |
| Observations  | 70                   | 70                   | 70                   |
| R-squared   | 0.010                | 0.010                | 0.278                |
| Standard errors in parentheses: *** p<0.01, **<br>p<0.05, * p<0.1 |                      |                      |                      |