# Poor Health Reporting? Using Vignettes to Recover the Health Gradient by Wealth

*Laura Rossouw[1]*
*Teresa Bago d'Uva[2,3]*
*Eddy van Doorslaer[2,3,4,5]*

[1]*ReSEP, Economics Department, Stellenbosch University, South Africa*
[2] *Erasmus School of Economics, Erasmus University Rotterdam, The Netherlands*
[3]*Tinbergen Institute, Rotterdam, The Netherlands*
[4]*Institute of Health Policy and Management, Erasmus University Rotterdam*
[5]*Stellenbosch Institute for Advanced Studies, Wallenberg Research Centre at Stellenbosch University, South Africa*

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

Contact: discussionpapers@tinbergen.nl

More TI discussion papers can be downloaded at http://www.tinbergen.nl

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 598 4580

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900

# Poor health reporting?

## Using vignettes to recover the health gradient by wealth

Laura Rossouw[1], Teresa Bago d'Uva[2, 3], Eddy van Doorslaer[2,3,4,5]

27 February 2017

## Abstract

In spite of the well-known wide disparities in wealth and in objective measures of health like mortality in countries like South Africa, health inequality by wealth in self-reported health measures appears to be nearly non-existent. We test and correct for reporting heterogeneity in sixteen domains of self-assessed health by wealth and race among elderly South Africans using anchoring vignettes. We find that significant reporting differences between high and low wealth groups lead to severe underestimation of the health-wealth gap: poorer individuals rate the same health relatively higher than richer. Using hierarchical ordered probit (HOPIT) modeling, we show that a significant and substantial health disadvantage of the poor emerges after correction. We also address the question whether and how health inequality and reporting heterogeneity are confounded by race. We find that within race groups - especially among Blacks but also among Whites - reporting heterogeneity leads to the underestimation of the health inequalities between richest and poorest. Finally, we show that the apparent Black (*vs* White) health disadvantage within the top wealth quintile disappears once we correct for reporting tendencies. All in all, our findings suggest that reporting tendencies are an important source of bias in the measurement of health disparities and that anchoring vignettes and HOPIT models can play a role in correcting for these biases.

Key words: self-assessed health; vignettes; health measurement; inequality; South Africa

Affiliations:
[1] ReSEP, Economics Department, Stellenbosch University
[2] Erasmus School of Economics, Erasmus University Rotterdam
[3] Tinbergen Institute, Rotterdam
[4] Institute of Health Policy and Management, Erasmus University Rotterdam
[5] Stellenbosch Institute for Advanced Studies, Wallenberg Research Centre at Stellenbosch University

**Introduction**

Health is distributed unequally across income groups. The phenomenon that the better off also enjoy on average better health is pervasive and persistent globally and South Africa is no exception. Studies in South Africa have consistently found worse health outcomes amongst the poor relative to the wealthier (Ardington and Gasealahwe 2014; Ataguba et al. 2011; Ataguba and McIntyre 2013; Ataguba 2013; Zere and McIntyre 2003).

Studies using household data to measure health outcomes usually rely on self-reported information rather than objectively measured health data or biomarkers. Consistently, significant associations have been found between subjective ratings of overall health status of individuals and their subsequent survival in both the industrialized and developing world (Idler and Benyamini 1997; Van Doorslaer and Gerdtham 2003; De Salvo et al. 2006; Jylhä et al. 2006; Frankenberg and Jones 2004). Similarly, Ardington and Gasealahwe (2014) have found self-assessed health (SAH) to be a significant predictor of two-year mortality in South Africa. These self-reports have also been found to be highly correlated with clinically measured biomarkers that lead to mortality (Jylhä et al. 2006) and predictors of health care utilization (Van Doorslaer et al. 2004).

While the relationship between SAH measures and mortality seems well-established, and as they are less costly and less invasive to collect than objective measures of health, health self-reports have also been found to be prone to systematic differences in reporting behaviour across different socioeconomic groups. Such differences imply that health disparities measured using SAH outcomes could possibly be biased (Etilé and Milcent 2006; Hernández-Quevedo et al. 2004; Lindeboom and Van Doorslaer 2004). The most common method of capturing overall SAH is through a categorical and ordinal variable. An individual is asked to classify health as either "Very poor", "Poor", "Fair", "Good" or "Excellent". Persons from different sub-groups could have a different interpretation of what it means to have "poor" or "excellent" health. One possible reason is the use of different comparison or reference groups (Boyce and Harris 2011; Harris et al. 2011). Individuals from poorer health communities may report themselves to be relatively well-off compared to their reference group, even if their health compares poorly to the overall population (Bago d'Uva et al. 2008b; Etilé and Milcent 2006).

These differences in the evaluation of SAH are usually referred to as *reporting heterogeneity*. If they are systematic across certain sub-groups, it then becomes problematic to use SAH for health comparisons (Etilé and Milcent 2006; Hernández-Quevedo et al. 2004; Lindeboom and Van Doorslaer 2004). Reporting heterogeneity is present when, at a fixed level of health, certain

population sub-groups systematically under- or over-report compared to their true, unobserved level of health. Importantly, systematic reporting heterogeneity by socioeconomic status will bias the measurement of the socioeconomic gradient in health. A first goal of this paper is to test whether inequalities in SAH by wealth in South Africa are underestimated as a result of reporting tendencies. This is done by using ratings of so-called health anchoring vignettes, combined with individuals' ratings of their own health, to estimate hierarchical ordered probit models. We do so using data taken from SAGE (the WHO Study on global AGEing and health), a nationally representative sample of persons aged 50 and older in South Africa, collected in 2008 (World Health Organization 2008).

The second objective is to investigate the presence of health reporting heterogeneity within and across race groups in South Africa. South Africa has a history of economic and political segregation by racial lines that was institutionalized during the Apartheid era. During this period, the mobility of race groups other than the minority White population was severely restricted. Particularly the Black African population group was disadvantaged, with a large part of this population's movement restricted to the homelands. These were areas demarcated by the South African government, situated along the countries peripheries with their own (severely underfunded) health departments. Their access to urban areas where better health care and economic opportunities were available was restricted and regulated. The apartheid policy only came to an end in 1994 when the first democratically elected government came to power (Coovadia et al. 2009). This history of racial segregation led to deeply entrenched income/wealth, racial and health disparities among South Africans. These disparities are well-known and documented, for instance, by Leibbrandt et al (2011) who show that the bottom quintiles mostly consist of members of the Black African population group while the White, Coloured and Asian/Indian groups are much more concentrated in the top quintiles.[1]

A number of studies have reported racial health disparities favouring the White population in South Africa (Ardington and Gasealahwe 2014; Charasse-Pouélé and Fournier 2006; Lau and Ataguba 2015), even after controlling for income. Ardington and Gasealahwe (2014), for instance, using nationally representative data for all age groups from the National Income Dynamics Study (NIDS), and after controlling for household expenditure and asset wealth, find that the White and Coloured population report significantly better overall SAH than the Black African population. Similar results of the Black African and White difference in self-reported health are found in Lau and Ataguba's (2015) analysis of NIDS. The intertwined relationship between race, wealth and

---

[1] And the same can be seen in the description of our data in Table 1 in the next section.

health in South Africa therefore means that wealth-related reporting heterogeneity in SAH may also be confounded by race.

**Reporting Tendencies and the Health Gradient**

Several authors have tested for reporting heterogeneity in self-reported health measures. The majority of these have used data from high income countries (e.g. Etilé and Milcent 2006; Hernández-Quevedo et al. 2004; Humphries and Van Doorslaer 2000; Lindeboom and Kerkhofs 2009; Lindeboom and Van Doorslaer 2004), while far fewer studies have used data from developing countries (e.g. Bago d'Uva et al. 2008b; Zhang et al. 2015; Molina 2016). Vulnerable sub-groups are often found to systematically over-report their health (or underestimate ill health). Mu (2014) explores health reporting differences between two provinces in China, one poor and one more affluent. She finds that persons from the poor province systematically underestimate their ill-health. Using French data, Etilé and Milcent (2006) find that the poor are relatively optimistic about their health, as do Bago d'Uva et al. (2008b) for Indonesia, India and China. Some authors have also found that individuals with low levels of education are likely to report better health levels than the higher educated, given the same true latent level of health (Bago d'Uva et al. 2008a; Molina 2016).

Researchers have cited various reasons for these discrepancies between true health status and SAH status. One reason for this underestimation is due to a comparison with different reference groups (as explained in the Introduction). A second possibility is the asymmetry of health information to which various sub-groups have access (Bonfrer et al. 2014). Wealthier individuals in South Africa may have access to private health care which would enable them to go for regular doctors' visits and check-ups. This access will allow them to be diagnosed with chronic conditions of which they were unaware. Better access to health information could lead to greater awareness of their ill-health. Better health knowledge may in turn also affect health expectations (Bonfrer et al. 2014). Sen (2002) gives an extreme example of a person growing up in a poor community where disease incidence is high and health facility access low. Such a person might view symptoms as part of a normal, healthy condition, while they could perhaps easily be prevented or remedied with appropriate treatment.

If poorer individuals systematically underestimate their ill health, this will be reflected in the reporting of SAH questions. Health inequalities based on such measures will then underestimate the gap with wealthier quintiles. Several authors have noted this phenomenon. Some have relied on health vignettes to directly examine reporting behaviour, others have provided more indirect

evidence of reporting tendencies by comparing gradients using more objective, observed measures. Bago d'Uva et al. (2008b) is an example of the first approach. Using vignettes data, they directly test for systematic reporting differences across various socioeconomic groups in India, Indonesia and China. In all three countries, they find that there are systematic differences in the reporting behaviour of the poor and the non-poor, and that the impact of income on health is underestimated if self-reported data is used. However, except for China effects are quite small. Nonetheless, they find that there is reason for concern that reporting heterogeneity could lead to bias in measuring health disparities across income groups.

Bonfrer et al. (2014) provide indirect evidence. They are concerned with measuring the "need for care", and demonstrate that there are reporting tendencies by comparing the measured degree of inequality (using concentration indices) in objective health measures (like stunting and underweight) and in self-reported health measures. In World Health Survey data from 18 countries in Sub-Saharan Africa (including South Africa), they find much higher degrees of health inequality by wealth when using objective health measures, concluding that the use of subjective health measures leads to a severe underestimation of health disparities across income groups in SSA.

In what follows, we formally test for wealth- and race-related reporting heterogeneity in SAH measures in South Africa and examine its implications for measuring health inequalities. This is done using the vignettes approach, as in Bago d'Uva et al. (2008a,b). If reporting heterogeneity is present, and systematically associated with wealth and race, then health disparity measurement across these sub-groups may be biased.

**Methodology: Anchoring Vignettes and HOPIT model**

In the presence of reporting heterogeneity, analyses of inequalities in SAH face an identification problem: any measured inequalities in SAH represent a mix between associations with true health status and reporting heterogeneity (Bago d'Uva et al. 2008a,b; King et al. 2004). This identification problem can only be solved with additional data, either on 'true' health status or on reporting behaviour.

One approach to identifying such reporting heterogeneity is to use more objective measures of health, as proxy indicators for true health status (Etilé and Milcent 2006; Lindeboom and Van Doorslaer 2004). This method assumes that the proxy indicators capture all variation in true health status associated with the individual characteristics of interest, such that any remaining systematic variation in SAH with these characteristics is attributable to health reporting behaviour (Kerkhofs

and Lindeboom 1995; Lindeboom and Van Doorslaer 2004). The applicability of this method in household surveys is, however, problematic if the available proxy indicators are themselves self-reported, and so possibly also biased (Baker et al. 2004). Of particular importance for our research question, in our data the proportion of richer individuals who reports having at least one chronic condition is larger than that among the poor, suggesting lower awareness for the latter group.[2] Truly objective measures, like biomarkers, are often too expensive to collect in household surveys and therefore not available. And, even if available, they might not enable a sufficiently comprehensive measurement of health, leaving room for remaining association between SAH and individual characteristics that cannot be attributed solely to reporting heterogeneity.

In this paper, we use the alternative anchoring vignettes approach. An anchoring vignette is a description of the level of health of a hypothetical person. Since this description is fixed across individuals, all systematic variation in vignette ratings with respect to individual characteristics is attributed to reporting heterogeneity. Self-assessed (inequalities) can then be purged from this reporting heterogeneity (King et al. 2004). As mentioned above, the vignettes approach has been used to identify and correct for reporting heterogeneity in SAH in studies on Asia (Bago d'Uva et al. 2008b; Guindon and Boyle 2012); several countries in Europe (Bago d'Uva et al. 2008a; Jürges 2007; Peracchi and Rossetti 2012) and the USA (Dowd and Todd 2011). In the remainder of this section, we describe in detail the data and methods used in this paper.


Data

We use data representative of the South African elderly (50+) population, taken from the WHO's Study on global AGEing and adult health (SAGE). This is a multi-country study, recorded in 2008 and containing approximately 3400 observations for South Africa. The data were collected with the aim of exploring and comparing the health and socioeconomic characteristics of several low and middle-upper income country populations. Data were collected on health status, chronic conditions, disability, life expectancy, health behaviour and health care utilization (He et al. 2012). South African aged 50 and older in 2008 grew up in the period of racial segregation prior to democratization in 1994, which will have had a profound effect on their wealth creation and access to health services during adolescence and early adulthood.

---

[2] Namely, the proportion of individuals who reported at least one of various chronic conditions (including arthritis, angina, stroke, hypertension, diabetes, lung disease, asthma and depression) is 5.65 amongst the poorest 20% of the sample and 8.3 amongst the richest 20%.

*Self-Assessed Health and Health Vignettes*

The SAGE questionnaire asks respondents to rate the difficulty they have performing in each of eight health domains on a 5-point scale as: 1= No difficulty, 2 = Mild difficulty, 3= Moderate difficulty, 4 = Severe difficulty, or 5 = Extreme difficulty. The domains include mobility, self-care, pain, cognition, interpersonal activities, sleep and energy, affect, vision.[3] In the case of mobility, for instance, the respondent is asked how much difficulty s/he had with moving around in the last 30 days. A similar question structure is applied to the other domains. Similar measures of health, such as depression or limitations in activities of daily living indices have been found to be highly correlated with mortality in South African data (Ardington and Gasealahwe 2014). SAGE collects information on two aspects within each of the eight domains. For instance, in the domain of vision, on farsightedness and nearsightedness. A detailed description of the 16 health aspects considered in this study, as well as their specific questions can be found in the supplementary material (Table S1). For ease of reference, we refer to these as 16 health domains from here onwards.

Subsets of randomly chosen respondents are presented with a selected set of vignettes. For each health domain, the respondent is asked to rate five different vignettes, each representing a different level of health and functionality. One example in the domain of mobility is: "Alan is able to walk distances of up to 200 meters without any problems but feels tired after walking one kilometer or climbing up more than one flight of stairs".[4] Respondents are then asked to rate the health of each of the vignettes in the respective domain, using the same one to five scale as is used in the self-assessment questions.

*Socio-Demographic Variables*

Our covariates include gender, age, level of educational attainment, marital status, race and urban residence. We account for four race categories as defined by Statistics South Africa: Black African, White, Coloured and Indian/Asian. Wealth is accounted for by including four quintile dummies. The wealth variable is created using the household's durable assets, characteristics of their dwelling and whether they had access to basic services such as sanitation and water (He et al. 2012). This is considered a better representation than income of a person's living standards when the sample consists of both retired and non-retired individuals (Grol-Prokopczyk et al. 2011).

---

[3] The selection of domains was based on the World Health Survey (WHS), and was guided by validity in terms of intuitive, clinical, and epidemiological concepts of health; correspondence to the conceptual framework of the International Classification of Functioning, Disability and Health; and comprehensiveness (Salomon et al. 2003).

[4] The full description of all vignettes can be found on the WHO website http://www.who.int/healthinfo/sage/en/

Table 1 shows sample averages of the covariates used in our analyses, by wealth quintile. Females are more likely to be in the lower quintiles, while age is fairly equally distributed across wealth quintiles. Individuals in higher quintiles are significantly more likely to be married, urban and higher educated. Predictably, race is very unequally distributed across quintiles: in the poorest quintile, the great majority (94%) of respondents are Black Africans, while this population group only represents 42% of the richest quintile. Asian, Indian and White, on the other hand, are more concentrated in the top wealth quintiles.


[Table 1 here]


Hierarchical Ordered Probit Model – HOPIT

We use the hierarchical ordered probit model (HOPIT) proposed by King et al. (2004) to identify and correct for reporting heterogeneity. This model is an extension of the ordered probit model and consists of two components - the *reporting behaviour component* and the *own health component*. Each of these is modeled as a generalized ordered probit model, with heterogeneous cut-points (rather than assumed constant as in the standard ordered probit).

The *reporting behaviour component* uses vignette ratings in order to identify cut-points as functions of individual characteristics. Formally, suppose that $H_{ij}^v$ represents the true latent level of health for hypothetical vignette $j$ ($j$=1,…,5), for respondent $i$. $H_{ij}^v$ is assumed to be the same for all individuals, apart from random error:

$$H_{ij}^v = \alpha_j + \varepsilon_{ij}^v, \quad \text{with} \quad \varepsilon_{ij}^v \sim N(0, \sigma_v^2) \tag{1}$$

This reflects the first identifying assumption of the vignette methodology: the *vignette equivalence* assumption requires that there are no systematic differences across individuals in their perceptions of the level of functioning described in the vignettes. We denote the observed categorical rating of the health of vignette $j$ by respondent $i$ as $AH_{ij}^v$. This relates to the latent true health level, $H_{ij}^v$, in the following way:

$$AH_{ij}^v = m \ if \ s_i^{m-1} \leq H_{ij}^v \leq s_i^m \tag{2}$$

where $m = 1, \dots, 5, s_i^0 < s_i^1 < s_i^2 < s_i^3 < s_i^4 < s_i^5$, and $s_i^0 = -\infty, s_i^5 = \infty$.

Finally, reporting heterogeneity is accommodated in this model by defining the cut-points $s_i^m$ as functions of the vector of individual characteristics $X_i$ (which includes wealth, race and other socio-demographic variables, besides a constant term). Identification of reporting heterogeneity in this component derives from the vignette equivalence assumption, which enables the exclusion of individual characteristics from Eq. 1 above and, consequently, inclusion of them in the following:

$$s_i^m = X_i'\gamma^m \tag{3}$$

A special case of this model is one with constant cut-points, i.e., no reporting heterogeneity. Testing *reporting homogeneity* according to one or a subset of variables included in *X* can therefore be done by testing significance of the respective (sets of) coefficients in the vectors $\gamma^m$, m=1,…,4.

The second – *own health* - component of the HOPIT model is specified as a generalized ordered probit with variable cut-points identified by the vignettes in the *reporting behaviour* component. The true own latent health level is typically modeled as a function of the same individual characteristics included in the cut-points:[5]

$$H_i^S = X_i'\beta + \varepsilon_i^s, \qquad \varepsilon_i^s \sim N(0, \sigma^2) \tag{4}$$

Finally, similar to the vignette ratings, the own health ratings relate to own latent true level as:

$$SAH_i = m \; if \; s_i^{m-1} \leq H_i^s \leq s_i^m \tag{5}$$

where the cut-points are as defined in Eq. 3 above. This equality reflects the *response consistency* assumption - individuals are assumed to use the same response scales when rating the vignettes and their own health. Under the two assumptions, the HOPIT model uses vignette ratings to identify and correct for reporting heterogeneity and returns associations between $X_i$ and health that are purged of reporting heterogeneity, i.e., represented by the coefficients in the vector $\beta$ in Eq. 4.

Evidence on the identifying assumptions of the vignette methodology is limited and shows mixed results. On vignette equivalence, see Bago d'Uva et al. (2011), Kristensen and Johansson (2008), Murray et al. (2003), Grol-Prokopczyk et al. (2011, 2015) and Rice et al. (2012); on response

---

[5] Put another way, applications of the HOPIT model typically allow, and so correct, for any potential reporting heterogeneity according to all variables included in the own health Eq. 4.

consistency, see Bago d'Uva et al. (2011), Datta Gupta et al. (2010), Grol-Prokopczyk et al. (2011, 2015), and Van Soest et al. (2011). In the final section of this paper, we discuss possible implications for our results of departures from these assumptions. Identification of the HOPIT model specified by Eqs. 1-5 further requires scale and location normalizations. We normalise $\sigma_v^2$ to 1 and $\alpha_1$ to zero, with no loss of generalization. The two components of the model are estimated jointly. The own health component makes use of ratings of own health for the whole sample, while the reporting behaviour component uses data from the random subsamples of individuals who rate vignettes in the respective domain.

**Results**

We analyse inequalities in SAH among elderly South Africans and the extent to which they might be affected by different reporting tendencies across sub-populations. We first focus on wealth-related health inequalities, which are expected to be substantial given the large economic inequalities which resulted from the Apartheid regime (Özler 2007). In the first two sub-sections, we address the question: does reporting heterogeneity cause an underestimation of health inequalities by wealth in South Africa? Since wealth inequality emanated from a regime that enforced the separation of race groups to the advantage of the White population, we aim to gain a deeper understanding of wealth-related health inequalities by exploring the role of race. To this end, we use in the subsequent sections a more complete specification to analyse in greater detail: *i)* health-wealth associations within race groups, as well *ii)* health-race associations among equally wealthy individuals. In the latter, we control for the other wealth-health groups but focus our attention on the Black African versus White population groups. This is because of the enduring overlap between race and patterns of privilege and poverty, with the lion's share of the poor being Black African and White South Africans being predominantly affluent.

We apply both specifications of the HOPIT model to vignettes and SAH ratings in the 16 domains described above. We use these estimated models by domain to: *i)* test for reporting heterogeneity and *ii)* to estimate health inequalities (un)corrected for reporting heterogeneity (more detail about all of these procedures given below). By comparing corrected with uncorrected health inequalities, we are able to assess the importance and the direction of any reporting bias (King et al. 2004; Rice et al. 2012; Tandon et al. 2003).

Inequalities in Self-Assessed Health by Wealth

We start with a naïve analysis of inequalities in SAH by wealth, ignoring any reporting heterogeneity. This uses a standard ordered probit model with constant cut-points, with the covariates wealth, race, age, gender, marital status, urbanization and educational achievement as defined above.[6] We use this model to predict the probability that an individual reports any difficulty in the respective health domain (i.e., categories 2, mild difficulty, to 5, extreme difficulty). Figure 1 shows average probabilities for Q1 and Q5, keeping other variables constant, for all health domains. We see very small differences between the levels of self-reported difficulties of these groups. For certain domains, such as vigorous activity, depression and anxiety, the wealthier report even *more* difficulties than the poor. Taking these self-reports at face value would lead to the overall conclusion that there is little or no health disadvantage for poor South Africans, compared to their richer counterparts. In the next sub-section, we examine whether these patterns may be related to reporting biases or tendencies.

[Fig. 1 here]

*Inequalities in Health by Wealth, Corrected for Reporting Heterogeneity*

We now extend the analysis of inequalities in health by wealth, by relaxing the assumption of reporting homogeneity, making use of vignettes data and the HOPIT model for each of the 16 domains. We keep the same covariates as in the previous section. The focus is on health inequalities and reporting heterogeneity by wealth but our specification allows for heterogeneity according to *all* covariates (Eq. 3 of the HOPIT model defined above).[7] We present the results for the poorest (Q1) relative to the richest (Q5) quintile, to illustrate and highlight the differences between the two extremes in the wealth distribution in South Africa.

Estimating the reporting (or vignette) component of the HOPIT model enables a direct test of the presence of reporting heterogeneity: we test the null hypothesis that cut-points of individuals in Q5 are the same as those in Q1, conditional on the remaining covariates. This is a test of the equality of the respective coefficients in the vectors $\gamma^m$ ($m$=1,…,4) in Eq. 3.[8] Table 2 presents p-

---

[6] Full estimation results of the ordered probit model for the mobility domain can be found in Table S2 of the supplementary material.

[7] For illustration, estimation results of the HOPIT model for the mobility domain are shown Table S2 of the supplementary material.

[8] In practice, since Q5 is the wealth reference category in our model, this corresponds to testing the significance of the coefficients of Q1, jointly across the four cut-points.

values of this test by health domain. At a 5% significance level, we can reject the null hypothesis that the wealthiest and the poorest use the same cut-points, in 9 of the 16 health domains.[9]

[Table 2 here]

Given the presence of reporting heterogeneity, self-reported data are likely to biased and, consequently, so are estimated health inequalities by wealth. The own health component of the HOPIT model uses the cut-points identified in the reporting component, Eq. 3, and returns associations between wealth (and other covariates) and true latent health status that are purged of reporting heterogeneity, Eq.4. We then use results of the HOPIT model to estimate and compare health inequalities corrected/uncorrected for reporting heterogeneity.

The measure of wealth-related health inequality that we present here is the average marginal effect of belonging to Q1 *vs* Q5 on the probability of having any difficulty (i.e., categories 2 "Mild difficulty" to 5 "Extreme difficulty"), keeping the other covariates fixed.[10] For each domain, we compute the average marginal effect on that probability: *i)* using ordered probit models – that is, uncorrected for reporting heterogeneity - and *ii)* using the estimated HOPIT models and imputing the same fixed cut-points across individuals – and so correcting for reporting heterogeneity. The fixed cut-points are those of a reference individual (an unmarried Black African male; in wealth quintile one; aged 62, the average age in the sample; who did not complete primary school; and lives in a rural area).

To summarize the large number of estimates generated by this procedure in a concise graphical way, the results obtained are depicted in Fig. 2 for all health domains in radar chart (with estimates detailed in supplementary results, Table S4). For instance, for the domains of depression and anxiety, according to the ordered probit model, individuals in Q1 are 8 and 6 percentage points less likely to report any difficulty than individuals in Q5, keeping other variables fixed. Correcting for reporting heterogeneity using the HOPIT model, we do not find a significant gap between the richest and the poorest quintiles in those health domains. The graph shows that across all

---

[9] The detailed cut-points coefficients are reported in the supplementary material (Table S3). The table shows the position of the cut-points between the categorical options of the vignettes for individuals in Q1 relative to individuals in Q5. For instance, the positive and significant coefficient for cut-point 1 in the domain mobility can be interpreted as individuals in Q1 having a significantly higher cut-point between the categories none and mild health difficulties, than those in Q5. This means that, given the true level of health of the vignette, $H_{ij}^v$, individuals in Q1 are systematically more likely to assess the vignette as having no health problems than individuals in Q5, indicating a relative optimism in their health evaluation.

[10] We follow the usual terminology by referring to the magnitudes of the associations of health with covariates as marginal effects, even if these should not be interpreted as causal effects.

health domains, the health gap by wealth (i.e., the health advantage in favour of the rich) grows after reporting correction. Before correction, the poorest are only significantly more likely to report health problems in the appearance domain. A very different picture emerges after correction: in eight of the 16 domains (mobility, relationship, conflict, farsightedness, nearsightedness, self-care, appearance, and learning), the health by wealth gap becomes significant (at a 10% level). Moreover, the only instances where the poorest reported better health than the wealthiest disappear. These results clearly demonstrate that wealth-related health inequalities can be substantially underestimated when using uncorrected SAH measures.

[Fig. 2 here]

Reporting bias in Health Inequalities by Wealth Among Black Africans

In order to further unravel the complex relationships between race, wealth, health and health reporting, we now turn to a more complete specification. In this sub-section, we examine the race-specific health-wealth associations, focusing on the Black African and White population groups, the two extremes of those disadvantaged and advantaged during the apartheid regime. We categorize wealth of the White population as Q5 versus Qs 2-4.[11] For comparability across races groups, and also crucial for analyses in subsequent sections, we categorize wealth of the Black African population as Q1, Q2-4 and Q5. We consider the same wealth categories for the remaining races (Coloured, Asian, and Indian). Small samples sizes do not make it possible to distinguish between the wealth effect of Asian/Indian *vs* Coloured but we do include an additional dummy variable for Coloured to allow for a differential health (reporting) effect for this group. In sum, we consider the following race/wealth variables: Coloured/Asian/Indian in Q1; Coloured/Asian/Indian in Q2-4; Coloured/Asian/Indian in Q5; Coloured; Black African in Q1; Black African in Q2-4; Black African in Q5; White in Q2-4 and White in Q5. The remaining covariates - age, gender, education, marital status and degree of urbanization - are defined as above. Again, we estimate, for each health domain: *i)* a standard ordered probit model that does not allow for reporting heterogeneity and *ii)* a HOPIT model including all the covariates in the own health equation and in the cut-points, and so allowing and correcting for reporting heterogeneity according to all of them.

---

[11] As seen in Table 1, there are no White individuals in Q1 in our sample; it is also not possible to further disaggregate wealth quintiles for this population, given the small size of the respective subsample, 238.

Average estimated probabilities of reporting any difficulty (mild to extreme) by wealth category, for the Black African population obtained from ordered probit models show that wealthiest (Q5) Black African population often reports *worse* health than the poorest Black African population (Q1), keeping other variables constant. Differences between the affluent and middle category Q2-4) are much smaller and not always in the same direction (see supplemental material, Figure S1).

We now use the reporting behaviour Eq. 3 of the HOPIT model to formally test for reporting heterogeneity across these population groups, and are able to reject at a 5% significance level the null hypothesis that Black African Q1 respondents use the same cut-points as Black African Q5 respondents, for no less than 10 of the 16 domains (detailed results available in supplemental material, Table S5, column 1). The same is true when comparing Black African Q2-4 respondents to Black African Q1 respondents for 7 of the 16 domains (Table S5, column 2).

As in the previous section, but now for the Black African population only, we use the HOPIT model to estimate health by wealth gaps purged of reporting heterogeneity and compare them to the uncorrected gaps. Again, we measure these using the average marginal effects of wealth on the probability of having any difficulty in a given health domain. Vignette-corrected probabilities are calculated using the cut-points of reference individual (an unmarried Black African male; aged 62, the average age in the sample; in quintiles two to four; who did not complete primary school; and lives in a rural area) for all individuals. The direction and size of the biases is illustrated in Fig. 3 on the left (comparing Q1 to Q5) and on the right (comparing Q2-4 to Q5), and detailed results can be found in the supplemental material (Table S6, columns 1-8). Figure 3 shows that, across all health domains, the health by wealth gap becomes (much) larger (or even reverses from negative to positive) once we correct for reporting differences. For instance, on the left we see that the poor Black African were 0.3 percentage points *less* likely than the rich, Black Africans to report difficulty with memory before correction. After correction, they are 10 percentage points *more* likely to do so, quite a spectacular difference. In certain domains, such as depression, heterogeneity correction leads to the removal of the health disadvantage of rich versus poor. A similar pattern is observed on the right when we look at the middle wealth category (Q2-4), albeit with smaller shifts.

[Fig. 3 here]

Reporting bias in Health Inequalities by Wealth Among Whites

A comparison of the estimated probabilities of reporting any difficulty in each of the 16 health domains for the two wealth categories defined for the White population, prior to applying the vignettes, reveals a stark contract to the Black African group (results available in the supplementary material, Figure S2). Across most domains, the less wealthy White report worse health (and in some cases considerably so) than the wealthier Whites. This is a first indication of a smaller role for reporting heterogeneity (and so a smaller bias) in wealth-related health gaps among Whites, compared to Black African population.

Results of the formal test using the HOPIT model indeed show little evidence of reporting heterogeneity in the self-evaluation of health by White individuals in Q2-4, compared to those in Q5 (results in supplementary material, Table S5, column 3). In only five of the 16 domains can reporting homogeneity be rejected at 5%.

Accordingly, for this population group, the marginal effects of being in Q2-4, compared to Q5, on the probability that someone has any difficulty in any of these health domains, are not as affected by correction for reporting heterogeneity (Fig. 4, and detailed results in Table S6, columns 9-12) as they are for the Black African population. Both before and after correction, the less wealthy Whites report to be less healthy than their wealthier counterparts.


[Fig. 6 here]


One possible concern with the specification used here for assessing health (reporting) differences is that it does not permit a comparison between the results obtained for the Black African and White population groups as, the distribution of wealth within quintiles is very different for the two groups.[12] We have therefore also estimated HOPIT and ordered probit models with a different wealth-race specification that avoids this issue. Namely, a flexible specification based on a polynomial of the continuous wealth variable interacted with race.[13] This enables wealthy *vs* poor comparisons using the exact same wealth levels for both races. Changing the wealth-race

---

[12] In particular, average wealth in Q5 (Q2-4) is about 13% (48%) larger for Whites than for Black Africans. And average wealth of Black Africans (Whites) in Q5 is 136% (80%) larger than that of Black Africans (Whites) in Q2-4.
[13] The preferred specification amongst those based on the continuous wealth variable was one that included the following wealth-race variables: dummy variables for White, Coloured, Indian/Asian (reference category Black African), a second order income polynomial interacted with the Black African dummy and interactions of income with White and with other races. Even more flexible specifications (for example, with second order income polynomials also for White and other races and with third order polynomial for Black Africans) often faced identification problems or otherwise tended to be outperformed according to the Akaike information criterion.

specification in this way does not affect our conclusions and so we prefer that based on quintiles which enables the more direct interpretations made above.

Inequalities in Health by Race, Within Top Wealth Quintile

Reporting bias in the measurement of wealth-related health inequalities is evident from our first model specification. Our second specification suggests that this may be less relevant among the White than the Black African population, with the poorest and middle quintile categories likely to underreport health problems relative to the richest Black Africans. Among both populations, the poor have worse actual health outcomes. One question that remains is: how does the health of the historically disadvantaged Black African population compare to that of the White population, in an era of decreasing economic segregation between races? We address this question by comparing health (reporting) of equally wealthy (Q5) Black Africans and Whites, using the same models as in the previous sub-section.[14]

As in previous sections, Fig. 5 compares average predicted probabilities of reporting some difficulty, estimated using a standard – homogenous reporting - ordered probit model. Across all domains, Black African wealthy individuals report on average worse levels of health than the White wealthy individuals. This is similar to the findings from Ardington and Gasealahwe (2014) and Lau and Ataguba (2015), who find that the Black African population have worse levels of general SAH than the White population.

[Fig. 5 here]

The following results, however, suggest that this relationship between race and health amongst the rich might be biased. There is clear evidence of reporting differences, with reporting homogeneity significantly rejected in 11 of the 16 health domains (results available in the supplementary material, Table S5, column 4). Figure 6 shows the marginal effects of being Black African and rich vs White and rich on the probability of reporting (being in) poor health (detailed results in Table S6, columns 13-16), keeping other variables fixed. Prior to controlling for reporting heterogeneity, Black Africans are significantly more likely to report poor health than White. Once fixed cut-points are applied, those health gaps are almost eliminated. Although the

---

[14] The conclusions obtained with this specification are also robust to the alternative specification described in footnote 13 above. In other words, they are not driven by the fact that White population group in Q1 is on average richer than the Black African group in the same quintile.

White population still shows better levels of health across most domains, the differences become much smaller and statistically insignificant.


[Fig. 6 here]


**Conclusion and Discussion**

Examination of general health differences cannot simply rely on the measurement of biomarkers and will always have to rely to a considerable extent on asking respondents to rate their health perception and experience. Measurement error in the answers to such questions can lead to substantial bias if reporting tendencies are systematically associated with characteristics of interest, like wealth and race. We make use of anchoring vignettes and HOPIT modeling to test and correct for such systematic reporting biases by wealth and race in a representative sample of elderly South Africans. Our findings are as follows.

First, we find that, for most health domains (10 out of 16), the hypothesis of reporting homogeneity by wealth is rejected. Rich (Q1) South African elderly rate the same health state descriptions as worse than their poor counterparts (Q5). This leads to a severe underestimation of health gaps by wealth: observed poor-rich health disparities are small and largely insignificant for all domains, except two (depression and anxiety) for which they are even significantly in favour the poor. Secondly, after correcting for these tendencies, substantial disparities between rich (Q5) and poor (Q1) emerge, most favouring the rich and significant for half (8 of 16) of the health domains rated. Race is however very unequally distributed across wealth quintiles and may play some role in this. Therefore, third, and given the interrelatedness of race and wealth, we examine health disparities by wealth *within* race groups. A similar picture emerges: also within race groups, reporting is heterogeneous and health gaps by wealth are severely underestimated when using observed, uncorrected reports. In the Black African race group, health disadvantages among the poorest (Q1) and the less poor (Q2-4) compared to the richest (Q5) always become larger and often significant after correction. Corrections in the same direction are observed among White population group. These have however a much smaller impact in the measurement of wealth-related inequalities as: *i)* they are usually less substantial and *ii)* unlike for the Black African population, significant health advantages for the rich are observed even before correction.

Finally, we look at health disparities by race within wealth groups, which can only be done for the top quintile. Interestingly, a very different finding emerges here: before reporting correction, Whites report less difficulties than Black Africans in *every* health domain (and significantly so in

17

most of them). But reporting homogeneity is also significantly rejected for most (11 out of 16) health domains: White rich are much more likely to rate the same vignettes as better health than rich Black Africans. When correcting for these biases, health disparities become smaller and lose significance. In other words: at similarly high wealth levels, no substantial health differences by race remain.

Our findings have important implications. They suggest that inequalities in health by wealth or race can be severely under- or overestimated if reporting tendencies are not taken into account. Given the dramatic inequalities in income and wealth in South Africa, it would indeed be very surprising if health were not similarly unequally distributed. But clear health inequality favouring the wealthier nonetheless only emerges after the differential health reporting between richer and poorer is accounted for. Health disparities across some groups may on the other hand also be overestimated: after correction of reporting tendencies, the wealthiest Black Africans are not found to be in worse health than White counterparts.

The vignette methodology relies on the assumptions of response consistency – individuals use the same standards for rating both own health and that of the vignettes – and vignette equivalence – there are no systematic differences across the sub-groups of interest in the way in which the vignette descriptions are understood. Evidence on these assumptions is limited and shows mixed results (Murray et al. 2003, Kristensen and Johansson 2008, Bago d'Uva et al. 2011, Van Soest et al. 2011, Grol-Prokopczyk et al. 2011, 2015, Rice et al. 2012). Strictly, our results may be driven by a departure from one or the other assumption or by true reporting heterogeneity, or a mix of the three. We believe however, that it is plausible that they are mainly driven by the latter, and so that vignette corrections enable a more accurate measurement of health disparities. For instance, for a departure from response consistency to be responsible for our results with respect to wealth, it would have to be the case that the poor are much more lenient than the rich when rating the health problems of vignettes but not their own, which seems implausible. The poor may understand the vignette descriptions differently from the rich. However, for such departure of vignette equivalence alone to explain our results, this bias would have to be predominantly in the same direction across very different health domains and very different vignette descriptions within domains. While it is not clear whether anchoring vignettes are sufficient to remove all of the bias, the correction is indeed in the direction to be expected given the widely socioeconomic inequalities in "harder" health outcomes like survival and disability, where we find disparities between the poor and the non-poor. These disparities are evident when we observe the longevity of individuals from different socio-economic backgrounds: while 22% of South Africans in

income Q5 are aged 50 and older, only 9% of South Africans in income Q1 reach this age category.[15] The evidence of underreporting of health problems by the poor using vignettes is also in line with the apparent lower awareness of chronic conditions mentioned above.

Does this mean that all health surveys should include health vignettes from now on? This recommendation may be premature, as there is still enough room for improvement to the anchoring vignettes methodology and not in every situation will the reporting differences and their impact be as large and systematic as observed in this elderly South African population. But in populations with important socio-economic and cultural differences like this one, it does indeed seem wise to include the possibility to test for reporting tendencies. And they are less expensive to collect than biomarker data from blood tests or other which, by definition, can only measure very specific aspects of (physical) health. Definitely, further research on the usefulness of vignettes is required, both on testing the underlying identifying assumptions of response consistency and vignette equivalence, as on its optimal implementation (wording, timing, ranking, etc) (Grol-Prokopczyk et al. 2011, 2015; Hopkins and King 2010).

Our findings also have important policy implications. South Africa is committed to obtaining universal health coverage of its population by 2025 (NHI South Africa 2015). It is currently still far from achieving the goal of providing equal treatment to those in equal needs (Ataguba and McIntyre 2012) but one first step in realizing this goal would be to make the worse off more aware of their health needs. Our research confirms that, at least among the older population, the lower socioeconomic groups overestimate their health and therefore underestimate their health care needs. A formal test of (horizontal) inequity would therefore be less likely to pick up a pro-rich pattern. This underestimation was demonstrated for elderly use of health care in several European countries (Bago d'Uva et al. 2008a) and is likely to hold *a fortiori* among the South African older population, given its far greater wealth and race disparities. One possible reason for why the poor in South Africa underestimate their health needs is their underperception of poor health and illness. If this is the mechanism that causes a systematic underestimation of ill health by the poor, then providing access to higher quality services at lower cost may work to decrease the reporting bias, even though the European evidence suggests it will not be entirely eliminated.

---

[15] These statistics were created using the NIDS wave 1 of 2008, collected by the Southern Africa Labour and Development Research Unit.

# References

Ardington, C., & Gasealahwe, B. (2014). Mortality in South Africa: Socio-economic profile and association with self-reported health. *Development Southern Africa*, *31*(1), 127–145.

Ataguba, J. E., Akazili, J., & McIntyre, D. (2011). Socioeconomic-related health inequality in South Africa: evidence from General Household Surveys. *International Journal for Equity in Health*, *10*(1), 48. doi:10.1186/1475-9276-10-48

Ataguba, J. E., & McIntyre, D. (2012). Paying for and receiving benefits from health services in South Africa: is the health system equitable? *Health policy and planning*, *27*(suppl 1), i35-i45.

Ataguba, J. E. (2013). Inequalities in multimorbidity in South Africa. *International journal for equity in health*, *12*(1), 64. doi:10.1186/1475-9276-12-64

Bago d'Uva, T., O'Donnell, O., & Van Doorslaer, E. (2008a). Differential health reporting by education level and its impact on the measurement of health inequalities among older Europeans. *International Journal of Epidemiology*, *37*(6), 1375–1383. doi:10.1093/ije/dyn146

Bago d'Uva, T., Van Doorslaer, E., Lindeboom, M., & O'Donnell, O. (2008b). Does Reporting Heterogeneity Bias the Measurement of Health Disparities? *Health Economics*, *17*, 351–375. doi:10.1002/hec

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

Baker, M., Stabile, M., & Deri, C. (2004). What Do Self-Reported, Objective, Measures of Health Measure? *The Journal of Human Resources*, *39*(4), 1067–1093. doi:10.3368/jhr.XXXIX.4.1067

Bonfrer, I., Van de Poel, E., Grimm, M., & Van Doorslaer, E. (2014). Does the distribution of health care utilization match needs in Africa? *Health policy and planning*, *29*(7), 921–37. doi:10.1093/heapol/czt074

Boyce, G., & Harris, G. (2011). A closer look at racial differences in the reporting of self-assessed health status and related concepts in South Africa. *Health SA Gesondheid*, *16*(1).

Charasse-Pouélé, C., & Fournier, M. (2006). Health disparities between racial groups in South Africa: a decomposition analysis. *Social science & medicine*, *62*(11), 2897–914. doi:10.1016/j.socscimed.2005.11.020

Coovadia, H., Jewkes, R., Barron, P., Sanders, D., & McIntyre, D. (2009). The health and health system of South Africa: historical roots of current public health challenges. *The Lancet*, *374*(9692), 817–834. doi:10.1016/S0140-6736(09)60951-X

Datta Gupta, N., Kristensen, N., & Pozzoli, D. (2010). External validation of the use of vignettes in cross-country health studies. *Economic Modelling*, *27*(4), 854–865.

De Salvo, K. B., Bloser, N., Reynolds, K., He, J., & Muntner., P. (2006). Mortality prediction with a single general self-rated health question. *Journal of general internal medicine*, *21*(3), 267–275. doi:10.1111/j.1525-1497.2005.0291.x

Dowd, J. B., & Todd, M. (2011). Does self-reported health bias the measurement of health inequalities in US adults? Evidence using anchoring vignettes from the Health and Retirement Study. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 66*(4), 478–489.

Etilé, F., & Milcent, C. (2006). Income-related reporting heterogeneity in self-assessed health: evidence from France. *Health economics*, *15*, 965–981.

Frankenberg, E., & Jones, N. R. (2004). Self-Rated Health and Mortality: does the Relationship Extend to a Low Income Setting? *Journal of health and social behaviour*, *45*(4), 441–452. doi:10.1177/002214650404500406

Grol-Prokopczyk, H., Freese, J., & Hauser, R. M. (2011). Using anchoring vignettes to assess group differences in general self-rated health. *Journal of health and social behaviour*, *52*(2), 246–261. doi:10.1177/0022146510396713

Grol-Prokopczyk, H., Verdes-Tennant, E., McEniry, M., & Ispany, M. (2015). Promises and Pitfalls of Anchoring Vignettes in Health Survey Research. *Demography*, *52*(5), 1703–1728. doi:10.1007/s13524-015-0422-1

Guindon, G. E., & Boyle, M. H. (2012). Using anchoring vignettes to assess the comparability of self-rated feelings of sadness, lowness or depression in France and Vietnam. *International Journal of Methods in Psychiatric Research*, *21*(1), 29–40. doi:10.1002/mpr

Harris, B., Goudge, J., Ataguba, J. E., McIntyre, D., Nxumalo, N., Jikwana, S., & Chersich, M. (2011). Inequities in access to health care in South Africa. *Journal of public health policy*, *32 Suppl 1*(S1), S102–S123. doi:10.1057/jphp.2011.35

He, W., Muenchrath, M. N., & Kowal, P. R. (2012). *Shades of Gray : A Cross-Country Study of Health and Well-Being of the Older Populations in SAGE Countries , 2007 – 2010 International Population Reports.* Washington DC.

Hernández-Quevedo, C., Jones, A. M., & Rice, N. (2004). Reporting bias and heterogeneity in self-assessed health. Evidence from the British Household Panel Survey. *Health, Econometrics and Data Group (HEDG) Working paper 05*, 4.

Hopkins, D. J., & King, G. (2010). Improving anchoring vignettes designing surveys to correct interpersonal incomparability. *Public Opinion Quarterly*, *74*(2), 201–222. doi:10.1093/poq/nfq011

Humphries, K. H., & Van Doorslaer, E. (2000). Income-related health inequality in Canada. *Social Science and Medicine*, *50*(5), 663–671. doi:10.1016/S0277-9536(99)00319-6

Idler, E. L. ., & Benyamini, Y. (1997). Self-Rated Health and Mortality : A Review of Twenty-

Seven Community Studies. *Journal of Health and Social Behaviour*, *38*(1), 21–37.

Jürges, H. (2007). True health vs response styles: exploring cross-country differences in self-reported health. *Health economics*, *16*(2), 163–178.

Jylhä, M., Volpato, S., & Guralnik, J. M. (2006). Self-rated health showed a graded association with frequently used biomarkers in a large population sample. *Journal of Clinical Epidemiology*, *59*(5), 465–471. doi:10.1016/j.jclinepi.2005.12.004

Kerkhofs, M., & Lindeboom, M. (1995). Subjective Health Measures and State Dependent Reporting Errors. *Health Economics*, 4(3), 221-235.

King, G., Murray, C. J. L., Salomon, J. A, & Tandon, A. (2004). Enchancing the Validty and Cross-Cultural Comparability of Measurement in Survey Research. *American Political Science Review*, *98*(1), 191–207. doi:10.1017/S000305540400108X

Kristensen, N. & Johansson, E. (2008). New evidence on cross-country differences in job satisfaction using anchoring vignettes. *Labour Economics,15*(1), pp.96-117.

Lau, Y. K., & Ataguba, J. E. (2015). Investigating the relationship between self-rated health and social capital in South Africa: a multilevel panel data analysis. *BMC public health*, *15*, 266. doi:10.1186/s12889-015-1601-0

Leibbrandt, M., Wegner, E. & Finn, A. 2011. The Policies for Reducing Income Inequality and Poverty in

South Africa. Southern Africa Labour and Development Research Unit. Working Paper Series Number 64.

Lindeboom, M., & Kerkhofs, M. (2009). Health and work of the elderly: Subjective health measures, reporting errors and endogeneity in the relationship between health and work. *Journal of Applied Econometrics, 24*(6), 1024–1046. doi:10.1002/jae

Lindeboom, M., & Van Doorslaer, E. (2004). Cut-point shift and index shift in self-reported health. *Journal of Health Economics*, *23*(6), 1083–1099. doi:10.1016/j.jhealeco.2004.01.002

Molina, T. (2016). Reporting Heterogeneity and Health Disparities Across Gender and Education Levels: Evidence From Four Countries. *Demography*, *53*(2), 295–323. doi:10.1007/s13524-016-0456-z

Murray, C. J., Ozaltin, E., Tandon, A., Salomon, J., Sadana, R., & Chatterji, S. (2003). Empirical evaluation of the anchoring vignettes approach in health surveys. *Health systems performance assessment: Debates, methods and empiricism*, *369*, 399.

NHI South Africa. (2015). *NHI South Africa.*
 http://www.mm3admin.co.za/documents/docmanager/f447b607-3c8f-4eb7-8da4-11bca747079f/00104782.pdf. Accessed 22 June 2016

Özler, B. (2007). Not separate, not equal: Poverty and inequality in post-Apartheid South Africa. *Economic development and cultural change*, *55*(739), 1–41. doi:10.1086/511191

Peracchi, F., & Rossetti, C. (2012). Heterogeneity in health responses and anchoring vignettes. *Empirical Economics*, *42*(2), 513–538. doi:10.1007/s00181-011-0530-8

Rice, N., Robone, S., & Smith, P. C. (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. doi:10.1111/j.1467-985X.2011.01021.x

Salomon, J. A., Mathers, C. D., Chatterji, S., Sadana, R., Ustun, T. B., & Murray, C. J. (2003). *Health Systems Performance Assessment_ Debates, Methods and Empiricism - Christopher J.* (C. Murray & D. Evans, Eds.). Geneva: World Health Organization.

Sen, A. (2002). Health: perception versus observation. *British Medical Journal*, *324*(August 2006), 860–861. doi:10.1136/bmj.324.7342.860

Tandon, A., Murray, C., Salomon, J., & King, G. (2003). Statistical models for enhancing cross-population comparability. *Health systems performance assessment: debates, methods and empiricism*, (42), 1–37.

Van Doorslaer, E., & Gerdtham, U. G. (2003). Does inequality in self-assessed health predict inequality in survival by income? Evidence from Swedish data. *Social Science and Medicine*, *57*(9), 1621–1629. doi:10.1016/S0277-9536(02)00559-2

Van Doorslaer, E., Koolman, X., & Jones, A. M. (2004). Explaining income-related inequalities in doctor utilisation in Europe. *Health Economics*, *13*(7), 629–647. doi:10.1002/hec.919

Van Soest, A., Delaney, L., Harmon, C., Kapteyn, A. & Smith, J.P. (2011). Validating the use of anchoring vignettes for the correction of response scale differences in subjective questions. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 174*(3), pp.575-595.

World Health Organization. (2008). WHO Study on global AGEing and adult health (SAGE). http://www.who.int/healthinfo/sage/en/. Accessed 1 January 2013

Zere, E., & McIntyre, D. (2003). Inequities in under-five child malnutrition in South Africa. *International journal for equity in health*, *2*(1), 7. doi:10.1186/1475-9276-2-7

Zhang, H., Bago d'Uva, T.M., & Van Doorslaer, E. (2015). The gender health gap in China: a decomposition analysis. *Economics and Human Biology 18,* 13–26.

**Table 1:** Sample averages of covariates by wealth quintile

| | | Quintile | | | | |
|---|---|---|---|---|---|---|
| | | **1** | **2** | **3** | **4** | **5** |
| **Female** | | 0.60 | 0.58 | 0.70 | 0.58 | 0.53 |
| **Education** | No formal schooling | 0.41 | 0.33 | 0.21 | 0.16 | 0.06 |
| | Less than primary school | 0.25 | 0.28 | 0.28 | 0.20 | 0.13 |
| | Primary school completed | 0.25 | 0.21 | 0.24 | 0.26 | 0.22 |
| | Secondary School completed | 0.07 | 0.12 | 0.20 | 0.21 | 0.17 |
| | High school completed | 0.02 | 0.04 | 0.04 | 0.07 | 0.25 |
| | College or university completed | 0.00 | 0.02 | 0.03 | 0.10 | 0.17 |
| **Age in years** | | 61.48 | 60.38 | 60.66 | 61.86 | 62.39 |
| **Marital status** | Never married | 0.24 | 0.16 | 0.15 | 0.14 | 0.08 |
| | Married | 0.31 | 0.41 | 0.39 | 0.53 | 0.69 |
| | Cohabitating | 0.11 | 0.10 | 0.04 | 0.02 | 0.02 |
| | Separated/divorced | 0.08 | 0.08 | 0.09 | 0.04 | 0.03 |
| | Widowed | 0.26 | 0.25 | 0.32 | 0.27 | 0.19 |
| **Urban** | | 0.44 | 0.59 | 0.61 | 0.76 | 0.87 |
| **Race** | Black African | 0.94 | 0.88 | 0.82 | 0.65 | 0.42 |
| | White | 0.00 | 0.01 | 0.02 | 0.12 | 0.30 |
| | Coloured | 0.05 | 0.10 | 0.13 | 0.18 | 0.20 |
| | Asian/Indian | 0.01 | 0.01 | 0.03 | 0.04 | 0.09 |

**N = 2968**

Notes: Statistics weighted with post-stratified individual weights

**Table 2:** Tests of reporting homogeneity between wealth Q1 and wealth Q5, by health domain (p-values)

| Health domain | | Health domain | |
|---|---|---|---|
| Moving around | **0.032** | Energy | 0.448 |
| Vigorous activity | 0.345 | Sleep | 0.650 |
| Depressed | **0.037** | Far sighted | **0.003** |
| Anxiety | 0.334 | Near sighted | **<0.001** |
| Relationships | **0.003** | Care | **0.002** |
| Conflict | **0.022** | Appearance | **<0.001** |
| Body pain | **0.016** | Memory | 0.184 |
| Body discomfort | 0.085 | Learning | 0.788 |

Note: p-values in bold text if they are significant at a 5% significance level; tests of joint equality of respective coefficients in the cut-points of HOPIT models by health domain; HOPIT model includes controls described in Table 1.

**Figure 1:** Estimated probability of reporting any difficulty (mild to extreme) before correcting for reporting bias.
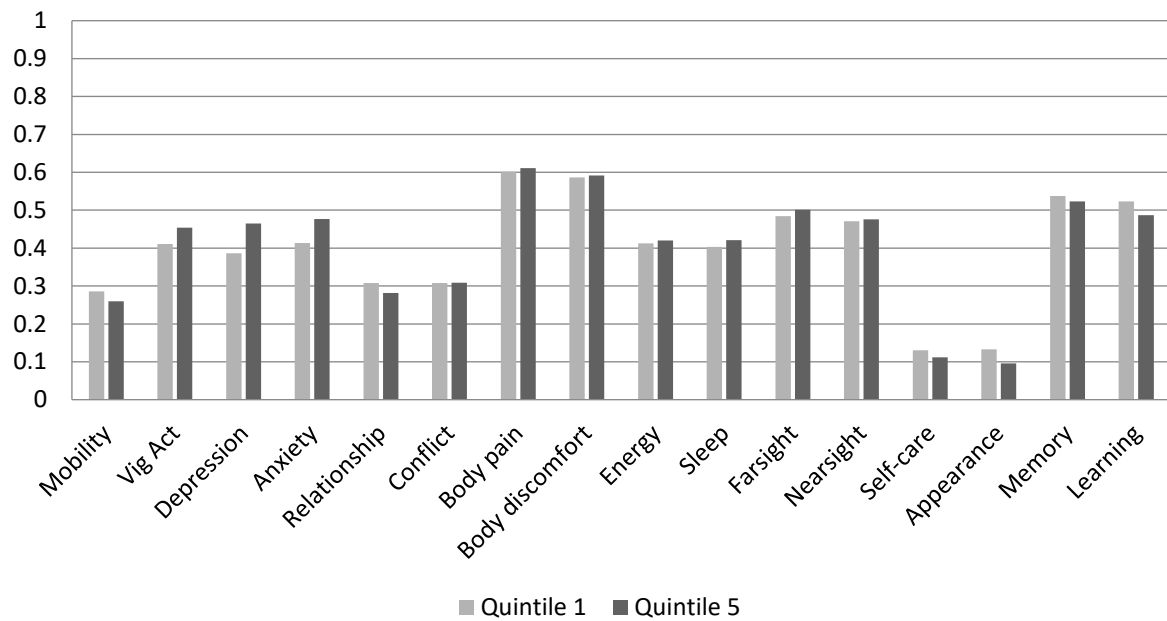


Notes: Average probabilities estimated from ordered probit models, varying wealth quintile, keeping fixed the other covariates (described in Table 1).

*Source:* WHO SAGE survey, 2008

**Figure 2:** Average marginal effects of being in Q1 on reporting any difficulty in health relative, compared to Q5.
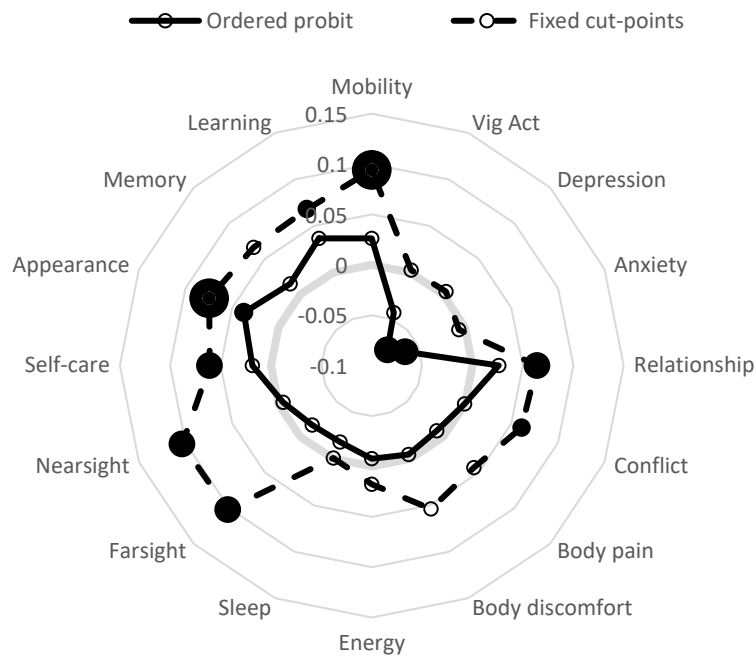


Notes: Average marginal effects from ordered probit models and HOPIT models including the covariates described in Table 1. For the HOPIT model, marginal effects use fixed cut-points of reference individual (unmarried Black African male; in wealth quintile 1; aged 62; who did not complete primary school; and lives in a rural area).

● Significant at 1%   ● Significant at 5%   ● Significant at 10% ○ Insignificant at 10%

*Source:* WHO SAGE survey, 2008

**Figure 3:** Average marginal effects of being in Q1 (left) and Q2-4 (right) on the probability of reporting any difficulty, relative to being in Q5 - Black African population.
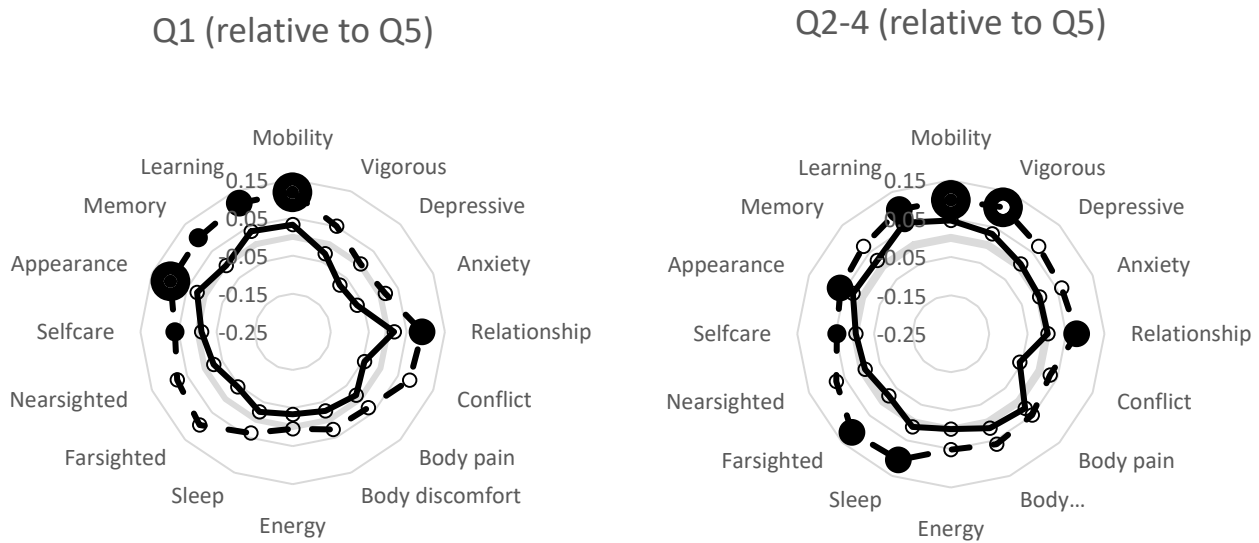


Notes: Average marginal effects defined as in Figure 3. Ordered probit and HOPIT models control for the same variables as in Table 1, except that race and wealth are controlled for in the following way: Black Q1, Black Q2-4; Black Q5; White Q2-4; White Q5; other races Q1, other races Q2-4; Other races Q5; and one dummy variable for Coloured.

● Significant at 1%  ● Significant at 5%  ● Significant at 10% ○ Insignificant at 10%

*Source:* WHO SAGE survey, 2008

**Figure 4:** Average marginal effects of being in wealth Q2-4 on the probability of reporting any difficulty in health domain (mild to extreme), relative to being Q5: White population.
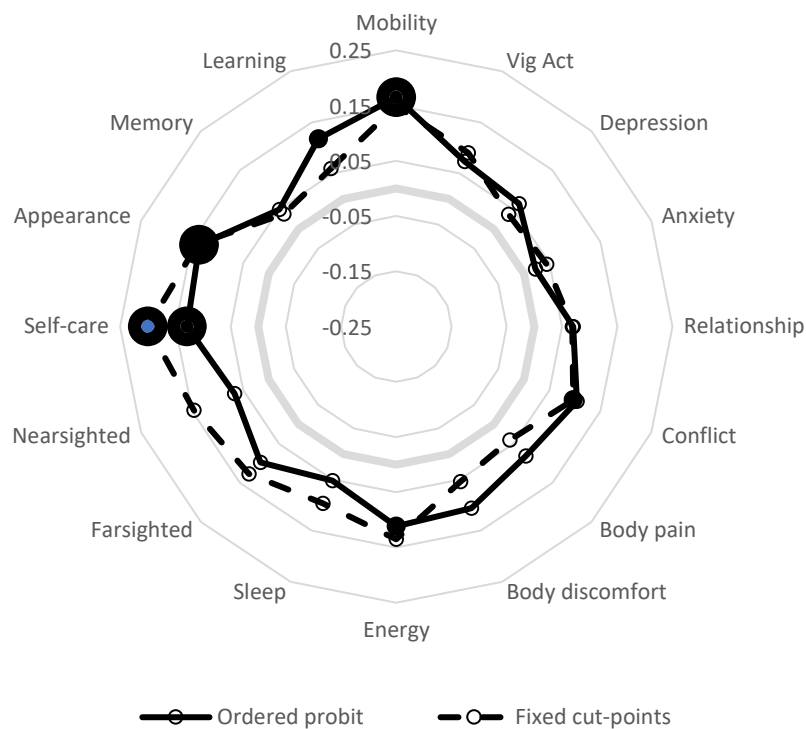


Notes: Average marginal effects computed varying wealth quintile within White population group, keeping other variables constant. Ordered probit and HOPIT models specified as in Figure 4.

⬤ Significant at 1%   ● Significant at 5%   • Significant at 10%   ○ Insignificant at 10%

*Source:* WHO SAGE survey, 2008

**Figure 5:** Estimated probability of reporting any difficulty (mild to extreme) before correcting for reporting bias for Black Africans and Whites: Q5.



Notes: Average probabilities estimated from ordered probit and HOPIT models, varying race within Q5, keeping other variables constant. Ordered probit and HOPIT models specified as in Figure 4.

*Source:* WHO SAGE survey, 2008

**Figure 6:** Average marginal effects of being White on reporting any difficulty in health (mild to extreme), relative to being Black African: Q5.
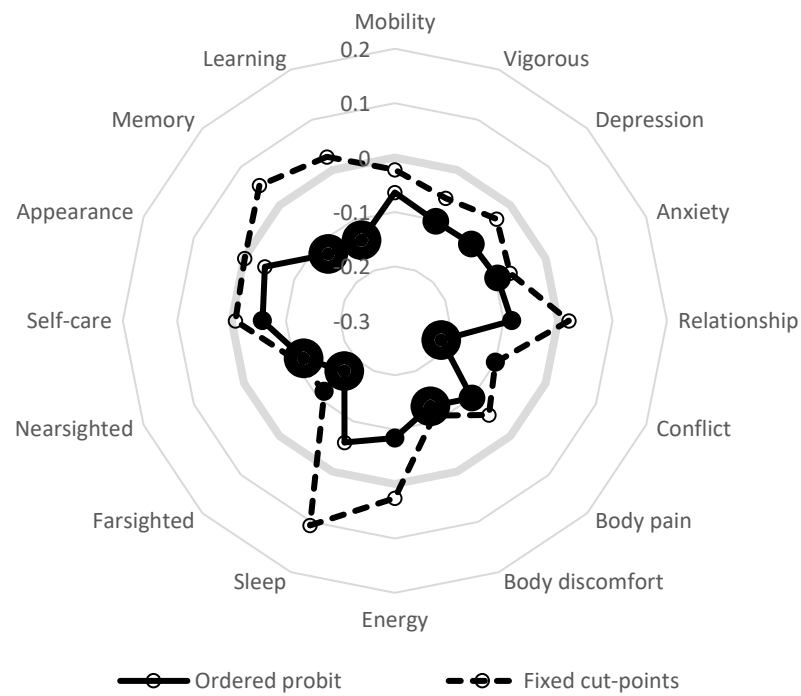


Notes: Average marginal effects computed varying race within Q5, keeping other variables constant. Ordered probit and HOPIT models specified as in Figure 4.
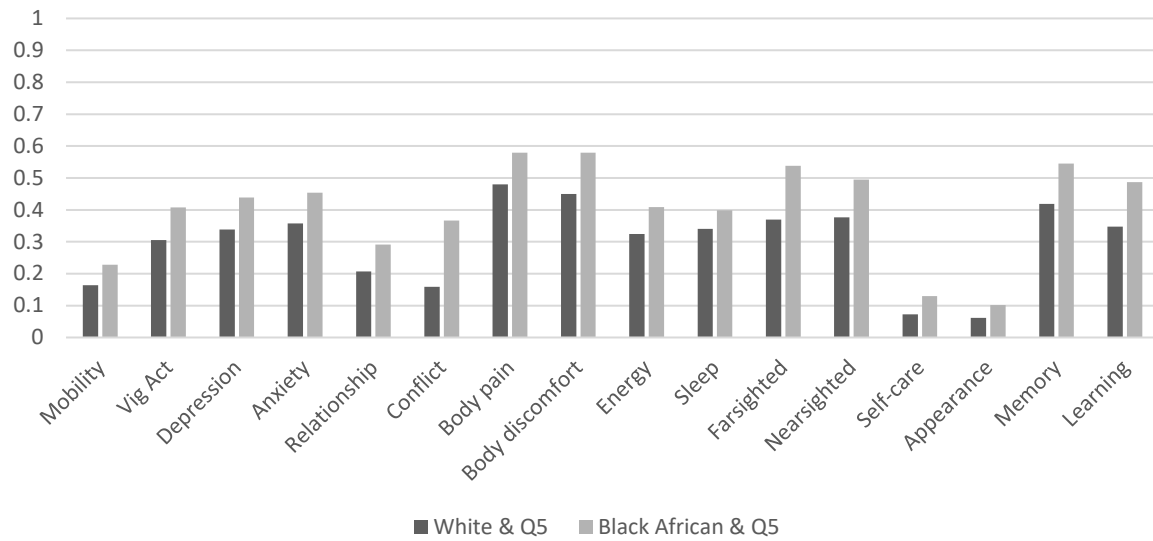
⬤ Significant at 1% ● Significant at 5% • Significant at 10% ○ Insignificant at 10%

*Source:* WHO SAGE survey, 2008

## Supplemental material

**Table S1**: Summary of health domains included in WHO SAGE data

| Domain | |
|---|---|
| **Mobility** | Overall in the last 30 days, how much difficulty did you have with moving around |
| **Vigorous activities** | Overall in the last 30 days, how much difficulty did you have in vigorous activities ('vigorous activities' require hard physical effort and cause large increases in breathing or health rate)? |
| **Depression** | Overall in the last 30 days, how much of a problem did you have with feeling sad, low or depressed? |
| **Anxiety** | Overall in the last 30 days, how much of a problem did you have with worry or anxiety? |
| **Relationships** | Overall in the last 30 days, how much difficulty did you have with personal relationships or participation in the community? |
| **Conflict** | Overall in the last 30 days, how much difficulty did you have in dealing with conflicts and tensions with others? |
| **Body pains** | Overall in the last 30 days, how much of bodily aches or pains did you have? |
| **Body discomfort** | Overall in the last 30 days, how much bodily discomfort did you have? |
| **Energy** | Overall in the last 30 days, how much of a problem did you have due to not feeling rested and refreshed during the day (for example, feeling tired, not having energy)? |
| **Sleep** | Overall in the last 30 days, how much of a problem did you have with sleeping, such as falling asleep, waking up frequently during the or waking up too early in the morning? waking up frequently during the night or waking up too early in the morning? |
| **Far sighted** | In the last 30 days, how much difficulty did you have in seeing and recognising an object or a person you know across the road (from a distance of about 20 meters)? |
| **Near sighted** | In the last 30 days, how much difficulty did you have in seeing and recognising an object at arm's length (for example, reading)? |
| **Self-care** | Overall in the last 30 days, how much difficulty did you have with self-care, such as bathing/washing or dressing yourself? |
| **Appearance** | Overall in the last 30 days, how much difficulty did you have in taking care of and maintaining your general appearance (for example, grooming, looking neat and tidy)? |
| **Memory** | Overall in the last 30 days, how much difficulty did you have with concentrating or remembering things? |
| **Learning** | Overall in the last 30 days, how much difficulty did you have in learning a new task. |

*Source*: WHO SAGE individual questionnaire

**Table S2**: Estimation results of ordered probit HOPIT models, mobility domain

| Variable | Ordered probit | HOPIT | | | | | | | | | |
| | Latent own health | Latent own health | | Latent vignette health | | | | Cut-points | | | |
| | | | Sigma | Vig2 | Vig3 | Vig4 | Vig5 | 1 - None vs. Mild | 2. Mild vs. Moderate | 3 -Moderate vs. Severe | 4 - Severe vs. Extreme |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Quintile 1 | 0.0828 | 0.368*** | | | | | | 0.277*** | 0.267*** | 0.278*** | 0.091 |
| | (0.091) | (0.137) | | | | | | (0.102) | (0.099) | (0.097) | (0.113) |
| Quintile 2 | 0.120 | 0.315** | | | | | | 0.172* | 0.178** | 0.204** | 0.141 |
| | (0.086) | (0.127) | | | | | | (0.093) | (0.090) | (0.088) | (0.106) |
| Quintile 3 | 0.136* | 0.269** | | | | | | 0.101 | 0.151* | 0.127 | -0.115 |
| | (0.083) | (0.124) | | | | | | (0.093) | (0.089) | (0.087) | (0.102) |
| Quintile 4 | 0.196** | 0.333*** | | | | | | 0.109 | 0.126 | 0.107 | -0.100 |
| | (0.078) | (0.116) | | | | | | (0.085) | (0.082) | (0.081) | (0.094) |
| Quintile 5 | (Ref.) | (Ref.) | | | | | | (Ref.) | (Ref.) | (Ref.) | (Ref.) |
| | | | | | | | | | | | |
| Female | 0.0277 | -0.00275 | | | | | | -0.0448 | 0.00818 | -0.0625 | -0.0560 |
| | (0.053) | (0.079) | | | | | | (0.058) | (0.056) | (0.055) | (0.066) |
| No formal schooling | (Ref.) | (Ref.) | | | | | | (Ref.) | (Ref.) | (Ref.) | (Ref.) |
| Less than primary school | 0.0907 | 0.0352 | | | | | | -0.041 | -0.135* | -0.096 | 0.040 |
| | (0.067) | (0.102) | | | | | | (0.077) | (0.074) | (0.073) | (0.088) |
| Primary school completed | 0.0741 | -0.0464 | | | | | | -0.121 | -0.151** | -0.167** | -0.111 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | (0.068) | (0.103) | | (0.077) | (0.075) | (0.074) | (0.087) |
| Secondary School completed | -0.0797 | -0.153 | | -0.066 | -0.092 | 0.036 | 0.130 |
| | (0.083) | (0.123) | | (0.090) | (0.087) | (0.085) | (0.103) |
| High school completed | -0.104 | -0.298* | | -0.208 | -0.128 | -0.115 | 0.120 |
| | (0.116) | (0.173) | | (0.127) | (0.122) | (0.122) | (0.146) |
| College or university completed | -0.308** | -0.512** | | -0.151 | -0.295** | -0.099 | 0.176 |
| | (0.140) | (0.206) | | (0.146) | (0.140) | (0.139) | (0.168) |
| Age in years | 0.021*** | 0.021*** | | -0.003 | -0.003 | -0.001 | -0.004 |
| | (0.003) | (0.004) | | (0.003) | (0.003) | (0.003) | (0.003) |
| Never married | (Ref.) | (Ref.) | | (Ref.) | (Ref.) | (Ref.) | (Ref.) |
| Married | 0.004 | -0.009 | | -0.001 | -0.033 | -0.044 | -0.069 |
| | (0.078) | (0.115) | | (0.084) | (0.080) | (0.079) | (0.094) |
| Cohabitating | -0.051 | -0.315* | | -0.253* | -0.178 | -0.439*** | -0.058 |
| | (0.126) | (0.191) | | (0.141) | (0.137) | (0.136) | (0.152) |
| Separated/divorced | 0.257** | 0.291* | | 0.024 | -0.038 | -0.075 | 0.205 |
| | (0.110) | (0.160) | | (0.114) | (0.109) | (0.108) | (0.132) |
| Widowed | 0.187** | 0.145 | | -0.037 | -0.112 | -0.144* | 0.026 |
| | (0.080) | (0.119) | | (0.087) | (0.083) | (0.082) | (0.097) |
| Urban | -0.034 | 0.057 | | 0.082 | 0.106* | 0.145** | 0.187*** |
| | (0.055) | (0.083) | | (0.062) | (0.060) | (0.059) | (0.069) |

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Black African | (Ref.) | (Ref.) | | | | | | (Ref.) | (Ref.) | (Ref.) | (Ref.) |
| White | -0.110 | -0.0735 | | | | | | 0.058 | 0.0829 | -0.0942 | -0.306** |
| | (0.110) | (0.160) | | | | | | (0.114) | (0.109) | (0.108) | (0.121) |
| Coloured | -0.182*** | -0.096 | | | | | | 0.121* | 0.0650 | 0.086 | 0.237*** |
| | (0.064) | (0.096) | | | | | | (0.070) | (0.067) | (0.067) | (0.081) |
| Asian/Indian | 0.420*** | 0.452*** | | | | | | -0.017 | -0.101 | 0.065 | 0.076 |
| | (0.084) | (0.127) | | | | | | (0.096) | (0.092) | (0.090) | (0.109) |
| Constant | 3.283*** | -2.354*** | 1.127*** | 1.586*** | 1.193*** | -1.164*** | 2.898*** | -0.028 | 0.435* | 1.157*** | 2.500*** |
| | (0.221) | (0.325) | (0.044) | (0.060) | (0.058) | (0.069) | (0.071) | (0.234) | (0.225) | (0.221) | (0.261) |
| Log-likelihood | -2659.29 | -6605.67 | | | | | | | | | |
| Observations | 2,968 | 2,968 | | | | | | | | | |

Notes: Standard errors in parentheses; *** p-value<0.01, **p-value<0.05, *p-value<0.1.

*Source*: WHO SAGE survey, 2008

**Table S3**: Estimated coefficients of quintile 1 (relative to quintile 5) in the cut-points of HOPIT models by health domain

| | Cut-point 1: None vs. Mild | Cut-point 2: Mild vs. Moderate | Cut-point 3: Moderate vs. Severe | Cut-point 4: Severe vs. Extreme |
|---|---|---|---|---|
| Mobility | 0.277*** | 0.267*** | 0.278*** | 0.0909 |
| | (0.102) | (0.099) | (0.097) | (0.113) |
| Vig Act | 0.190* | 0.169* | 0.137 | 0.145 |
| | (0.107) | (0.100) | (0.094) | (0.010) |
| Depression | 0.288*** | 0.121 | 0.133 | 0.144 |
| | (0.106) | (0.098) | (0.091) | (0.109) |
| Anxiety | 0.185* | 0.0916 | 0.146 | 0.0809 |
| | (0.109) | (0.099) | (0.093) | (0.108) |
| Relationship | 0.201* | 0.111 | 0.317*** | 0.0349 |
| | (0.105) | (0.101) | (0.094) | (0.106) |
| Conflict | 0.224** | 0.259** | 0.328*** | 0.129 |
| | (0.109) | (0.105) | (0.099) | (0.106) |
| Body pain | 0.265** | 0.0575 | 0.156* | 0.247* |
| | (0.103) | (0.092) | (0.093) | (0.127) |
| Body discomfort | 0.236** | 0.144 | 0.153* | 0.239* |
| | (0.103) | (0.092) | (0.093) | (0.126) |
| Energy | 0.0566 | 0.0818 | 0.155* | 0.105 |
| | (0.101) | (0.092) | (0.083) | (0.114) |
| Sleep | 0.0187 | 0.0551 | 0.119 | 0.0490 |
| | (0.090) | (0.085) | (0.083) | (0.132) |
| Far sight | 0.329*** | 0.196** | 0.171* | 0.195* |
| | (0.092) | (0.089) | (0.088) | (0.104) |
| Near-sight | 0.348*** | 0.114 | 0.175** | 0.228** |
| | (0.094) | (0.089) | (0.087) | (0.101) |
| Self-care | 0.392*** | 0.316*** | 0.0726 | -0.146 |
| | (0.109) | (0.102) | (0.106) | (0.120) |
| Appearance | 0.420*** | 0.319*** | -0.0234 | -0.163 |
| | (0.110) | (0.103) | (0.109) | (0.119) |
| Memory | 0.203** | 0.110 | -0.0117 | 0.0793 |
| | (0.097) | (0.092) | (0.094) | (0.133) |
| Learning | 0.122 | 0.0830 | 0.0673 | 0.0724 |
| | (0.098) | (0.094) | (0.094) | (0.125) |

Notes: Standard errors in parentheses; *** p-value<0.01, ** p-value<0.05, * p-value<0.1. HOPIT models specified as in Table S3.

*Source*: WHO SAGE survey, 2008

**Table S4**: Marginal effects of being in quintile 1 (poorest) on reporting any health difficulty (mild to extreme) in a health domain, compared to being in quintile 5, assuming reporting homogeneity and allowing for reporting heterogeneity

| | Ordered probit | | HOPIT | |
|---|---|---|---|---|
| Mobility | 0.026 | (0.029) | 0.094*** | (0.034) |
| Vig Act | -0.043 | (0.031) | 0.002 | (0.036) |
| Depression | -0.078** | (0.031) | 0.004 | (0.042) |
| Anxiety | -0.064** | (0.031) | -0.006 | (0.043) |
| Relationship | 0.026 | (0.03) | 0.064** | (0.031) |
| Conflict | <-0.001 | (0.03) | 0.061* | (0.033) |
| Body pain | -0.009 | (0.029) | 0.043 | (0.039) |
| Body discomfort | -0.004 | (0.029) | 0.054 | (0.039) |
| Energy | -0.008 | (0.032) | 0.018 | (0.042) |
| Sleep | -0.018 | (0.031) | -0.001 | (0.043) |
| Farsight | -0.016 | (0.031) | 0.102** | (0.047) |
| Nearsight | -0.005 | (0.031) | 0.104** | (0.046) |
| Self-care | 0.019 | (0.022) | 0.061** | (0.024) |
| Appearance | 0.037* | (0.021) | 0.075*** | (0.023) |
| Memory | 0.015 | (0.03) | 0.066 | (0.044) |
| Learning | 0.037 | (0.03) | 0.068* | (0.04) |

Notes: Standard errors in parentheses; *** p-value<0.01, ** p-value<0.05, * p-value<0.1. Ordered probit and HOPIT models specified as in Table S3.

*Source*: WHO SAGE survey, 2008

**Table S5**: Tests of reporting homogeneity across wealth and race groups, by health domain (p-values)

| | Black and Q1 (*vs* Black and Q5) | Black and Q2-4 (*vs* Black and Q5) | White and Q 2-4 (*vs* White and Q5) | White and Q5 (*vs* Black and Q5) |
|---|---|---|---|---|
| Mobility | **0.003** | **0.016** | 0.173 | 0.450 |
| Vig Act | **0.014** | **0.019** | 0.484 | 0.835 |
| Depression | **0.004** | **0.085** | **<0.001** | 0.227 |
| Anxiety | **0.077** | **0.096** | **<0.001** | **0.030** |
| Relationship | **<0.001** | **0.004** | 0.634 | **0.026** |
| Conflict | **0.002** | 0.179 | 0.138 | **0.036** |
| Body pain | **0.057** | 0.732 | **0.002** | 0.529 |
| Body discomfort | **0.094** | 0.231 | **0.026** | 0.483 |
| Energy | 0.799 | 0.179 | 0.569 | **<0.001** |
| Sleep | 0.617 | **0.038** | 0.804 | **<0.001** |
| Far-sight | **0.009** | **0.011** | **0.052** | **0.030** |
| Nearsight | **<0.001** | **0.011** | **0.013** | **<0.001** |
| Self-care | **<0.001** | **0.022** | 0.721 | **0.023** |
| Appearance | **<0.001** | 0.122 | 0.562 | **0.024** |
| Memory | **0.021** | 0.122 | 0.875 | **0.001** |
| Learning | 0.145 | 0.476 | 0.633 | **0.002** |

Notes: p-values in bold text if they are significant at a 10% significance level; tests of joint equality of respective coefficients in the cut-points of HOPIT models by health domain. Controls are the same as in Table 1, except that race and wealth are controlled for in the following way: Black Q1, Black Q2-4; Black Q5; White Q2-4; White Q5; other races Q1, other races Q2-4; Other races Q5; and one dummy variable for Coloured.

*Source:* WHO SAGE survey, 2008

**Table S6**: Marginal effects of race-wealth groups on the probability of reporting any health difficulty (mild to extreme) in a health domain

| | Black and Q1 (*vs* Black and Q5) | | | | Black Q 2-4 (*vs* Black and Q5) | | | | White Q 2-4 (*vs* White and Q5) | | | | White and Q5 (*vs* Black and Q5) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Ordered probit | | Fixed cut-points | | Ordered probit | | Fixed cut-points | | Ordered probit | | Fixed cut-points | | Ordered probit | | Fixed cut-points | |
| Mobility | 0.032 | (0.034) | 0.117*** | (0.043) | 0.045 | (0.03) | 0.099*** | (0.035) | 0.165*** | (0.059) | 0.149** | (0.072) | -0.064 | (0.039) | -0.023 | (0.045) |
| Vig Act | -0.028 | (0.037) | 0.051 | (0.044) | 0.031 | (0.034) | 0.106*** | (0.039) | 0.073 | (0.063) | 0.09 | (0.072) | -0.103** | (0.046) | -0.057 | (0.051) |
| Depression | -0.075* | (0.039) | 0.002 | (0.055) | 0.007 | (0.035) | 0.072 | (0.049) | 0.064 | (0.067) | 0.038 | (0.089) | -0.101** | (0.048) | -0.035 | (0.067) |
| Anxiety | -0.067* | (0.039) | 0.013 | (0.054) | 0.0001 | (0.035) | 0.063 | (0.048) | 0.022 | (0.067) | 0.045 | (0.088) | -0.096** | (0.048) | -0.071 | (0.066) |
| Relationship | 0.017 | (0.037) | 0.091** | (0.041) | 0.002 | (0.033) | 0.078** | (0.033) | 0.07 | (0.061) | 0.068 | (0.077) | -0.085* | (0.044) | 0.02 | (0.048) |
| Conflict | -0.046 | (0.039) | 0.082 | (0.051) | -0.056 | (0.034) | 0.03 | (0.042) | 0.104* | (0.059) | 0.097 | (0.078) | -0.207*** | (0.043) | -0.1* | (0.053) |
| Body pain | -0.015 | (0.037) | 0.031 | (0.047) | 0.024 | (0.033) | 0.049 | (0.042) | 0.082 | (0.064) | 0.041 | (0.085) | -0.099** | (0.047) | -0.055 | (0.062) |
| Body discomfort | -0.025 | (0.037) | 0.027 | (0.048) | 0.016 | (0.033) | 0.061 | (0.042) | 0.106 | (0.064) | 0.054 | (0.088) | -0.13*** | (0.047) | -0.111* | (0.063) |
| Energy | -0.034 | (0.039) | 0.005 | (0.051) | -0.002 | (0.035) | 0.052 | (0.046) | 0.112* | (0.066) | 0.135 | (0.09) | -0.084* | (0.047) | 0.026 | (0.064) |
| Sleep | -0.023 | (0.039) | 0.038 | (0.052) | 0.012 | (0.035) | 0.105** | (0.047) | 0.052 | (0.066) | 0.097 | (0.093) | -0.058 | (0.048) | 0.107 | (0.067) |
| Farsight | -0.046 | (0.039) | 0.096 | (0.06) | -0.021 | (0.035) | 0.114** | (0.053) | 0.098 | (0.069) | 0.127 | (0.099) | -0.169*** | (0.05) | -0.117 | (0.07) |
| Nearsight | -0.025 | (0.04) | 0.078 | (0.06) | -0.009 | (0.035) | 0.073 | (0.053) | 0.067 | (0.069) | 0.147 | (0.101) | -0.119** | (0.05) | -0.106* | (0.07) |
| Self-care | -0.01 | (0.027) | 0.059* | (0.033) | -0.002 | (0.024) | 0.047* | (0.027) | 0.129*** | (0.049) | 0.2*** | (0.079) | -0.057* | (0.029) | -0.007 | (0.037) |
| Appearance | 0.022 | (0.025) | 0.097*** | (0.034) | 0.026 | (0.022) | 0.062** | (0.026) | 0.136*** | (0.048) | 0.14** | (0.07) | -0.041 | (0.027) | -0.001 | (0.036) |
| Memory | -0.003 | (0.038) | 0.1* | (0.054) | 0.021 | (0.034) | 0.072 | (0.049) | 0.05 | (0.068) | 0.038 | (0.099) | -0.127*** | (0.049) | 0.052 | (0.069) |
| Learning | 0.036 | (0.038) | 0.116** | (0.051) | 0.063* | (0.034) | 0.1** | (0.046) | 0.118* | (0.066) | 0.06 | (0.095) | -0.14*** | (0.049) | 0.026 | (0.067) |

Notes: Standard errors in parentheses; *** p-value<0.01, ** p-value<0.05, * p-value<0.1. Average marginal effects. For the HOPIT model, marginal effects use fixed cut-points of reference individual (unmarried Black African male; in wealth quintile 1; aged 62; who did not complete primary school; and lives in a rural area). As in Table S6, control variables are the same as in Table 1, except that they race and wealth are controlled for in the following way: Black Q1, Black Q2-4; Black Q5; White Q2-4; White Q5; other races Q1, other races Q2-4; Other races Q5; and one dummy variable for Coloured.

*Source:* WHO SAGE survey, 2008

**Figure S1**: Estimated probability of reporting any difficulty (mild to extreme) before correcting for reporting bias: Black African population.
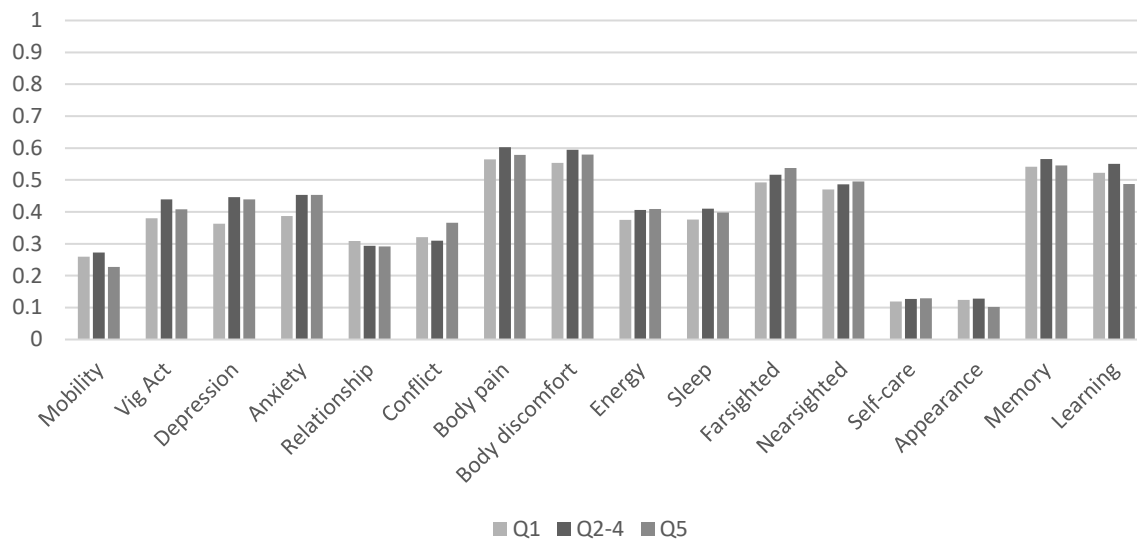


Notes: Average probabilities estimated from ordered probit models, varying wealth quintile within the Black population group, keeping fixed the other covariates (the same as in Table 1, except that race and wealth are controlled for in the following way: Black Q1, Black Q2-4; Black Q5; White Q2-4; White Q5; other races Q1, other races Q2-4; Other races Q5; and one dummy variable for Coloured).

*Source:* WHO SAGE survey, 2008

**Figure S2**: Estimated probability of reporting any difficulty (mild to extreme) before correcting for reporting bias: White population.
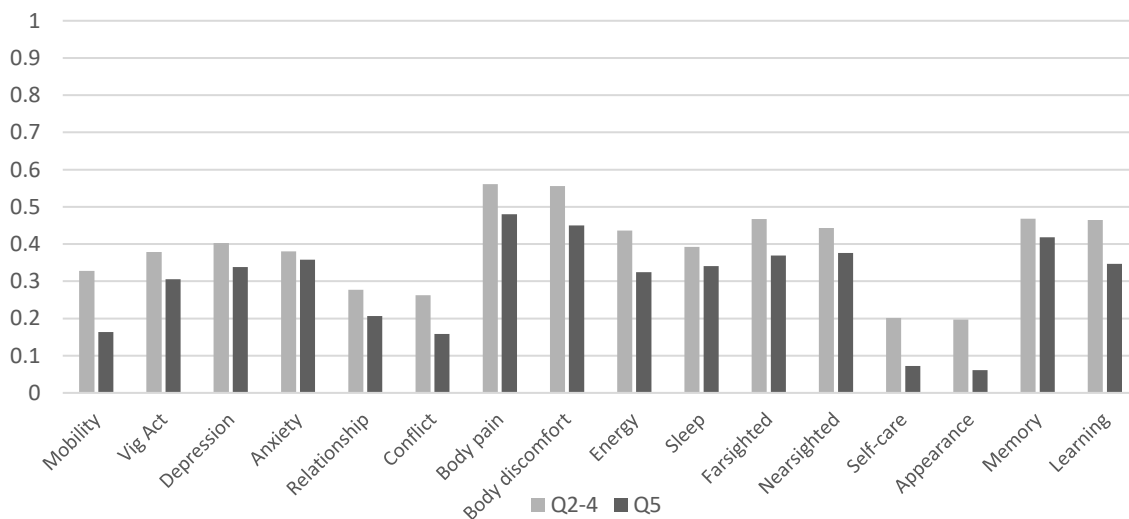


Notes: Average probabilities estimated from ordered probit models, varying wealth quintile within the White population group, keeping fixed the other covariates (the same as in Table 1, except that race and wealth are controlled for in the following way: Black Q1, Black Q2-4; Black Q5; White Q2-4; White Q5; other races Q1, other races Q2-4; Other races Q5; and one dummy variable for Coloured).

*Source:* WHO SAGE survey, 2008