Demography (2018) 55:1935–1956
https://doi.org/10.1007/s13524-018-0709-0

Poor Health Reporting? Using Anchoring Vignettes to Uncover Health Disparities by Wealth and Race

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

Published online: 25 September 2018

Abstract In spite of the wide disparities in wealth and in objective health measures like mortality, observed inequality by wealth in self-reported health appears to be nearly nonexistent in low- to middle-income settings. To determine the extent to which this is driven by reporting tendencies, we use anchoring vignettes to test and correct for reporting heterogeneity in health among elderly South Africans. Significant reporting differences across wealth groups are detected. Poorer individuals rate the same health state description more positively than richer individuals. Only after we correct for these differences does a significant and substantial health disadvantage of the poor emerge. We also find that health inequality and reporting heterogeneity are confounded by race. Within race groups—especially among black Africans and to a lesser degree among whites—heterogeneous reporting leads to an underestimation of health inequalities between richest and poorest. More surprisingly, we also show that the correction may go in the opposite direction: the apparent black African (vs. white) health disadvantage within the top wealth quintile almost disappears after we correct for reporting tendencies. Such large shifts and even reversals of health gradients have not been documented in previous studies on reporting bias in health inequalities. The evidence for South Africa, with its history of racial segregation and socioeconomic inequality, highlights that correction for reporting matters greatly when using self-reported health measures in countries with such wide disparities.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s13524-018-0709-0) contains supplementary material, which is available to authorized users.

✉ Eddy van Doorslaer
vandoorslaer@ese.eur.nl

[1] Research on Socio-Economic Policy, Economics Department, Stellenbosch University, Stellenbosch, South Africa

[2] Erasmus School of Economics, Erasmus University Rotterdam and Tinbergen Institute, Rotterdam, The Netherlands

[3] Institute of Health Policy and Management, Erasmus University, Rotterdam, The Netherlands

[4] Erasmus School of Health Policy and Management, Erasmus University Rotterdam, Rotterdam, The Netherlands

Keywords Self-assessed health · Vignettes · Health measurement · Inequality · South Africa

Introduction

Around the globe, those with higher incomes can expect to live longer lives and enjoy better health (Bloom and Canning 2000; Deaton 2003), and South Africa is no exception (Case 2004). Studies using household data to measure health outcomes usually have had to rely on self-reported information.[1] Consistently, significant associations have been found between such health measures and subsequent survival in both the industrialized and developing world (see, e.g., Idler and Benyamini 1997; Jylhä et al. 2006; van Doorslaer and Gerdtham 2003), including in South Africa (Ardington and Gasealahwe 2014).

Although it has been established that self-assessed measures contain important health information, they have also been found to be prone to systematic differences in reporting behavior across different socioeconomic groups (Bago d'Uva et al. 2008a, b; Etilé & Milcent 2006; Molina 2016), perhaps due to the use of different comparison or reference groups (Boyce and Harris 2011). Individuals from poorer health communities may report themselves to be relatively better off compared with their reference group, even if their health compares poorly with the overall population (Bago d'Uva et al. 2008b; Etilé and Milcent 2006).

Such differences in the evaluation of self-reported measures of health are usually referred to as *reporting heterogeneity* and imply that for the same health state, certain population subgroups systematically rate their health differently. Clearly, in the presence of reporting heterogeneity by socioeconomic status (SES), the measurement of the socioeconomic gradient in health will be biased (Lindeboom and van Doorslaer 2004).

South Africa is also known to have one of the highest levels of income inequality as measured by the Gini index (World Bank 2011). Therefore, one might expect a more substantial gap in the health reporting behavior of the most and least affluent. The first two objectives of our study are to provide estimates of (1) the extent of reporting heterogeneity and (2) the resulting bias in measured health inequalities by wealth in this particular setting.

We achieve those objectives using ratings of so-called health-anchoring vignettes combined with individuals' ratings of their own health (Bago d'Uva et al. 2008a, b; King et al. 2004). An *anchoring vignette* is a description of the level of health of a hypothetical person, which respondents are asked to evaluate using the same scale as for their own health. This acts as a benchmarking tool that makes it possible to identify reporting heterogeneity with respect to individual characteristics.[2] We use data taken from the WHO Study on Global AGEing and Health (SAGE), a nationally representative sample of persons aged 50 and older in South Africa, collected in 2008 (World Health Organization 2008).

South Africa has a history of economic and political segregation by racial lines that was institutionalized during the apartheid era. Apartheid came to an end only in 1994, when the first democratically elected government came to power (Coovadia et al.

[1] More objectively measured health indicators, such as body mass index (BMI) or biomarkers, do not capture overall health.

[2] An example of a vignette is as follows: "[Luvo] has a lot of swelling in his legs due to his health condition. He has to make an effort to walk around his home as his legs feel heavy."

2009). During the apartheid period, the mobility of race groups other than the minority white population was severely restricted. The black African population group was particularly 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 country's peripheries with their own health departments, which were severely underfunded (Coovadia et al. 2009; McIntyre et al. 1995; Neocosmos 2010). Their access to urban areas where better health care and economic opportunities were available was restricted and regulated.

The racial segregation led to deeply entrenched income/wealth, racial, and health disparities among South Africans. A number of studies have reported racial health disparities favoring the white population in South Africa (Ardington and Gasealahwe 2014; Charasse-Pouélé and Fournier 2006; Lau and Ataguba 2015), even after controlling for wealth. Income disparities by race are also well-known and documented; for instance, Leibbrandt et al. (2011) showed that the bottom income quintiles mostly consist of black Africans, and white, colored, and Asian/Indian respondents[3] are much more concentrated in the top quintiles.[4]

The intertwined relationship among race, wealth, and health in South Africa therefore means that wealth-related reporting heterogeneity in self-assessed health (SAH) may also be confounded by race. Our third and fourth objectives are to quantify the extent to which such tendencies bias estimates of health inequalities, respectively, within and across race groups in this country. A few studies using anchoring vignettes have focused on reporting heterogeneity by wealth or income (Bago d'Uva et al. 2008b; Grol-Prokopczyk et al. 2011; Guindon and Boyle 2012), and more rarely on race (Bzostek et al. 2016; Dowd and Todd 2011), but never on their interlinkage nor on South Africa. Furthermore, studies looking at race have focused primarily on the United States.

Reporting Tendencies and the Health Gradient

Several studies have tested for reporting heterogeneity in self-reported health measures. The majority of these studies 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), but far fewer have covered developing countries (e.g., Bago d'Uva et al. 2008b; Molina 2016; Zhang et al. 2015). Vulnerable subgroups are often found to systematically rate a given level of health as better than do less-vulnerable subgroups (Bago d'Uva et al. 2008a,b; Etilé & Milcent 2006; Molina 2016). This was found, for example, in comparisons of poor and rich individuals in France (Etilé and Milcent 2006) and in Indonesia, India, and China (Bago d'Uva et al. 2008b). Individuals with lower levels of education have

[3] These are self-reported racial categories as identified by Statistics South Africa. Given the country's past, the use and choice of racial categories is contentious and complicated but are considered necessary to address and eradicate social injustices that remain after apartheid (Posel 2001). "Colored" as a racial category consists of numerous individual and distinct groups but was used collectively to describe a mixed ancestry population from indigenous South African, Asian, African, and European descent (Coovadia et al. 2009; de Wit et al. 2010).

[4] The same can be seen in the description of our data in Table 1.

also been found more likely to give a higher rating to a certain health level than the higher-educated (Bago d'Uva et al. 2008a; Molina 2016).

Various possible reasons have been suggested for the observed reporting tendencies. One reason is a comparison with different reference groups, as explained in the Introduction. A second possibility is the asymmetry of health information to which various subgroups have access. If, say, wealthier individuals have better access to health care, which then allows them to be diagnosed with chronic conditions, such access could lead to greater awareness of their ill health. Better health knowledge may in turn affect health expectations (Bonfrer et al. 2014). Sen (2002) offered 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 be easily prevented or remedied with appropriate treatment.

If poorer individuals systematically underestimate their ill health relative to more-affluent individuals, this will be reflected in their health reporting. Health inequality measures will then underestimate the health-by-wealth gap. Various studies have noted this phenomenon. Some have relied on health-anchoring vignettes to directly correct for reporting behavior, whereas others have provided more indirect evidence by comparing with socioeconomic gradients obtained using more objective, observed measures. Bago d'Uva et al. (2008b), for instance, used anchoring vignettes to correct for systematic reporting differences across various socioeconomic groups in India, Indonesia, and China. In all three countries, systematic differences in the reporting behavior of the poor and the nonpoor are found to lead to underestimation of income-related inequality in self-reported health.[5]

In what follows, we formally test for wealth- and race-related reporting heterogeneity in SAH measures in South Africa and examine the implications of their occurrence for measuring health inequalities, using the anchoring vignettes approach.

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 actual associations with health status and reporting heterogeneity (Bago d'Uva et al. 2008a, b; King et al. 2004). This identification problem can be solved with additional data on reporting behavior using anchoring vignettes.[6]

An *anchoring vignette* is a description of the level of health of a hypothetical person. Because this description is fixed across individuals, all systematic variation in vignette ratings with respect to individual characteristics is attributed to reporting heterogeneity.

[5] Bonfrer et al. (2014) provided more indirect evidence of systematic reporting tendencies by comparing the inequalities calculated using objective and self-reported health measures. They found much higher degrees of health inequality by wealth when using objective health measures.

[6] Another approach to identifying reporting heterogeneity is to use more objective measures of health, if available, as proxy indicators for true health status (see, e.g., Etilé and Milcent 2006; Lindeboom and van Doorslaer 2004). This can, however, be problematic if the available proxy indicators are themselves self-reported (Baker et al. 2004). Importantly, in our data, a higher proportion of wealthier individuals report having at least one chronic condition (compared with the poor), suggesting lower awareness for the latter.

Measures of self-assessed health (inequalities) can then be corrected for this reporting heterogeneity (King et al. 2004).

Data

We use data representative of the South African elderly (aged 50 +) population, taken from the WHO's SAGE study, a multicountry study recorded in 2008 and containing 3,840 observations for South Africa. Observations with missing values in any of the variables used in the analysis and individuals older than age 90 are dropped, leaving a remaining analytic sample of 2,968.[7] Data were collected on health status, chronic conditions, disability, health behavior, and health care utilization (He et al. 2012).

Self-assessed Health Domains and Anchoring Vignettes

SAGE asks respondents to rate the difficulty, on a 5-point scale, that they have in (each of) eight health domains: 1 = no difficulty, 2 = mild difficulty, 3 = moderate difficulty, 4 = severe difficulty, or 5 = extreme difficulty. The domains are mobility, self-care, pain, cognition, interpersonal activities, sleep and energy, affect, and vision.[8] In the case of mobility, for instance, the respondent is asked how much difficulty she/he had with moving around in the last 30 days. A similar question structure is applied to the other domains. SAGE collects information on two aspects within each of the eight domains: for instance, in the domain of vision, on far-sightedness and near-sightedness. A detailed description of the 16 health aspects considered in this study, as well as their specific questions, can be found in the online appendix (Table S1). For ease of reference, we refer to these as 16 health domains from here onward.

Subsets of randomly chosen respondents are presented with a selected set of vignettes.[9] For each health domain, the respondent is asked to rate five 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."[10] Respondents are then asked to rate the health of each of the vignettes in the respective domain, using the same 5-point ordinal scale as used in the self-assessment questions.

[7] After listwise deletion, we retain 80 % of the observations within wealth Quantile 1, 79 % within wealth Quantiles 2–4, and 74 % within Quintile 5. The response rate of the race variable is 86 % overall, 88 % for wealth Quintile 1, 86 % within Quintiles 2–4, and 82 % for Quintile 5. Finally, we retain 91 % and 89 % of the black and white groups, respectively.

[8] 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).

[9] The total sample is randomly divided into four subsamples, each of which is asked to rate vignettes in four domains: (1) mobility, vigorous activity, depression, and anxiety; (2) relationships, conflict, body pain, and body discomfort; (3) energy, sleep, far-sighted, and near-sighted; and (4) self-care, appearance, memory, and learning.

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

Sociodemographic Variables

We measure wealth using quintiles of an index created from information on 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 measurement is considered to capture a person's living standards better than income, when the sample consists of both retired and nonretired individuals (Grol-Prokopczyk et al. 2011; Zhang et al. 2015).

Other covariates include gender, age, level of educational attainment of the household head, marital status, race, and urban residence. We account for four racial categories as defined by Statistics South Africa: black African, white, colored, and Indian/Asian.[11]

Table 1 shows distributions of covariates by wealth quintile. Females are more likely to be in the lower quintiles, and 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 wealth quintiles. In the poorest wealth quintile, the great majority (94 %) of respondents are black Africans, but this population group represents only 42 % of the richest wealth quintile. Conversely, Asian, Indian, and white are more concentrated in the top wealth quintiles.

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, a standard model for ordinal dependent variables, and the most common approach in analyses of Likert scale self-assessed health questions (Etilé and Milcent 2006; Jürges 2007). Standard ordered probit models assume that individuals use a common scale when rating their own health and thus do not distinguish between health and reporting differences, which is the aim of our study.

The HOPIT model consists of two components: the reporting behavior component and the own health component. Each is modeled as a generalized ordered probit model, with allowance for heterogeneous cut points (rather than assuming that they are constant, as in the standard ordered probit). The *reporting behavior component* uses anchoring vignette ratings 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:

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

This reflects the first identifying assumption of the anchoring vignette methodology: the *vignette equivalence* assumption, which requires that no systematic differences exist across individuals in their perceptions of the level of functioning described in the

[11] See footnote 3 for a description of racial categories in South Africa.

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:

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

where $m = 1, \ldots, 5, s_i^0 < s_i^1 < s_i^2 < s_i^3 < s_i^4 < s_i^5$, 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 $\mathbf{X}_i$ (which includes wealth, race, and other sociodemographic 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) and, consequently, inclusion of them in the following:[12,13]

$$s_i^m = \mathbf{X}_i^{'}\boldsymbol{\gamma}^m. \tag{3}$$

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

The second component of the HOPIT model—*own health*—is specified as a generalized ordered probit with variable cut points identified by the vignettes in the reporting behavior component. The true own latent health level is typically modeled as a function of the same individual characteristics included in the cut points:[14]

$$H_i^S = \mathbf{X}_i^{'}\beta + \varepsilon_i^s, \qquad \varepsilon_i^s \sim N\left(0, \sigma^2\right). \tag{4}$$

The error terms in Eqs. (1) and (4) are assumed uncorrelated with the observed characteristics in $\mathbf{X}_i$. Finally, similar to the vignette ratings, the own health ratings relate to own latent true level as

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

[12] Some authors have used the same linear specification of the first cut point but the following alternative specification for the subsequent ones: $s_i^m = s_i^{m-1} + exp\left(\mathbf{X}_i^{'}\boldsymbol{\gamma}^m\right)$, $m = 2, \ldots, 4$. Such a HOPIT model fits the data slightly worse/better than one with linear cut points in more/less than one-half of the cases covered here. Although the linear specification does not ensure monotonous cut points, this always holds in our analyses. Finally, the two specifications produce very similar results of tests of reporting heterogeneity and HOPIT–corrected health disparities (results available from the authors upon request).

[13] Following Kapteyn et al. (2007), some authors have also included an unobserved heterogeneity term that explicitly accounts for within-individual correlation of vignette ratings. Our inferences allow for such correlation in an alternative way: namely, by using standard errors clustered at the individual level (see Eq. (4)). Compared with the model used here, the one of Kapteyn et al. (2007) permits an efficiency gain but does not introduce additional flexibility in the identification and correction of reporting heterogeneity: it assumes that the added unobserved heterogeneity term is also uncorrelated with observed characteristics. Kapteyn et al. (2007) reported estimates of their effect of interest resulting from models with and without unobserved heterogeneity. Their two estimates are almost identical.

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

Table 1 Descriptive statistics of covariates by wealth quintile ($N = 2.968$)

	Wealth Quintile				
	1	2	3	4	5
Female Education	0.60	0.58	0.70	0.58	0.53
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
Colored	0.05	0.10	0.13	0.18	0.20
Asian/Indian	0.01	0.01	0.03	0.04	0.09

Note: The table presents sample averages of continuous and binary variables and relative frequencies of categorical variables, weighted with poststratified individual weights.

where the cut points are as defined in Eq. (3). 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 estimates associations between $\mathbf{X}_i$ and health that have been corrected for reporting heterogeneity—that is, represented by the coefficients in the vector $\boldsymbol{\beta}$ in Eq. (4). Because each individual rates multiple vignettes, we use standard errors clustered at the individual level in all inferences based on the HOPIT model.[15]

Evidence on the identifying assumptions of the anchoring vignette methodology is limited and shows mixed results. On vignette equivalence, see Bago d'Uva et al. (2011a), Grol-Prokopczyk et al. (2011, 2015), Kristensen and Johansson (2008), Murray et al. (2003), and Rice et al. (2012); on response consistency, see Bago d'Uva et al. (2011a), Datta Gupta et al. (2010), Grol-Prokopczyk et al. (2011, 2015), and van Soest et al. (2011). In the concluding section of this article, we discuss possible

[15] See footnote 13 for a discussion about an alternative way of accounting for within-individual correlation in vignette ratings.

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 normalize σ_v^2 to 1 and α_1 to 0, 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, and the reporting behavior component uses data from the random subsamples of individuals who rate vignettes in the respective domain.

Results

We analyze inequalities in 16 self-reported health domains among elderly South Africans and the extent to which they might be affected by different reporting tendencies across subpopulations. We first focus on wealth-related health inequalities, which are expected to be substantial given the large economic inequalities that resulted from the apartheid regime (Özler 2007). We address the question of whether reporting heterogeneity causes an underestimation of health inequalities by wealth in South Africa. Because wealth inequality emanated from a regime that enforced the separation of race groups to the advantage of the white population, we subsequently aim to gain a deeper understanding of wealth-related health inequalities by exploring the role of race. To that end, we use a more complete specification to analyze in greater detail (1) health-wealth associations within race groups as well as (2) health-race associations among equally wealthy individuals.

We apply both specifications of the HOPIT model to anchoring vignettes and own health ratings in all 16 domains. The estimated models by domain are used to (1) test for reporting heterogeneity, and (2) estimate health inequalities (un)corrected for reporting heterogeneity (find more detail about all these procedures in upcoming sections). By comparing corrected with uncorrected health inequalities, we can assess the importance and the direction of any reporting bias.

Inequalities in Self-assessed Health Domains by Wealth, Uncorrected

We start with an analysis of inequalities in health by wealth that ignores any reporting heterogeneity using a standard ordered probit model with constant cut points, with the covariates wealth, race, age, gender, marital status, urbanization, and educational achievement as defined earlier.[16] We use this model to predict the probability that an individual reports any difficulty in the respective health domain: that is, categories 2 (mild difficulty) to 5 (extreme difficulty). Figure 1 shows average probabilities for Quintile 1 (Q1) and Quintile 5 (Q5), keeping other variables constant, for all health domains. We observe very small differences between the levels of self-reported difficulties by wealth. 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 with their richer counterparts. In the

[16] Full estimation results of the ordered probit model for the mobility domain can be found in Table S2 of the online appendix.

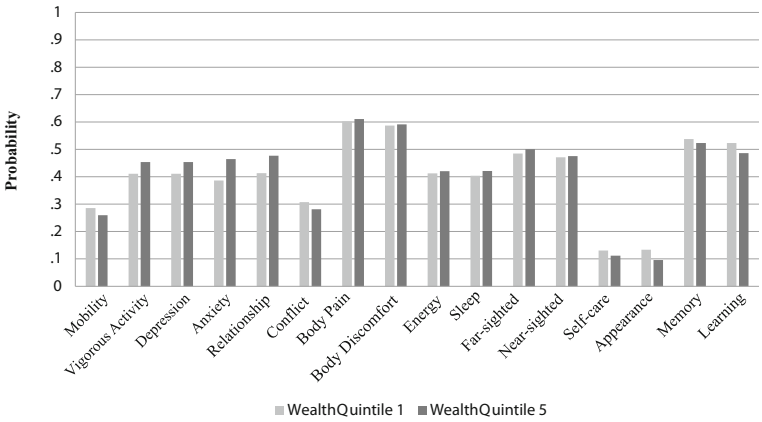

Fig. 1 Estimated probability of reporting any difficulty (mild to extreme) before correcting for reporting bias. Average probabilities estimated from ordered probit models, varying wealth quintile keeping fixed the other covariates (described in Table 1)

following section, we examine whether these patterns may be related to reporting tendencies.

Inequalities in Health by Wealth, Corrected for Reporting Heterogeneity

We now relax the assumption of reporting homogeneity by making use of the vignette ratings and the HOPIT model for each of the 16 domains and 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 earlier defined HOPIT model).[17] We present the results for the poorest (Q1) relative to the richest (Q5) wealth quintile to illustrate and highlight the differences between the two extremes in the wealth distribution in South Africa.

The reporting (or vignette) component of the HOPIT model provides 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 amounts to testing for equality of all coefficients in the vectors $\boldsymbol{\gamma}^m$ ($m = 1, \ldots, 4$) in Eq. (3).[18] Table 2 presents p 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 8 of the 16 health domains.[19]

[17] For illustration, estimation results of the HOPIT model for the mobility domain are shown in Table S2 of the online appendix.

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

[19] The estimated cut points coefficients are reported in the online appendix (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. Thus, 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.

Table 2 Significance tests (p values) of reporting homogeneity between wealth Q1 and wealth Q5, by health domain (N = 2,968)

Health Domain		Health Domain	
Mobility	**.031**	Energy	.771
Vigorous Activity	.380	Sleep	.834
Depression	**.019**	Far-sighted	**.001**
Anxiety	.209	Near-sighted	**<.001**
Relationships	**.017**	Self-care	**.002**
Conflict	.108	Appearance	**<.001**
Body Pain	**.011**	Memory	.091
Body Discomfort	.104	Learning	.730

Notes: Values in bold indicate $p < .05$ (based on standard errors clustered at the individual level). Tests of joint equality of respective coefficients in the cut points of HOPIT models are shown by health domain. HOPIT models include the covariates described in Table 1.

Given the presence of reporting heterogeneity, the self-reported data are likely to be biased and therefore the estimated health inequalities by wealth are also likely to be biased. The own health component of the HOPIT model uses the cut points identified in the reporting component (in Eq. (3)) to estimate partial associations between wealth (and other covariates) and true latent health status that are corrected for reporting heterogeneity (in Eq. (4)). The HOPIT model–based estimates are used to compare health inequalities with and without reporting heterogeneity correction.

To measure wealth-related health inequality, we use the average marginal effect of belonging to Q1 versus Q5 on the probability of having *any* difficulty (i.e., all categories from mild to extreme), keeping other covariates fixed.[20] For each domain, we compute the average marginal effect on that probability twice: (1) using ordered probit models: that is, uncorrected for reporting heterogeneity; and (2) using the estimated HOPIT models and imputing the same fixed cut points across individuals, and thus correcting for reporting heterogeneity. The fixed cut points are those of a reference individual (an unmarried black African male; in wealth quintile 1; aged 62, the average age in the sample; who did not complete primary school; and who lives in a rural area).

To summarize the large number of estimates generated by this procedure graphically, Fig. 2 presents results for all health domains in a radar chart (with estimates detailed in the online appendix, 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 (respectively) to report any difficulty than individuals in Q5, keeping other variables fixed. After correcting for reporting heterogeneity, we do not find a significant gap between the richest and the poorest wealth quintiles in those health domains. The graph shows that across all health domains, the measured health gap by wealth (i.e., the health advantage in favor of the rich) grows after reporting correction. Before correction, the poorest are significantly more likely to report health problems only in the

[20] 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.

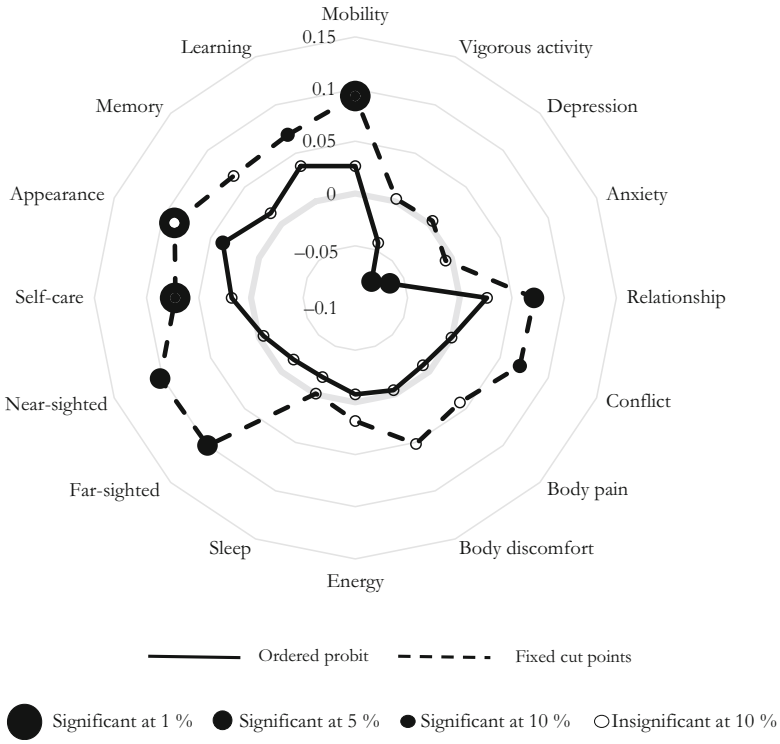

Fig. 2 Average marginal effects of being in wealth Q1 on reporting any difficulty compared with wealth Q5. Average marginal effects from ordered probit models and HOPIT models including the covariates described in Table 1. For the HOPIT model, marginal effects use fixed cut points of reference individual (unmarried black African male; in wealth Quintile 1; aged 62; who did not complete primary school; and who lives in a rural area). Standard errors are clustered at the individual level

appearance domain. However, a very different picture emerges after correction: in 8 of the 16 domains (mobility, relationship, conflict, far-sightedness, near-sightedness, self-care, appearance, and learning), the health by wealth gap becomes significant (at a 10 % level). Moreover, instances in which the poorest reported better health than the wealthiest disappear. These results clearly demonstrate that wealth-related health inequalities are substantially underestimated when uncorrected health measures are used.

Reporting Bias in Health Inequalities by Wealth Among Black Africans

To further unravel the relationships among race, wealth, health, and health reporting, we use a more complete specification. In this section, we examine the race-specific health-wealth associations, focusing on the black African and white population groups, the two most disadvantaged and advantaged groups during apartheid. We categorize wealth of the white population as Q5 versus Q2–Q4.[21] For comparability across racial

[21] As shown in Table 1, we have 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). The subsample of white individuals in Q5 contains 174 observations.

groups, and also crucial for analyses in subsequent sections, we categorize wealth of the black African population as Q1, Q2–Q4, and Q5; these categories correspond to 512, 1,151, and 194 observations, respectively. Small samples sizes do not allow us to distinguish between the wealth effect of Asian/Indian versus colored, but we do include an additional dummy variable for colored to allow for a differential health (reporting) effect for this group. In sum, we consider the following race/wealth variables: colored/Asian/Indian in Q1; colored/Asian/Indian in Q2–Q4; colored/Asian/Indian in Q5; colored; black African in Q1; black African in Q2–Q4; black African in Q5; white in Q2–Q4; and white in Q5. The remaining covariates are defined as earlier, and again we estimate the following for each health domain: (1) a standard ordered probit model (no reporting heterogeneity), and (2) a HOPIT model including all covariates in both the own health equation and in the cut points.

Average estimated probabilities of reporting any difficulty (mild to extreme) by wealth category for black Africans obtained from ordered probit models show that the wealthiest (Q5) often report *worse* health than the poorest (Q1). Differences between the wealthiest (Q5) and middle category (Q2–Q4) are much smaller and not always in the same direction (see the online appendix, Fig. S1).

Using the reporting behavior equation of the HOPIT model, we formally test for reporting heterogeneity and 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 most (10 of 16) domains (detailed results available in the online appendix, Table S5, column 1). The same is true when comparing black African Q2–Q4 respondents with Q1 respondents for 7 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 corrected for reporting heterogeneity and compare these with the uncorrected gaps. Again, we measure these gaps 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 a reference individual (an unmarried black African male; aged 62; in wealth Q2–Q4; who did not complete primary school; and who lives in a rural area) for all respondents. The direction and size of the biases is illustrated in the left panel of Fig. 3 (comparing Q1 with Q5) as well as the right panel (comparing Q2–Q4 with Q5), and detailed results can be found in the online appendix (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—after we correct for reporting differences. For instance, the left panel of Fig. 3 shows that poor black Africans were 0.3 percentage points less likely than rich black Africans to report difficulty with memory before correction. After correction, they are 10 percentage points more likely to do so—a rather 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 in the right panel of Fig. 3 for the middle wealth category (Q2–Q4), albeit with smaller shifts.

Reporting Bias in Health Inequalities by Wealth Among Whites

Prior to the vignette correction, a comparison of the estimated probabilities of reporting any difficulty in each of the 16 health domains between the two wealth categories

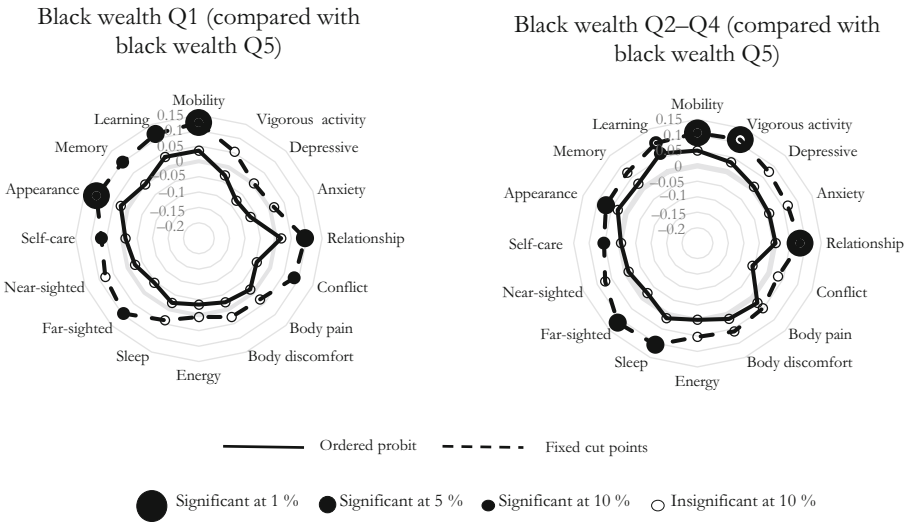

Fig. 3 Average marginal effects of being in Q1 (left) and Q2–Q4 (right) on the probability of reporting any difficulty compared with being in Q5 black African population. Average marginal effects defined as in Table S6 in the online appendix. Ordered probit and HOPIT models control for the same variables as in Table 1 except that race and wealth are controlled for in the following way: black Q1; black Q2–Q4; black Q5; white Q2–Q4; white Q5; other races Q1; other races Q2–Q4; other races Q5; and one dummy variable for colored. Standard errors are clustered at the individual level

defined for the white population reveals a stark contrast to the black African group (results available in the online appendix, Fig. 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 with black Africans.

Results of the formal test using the HOPIT model indeed show little evidence of reporting heterogeneity in the self-evaluation of health by whites in Q2–Q4 compared with those in Q5 (results in the online appendix, Table S5, column 3). In only 4 of the 16 domains can reporting homogeneity be rejected at 5 %.

Thus, for the white population group, the marginal effects of being in Q2–Q4, compared with Q5, on the probability that someone has any difficulty in any of these health domains, are not as affected by heterogeneity correction as they are for black Africans (Fig. 4, and detailed results in Table S6, columns 9–12). Both before and after correction, the less-wealthy whites report to be less healthy than their wealthier counterparts.

One might be concerned that specifying wealth in quintiles impairs the comparison between the results obtained here for the black African and white population groups given that the wealth distribution within quintiles is very different by race.[22] We therefore also estimate HOPIT and ordered probit models with alternative wealth-race

[22] Average wealth in Q5 (Q2–Q4) is approximately 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–Q4.

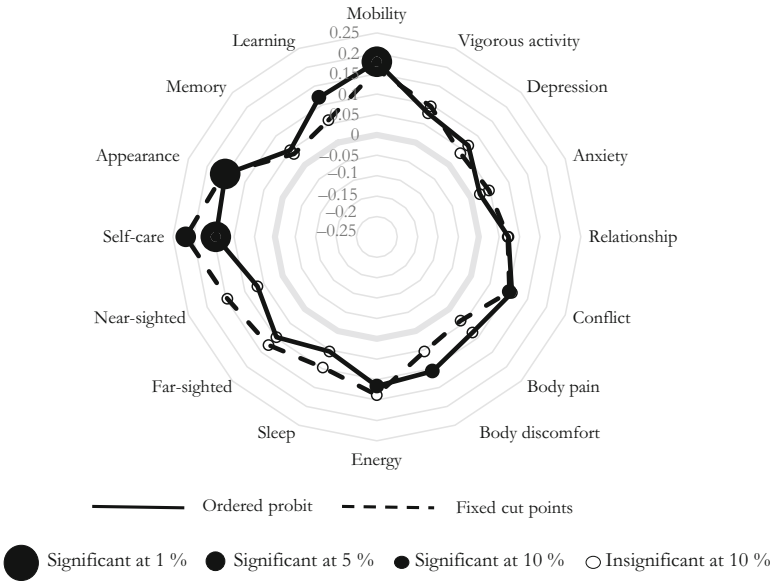

Fig. 4 Average marginal effects of being in wealth Q2–Q4 on the probability of reporting any difficulty (mild to extreme), compared to being wealth Q5 white population. Average marginal effects computed varying wealth quintile within white population group, keeping other variables constant. Ordered probit and HOPIT models specified as in Fig. 3. Standard errors are clustered at the individual level

specifications—polynomials of wealth interacted with race, which enables wealthy versus poor comparisons using the exact same wealth levels for both races. This does not affect our conclusions, and thus we prefer that based on quintiles, which enables the more direct interpretations made earlier.

Inequalities in Health by Race, Within Top Wealth Quintile

From our previous results, reporting bias in the measurement of wealth-related health inequalities is evident and appears to be more problematic among the black African than the white population. Within both populations, the poor have worse actual health outcomes. One question that remains is how the health of the historically disadvantaged black African population compares with that of the white population. We address this question by comparing health (reporting) of equally wealthy (Q5) black Africans and whites, using the same models as in the previous section.[23]

As in previous sections, we compare average predicted probabilities of reporting some difficulty, estimated by using a standard—homogenous reporting—ordered probit model. As shown in Fig. 5, across all domains, black African wealthy individuals report, on average, worse levels of health than the white wealthy individuals.

[23] The conclusions obtained with this specification are also robust to the alternative specifications based on polynomials of wealth interacted with race described earlier. In other words, they are not driven by the fact that white population group in wealth Q1 is, on average, richer than the black African group in the same wealth quintile.

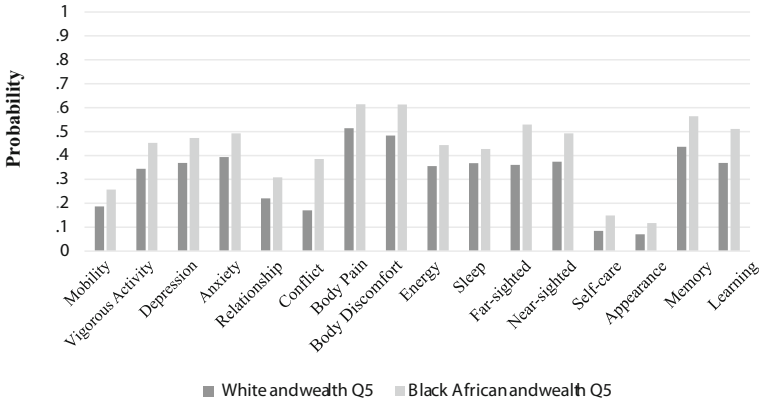

Fig. 5 Estimated probability of reporting any difficulty (mild to extreme) before correcting for reporting bias for black Africans and whites, wealth Q5. Average probabilities estimated from ordered probit models, varying race within wealth Q5, keeping other variables constant. Ordered probit models specified as in Fig. 3

The following results, however, suggest that this relationship between race and health amongst the rich is severely biased. We detect clear evidence of reporting differences, with reporting homogeneity significantly rejected at a 5 % level in 8 of the 16 health domains (results available in the online appendix, Table S5, column 4). Figure 6 shows the marginal effects of being black African and rich versus 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

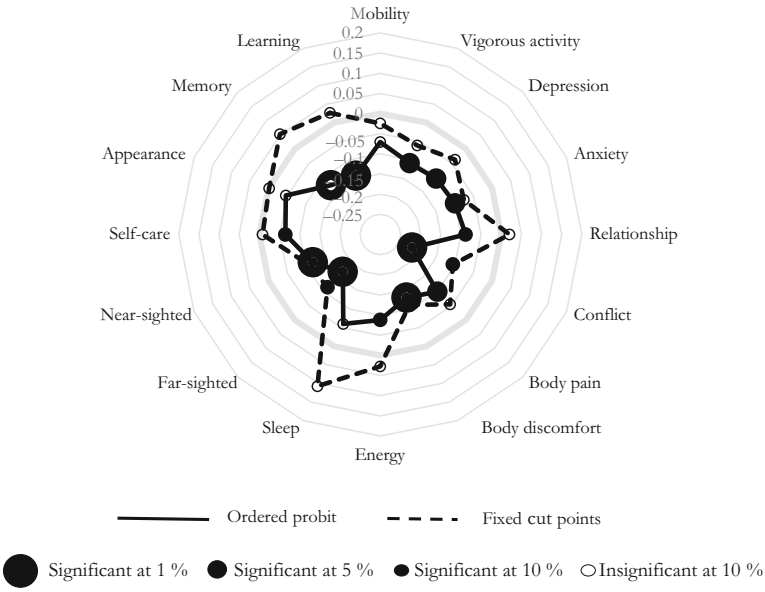

Fig. 6 Average marginal effects of being white on reporting any difficulty (mild to extreme) compared with being black African, wealth Q5. Average marginal effects computed varying race within wealth Q5, keeping other variables constant. Ordered probit and HOPIT models specified as in Fig. 3. Standard errors are clustered at the individual level

to controlling for reporting heterogeneity, black Africans are significantly more likely to report poor health than whites. After fixed cut points are applied, those health gaps are substantially reduced. Although the white population still shows better levels of health across most domains, the differences become much smaller and statistically insignificant.

Conclusion and Discussion

Examination of health differences relies to a considerable extent on asking respondents to rate their health perception and experience. Measurement error in these answers can lead to substantial bias in observed disparities if reporting tendencies are systematically associated with individual characteristics, such as wealth and race. This is particularly worrisome in a country like South Africa, given its history of racial segregation and with income inequalities among the highest in the world. To the extent that factors such as differential health knowledge and reference groups influence health perceptions, reporting bias is likely to be more substantial in such a setting. Furthermore, there is probably no other country where the relationship between wealth and health is so intertwined with race as in South Africa. To our knowledge, our study is the first to examine and correct for health reporting heterogeneity by both wealth and race in South Africa and for the interlinkage between the two.

Using anchoring vignettes and HOPIT modeling, we test and correct for 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 one-half of the health domains (8 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 (Q5) counterparts, leading to a severe underestimation of health gaps by wealth. Observed poor-rich health disparities are small and largely insignificant for all but two domains (depression and anxiety), for which they are even significantly in favor of the poor.

Second, after we correct for these tendencies, substantial disparities between rich (Q5) and poor (Q1) emerge, mostly favoring the rich and significant for one-half the health domains rated (8 of 16). Race is, however, very unequally distributed across wealth quintiles and may play some role in this.

Third, 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–Q4) compared with the richest (Q5) always become larger and are often significant after correction.

Correction of reporting heterogeneity by wealth among white population goes in the same direction but has a much smaller effect in the measurement of wealth-related inequalities. First, the extent of the bias is usually less substantial. Second, unlike for the black African population, significant health advantages of white rich versus the white poor are already observed before the correction.

Finally, we investigate health disparities by race within wealth groups, which can be done only for the top wealth quintile. Interestingly, a very different finding emerges:

before reporting correction, whites report less difficulties than black Africans in *every* health domain (and significantly so in most of them). However, reporting homogeneity is also significantly rejected for one-half the health domains: rich whites 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, the health differences by race are almost removed.

The anchoring 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—with no systematic differences across the subgroups of interest in how the vignette descriptions are understood. Evidence on these assumptions is limited and shows mixed results (Bago d'Uva et al. 2011a; Grol-Prokopczyk et al. 2011, 2015; Kristensen and Johansson 2008; Murray et al. 2003; Rice et al. 2012; van Soest et al. 2011). 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. However, we believe it plausible that they are mainly driven by the latter and 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 as well as across very different vignette descriptions within domains. Although it is not clear whether anchoring vignettes are sufficient to remove *all* bias, the correction is in the expected direction given observed socioeconomic inequalities in "harder" health outcomes, such as survival and disability. These disparities are evident from the longevity disparities that we observe: although 22 % of South Africans in wealth Q5 are aged 50 and older, only 9 % of South Africans in wealth Q1 reach this age category.[24] The evidence of underreporting of health problems by the poor relative to the more-affluent using vignettes is also in line with the apparent lower awareness of chronic conditions mentioned earlier.

Our findings have some important implications. First, they demonstrate 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 wealth in South Africa, it would indeed be very surprising if health were not similarly unequally distributed. Nonetheless, clear health inequality favoring the wealthier emerges only after the differential health reporting between richer and poorer is accounted for. Such consistently large shifts and even reversals of health gradients have not been documented in previous studies. The evidence for South Africa demonstrates that correction for reporting matters greatly when using self-reported health measures in countries with such wide socioeconomic disparities.

Second, other health disparities may be overestimated. After correction for reporting tendencies, the wealthiest black Africans are no longer found to be in worse health than their white counterparts. Our results, therefore, shed light not only on the role of wealth

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

and race in reporting heterogeneity separately but also on the interlinkage between race and wealth. Such evidence is also of relevance for other countries that have experienced racial segregation in the past, which has led to a persistent unequal distribution of wealth.

The implications of our findings go beyond the demonstration of socioeconomic health inequalities in the South African setting. Measures of inequity in health care utilization that rely on self-reports of need have also been shown to be underestimated in European countries (Bago d'Uva et al. 2011b). Correction of this bias reveals that even some of these European countries have not achieved the goal of equal use for equal need. The results of our study suggest that such biases might be much more serious in South Africa. Reporting heterogeneity is potentially relevant for many other research questions in the field of demography, such as neighborhood effects on health. For example, Entwisle (2007) reviewed numerous articles on neighborhoods and health, many of which relied on self-reports. Reporting differences may also contribute to the so-called immigrant health advantage (Riosmena et al. 2017). Cultural differences in the interpretation of self-assessed health questions (Viruell-Fuentes et al. 2011) and lower awareness of chronic conditions due to poor access to health care (e.g., Derose et al. 2009; Jurkowski and Johnson 2005) likely bias the immigrant advantage with respect to native-born populations (Riosmena et al. 2017). To our knowledge, there has not yet been a formal assessment of reporting heterogeneity and its consequences for that type of study.

Last but not the least, our findings also have important implications for health policy. One possible reason why the poor in South Africa underestimate their health needs, compared with the wealthiest, is their underperception of poor health and illness. Such perceptions likely drive the demand for health care (Bago d'Uva et al. 2011b). This poses an additional challenge to the goals of universal health care systems and may, ultimately, contribute to the persistence of real health inequalities. South Africa is committed to obtaining universal health coverage of its population by 2025 (National Department of Health South Africa 2018), but it is currently still far from achieving the goal of providing equal treatment to those in equal needs (Ataguba and McIntyre 2012). One first step in realizing this goal could be to make the poorer more aware of their health needs by providing better access to higher-quality services at lower cost.

Acknowledgments van Doorslaer acknowledges the hospitality and support received while he was a Visiting Fellow at the Stellenbosch Institute for Advanced Studies (STIAS). Rossouw acknowledges the support and feedback received from members of ReSEP and feedback from delegates at the UNU-WIDER conference on inequality (2014), the International Health Economics Associate Conference (2015), and the Economic Society of South Africa biennial conference (2015).

References

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

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.

Bago d'Uva, T., Lindeboom, M., O'Donnell, O., & van Doorslaer, E. (2011a). Slipping anchor? Testing the vignettes approach to identification and correction of reporting heterogeneity. *Journal of Human Resources, 46,* 875–906.

Bago d'Uva, T., Lindeboom, M., O'Donnell, O., & van Doorslaer, E. (2011b). Education-related inequity in health care with heterogeneous reporting of health. *Journal of the Royal Statistical Society: Series A, Statistics in Society, 174,* 639–664.

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,* 1375–1383.

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.

Baker, M., Stabile, M., & Deri, C. (2004). What do self-reported, objective, measures of health measure? *Journal of Human Resources, 39,* 1067–1093.

Bloom, D. E., & Canning, D. (2000). The health and wealth of nations. *Science, 287,* 1207–1209.

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,* 921–937.

Boyce, G. D., & 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*. https://doi.org/10.4102/hsag.v16i1.559

Bzostek, S., Sastry, N., Goldman, N., Pebley, A., & Duffy, D. (2016). Using vignettes to rethink Latino-white disparities in self-rated health. *Social Science & Medicine, 149,* 46–65.

Case, A. (2004). Does money protect health status? Evidence from South African pensions. In D. A. Wise (Ed.), *Perspectives on the economics of aging* (pp. 287–312). Chicago, IL: University of Chicago Press.

Charasse-Pouélé, C., & Fournier, M. (2006). Health disparities between racial groups in South Africa: A decomposition analysis. *Social Science & Medicine, 62,* 2897–2914.

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. *Lancet, 374,* 817–834.

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

Deaton, A. (2003). *Health, income, and inequality* (NBER Reporter: Research Summary). Cambridge, MA: National Bureau of Economic Research.

Derose, K. P., Bahney, B. W., Lurie, N., & Escarce, J. J. (2009). Review: Immigrants and health care access, quality, and cost. *Medical Care Research and Review, 66,* 355–408.

de Wit, E., Delport, W., Rugamika, C. E., Meintjes, A., Möller, M., van Helden, P. D., . . . Hoal, E. G. (2010). Genome-wide analysis of the structure of the South African coloured population in the western cape. *Human Genetics, 128,* 145–153.

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

Entwisle, B. (2007). Putting people into place. *Demography, 44,* 687–703.

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

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,* 246–261.

Grol-Prokopczyk, H., Verdes-Tennant, E., McEniry, M., & Ispány, M. (2015). Promises and pitfalls of anchoring vignettes in health survey research. *Demography, 52,* 1703–1728.

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,* 29–40.

He, W., Muenchrath, M. N., & Kowal, P. R. (2012, May). *Shades of gray: A cross-country study of health and well-being of the older populations in SAGE countries, 2007–2010* (International Population Reports No. P95/12-01). Washington, DC: U.S. Census Bureau.

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 Working Paper 05/04). York, UK: University of York Department of Economics and Related Studies.

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

Idler, E. L., & Benyamini, Y. (1997). Self-rated health and mortality: A review of twenty-seven community studies. *Journal of Health and Social Behavior, 38,* 21–37.

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

Jurkowski, J. M., & Johnson, T. P. (2005). Acculturation and cardiovascular disease screening practices among Mexican Americans living in Chicago. *Ethnicity & Disease, 15,* 411–417.

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,* 465–471.

Kapteyn, A., Smith, J. P., & van Soest, A. (2007). Vignettes and self-reports of work disability in the United States and the Netherlands. *American Economic Review, 97,* 461–473.

King, G., Murray, C. J. L., Salomon, J. A., & Tandon, A. (2004). Enhancing the validity and cross-cultural comparability of measurement in survey research. *American Political Science Review, 98,* 191–207.

Kristensen, N., & Johansson, E. (2008). New evidence on cross-country differences in job satisfaction using anchoring vignettes. *Labour Economics, 15,* 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. https://doi.org/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 No. 64). Cape Town, South Africa: Southern Africa Labour and Development Research Unit, University of Cape Town.

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,* 1024–1046.

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

McIntyre, D., Bloom, G., Doherty, J., & Brijlal, P. (1995). *Health expenditure and finance in South Africa*. Durban, South Africa: Health Systems Trust; and Washington, DC: World Bank.

Molina, T. (2016). Reporting heterogeneity and health disparities across gender and education levels: Evidence from four countries. *Demography, 53,* 295–323.

Murray, C. J., Ozaltin, E., Tandon, A., Salomon, J., Sadana, R., & Chatterji, S. (2003). Empirical evaluation of the anchoring vignette approach in health surveys. In C. J. L. Murray & D. B. Evans (Eds.), *Health systems performance assessment: Debates, methods and empiricism* (pp. 369–399). Geneva, Switzerland: World Health Organization.

National Department of Health South Africa. (2018). *National health insurance.* Retrieved from http://www.health.gov.za/index.php/component/phocadownload/category/383

Neocosmos, M. (2010). *From "foreign natives" to "native foreigners": Explaining xenophobia in post-apartheid South Africa. Citizenship and nationalism, identity and politics*. Dakar, Senegal: Council for the Development of Social Science Research in Africa.

Özler, B. (2007). Not separate, not equal: Poverty and inequality in post-apartheid South Africa. *Economic Development and Cultural Change, 55,* 1–41.

Posel, D. (2001). Race as common sense: Racial classification in twentieth-century South Africa. *African Studies Review, 44*(2), 87–114.

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,* 337–369.

Riosmena, F., Kuhn, R., & Jochem, W. C. (2017). Explaining the immigrant health advantage: Self-selection and protection in health-related factors among five major national-origin immigrant groups in the United States. *Demography, 54,* 175–200.

Salomon, J. A., Mathers, C. D., Chatterji, S., Sadana, R., Üstün, T. B., & Murray, C. J. L. (2003). Quantifying individual levels of health: Definitions, concepts, and measurement issues. In C. J. L. Murray & D. B. Evans (Eds.), *Health systems performance assessment: Debates, methods and empiricism* (pp. 301–318). Geneva, Switzerland: World Health Organization.

Sen, A. (2002). Health: Perception versus observation. *BMJ, 324,* 860. https://doi.org/10.1136/bmj.324.7342.860

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

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,* 575–595.

Viruell-Fuentes, E. A., Morenoff, J. D., Williams, D. R., & House, J. S. (2011). Language of interview, self-rated health, and the other Latino health puzzle. *American Journal of Public Health, 101,* 1306–1313.

World Bank. (2011). *The World Bank DataBank* [Database]. Retrieved from http://databank.worldbank.org/data/reports.aspx?source=2&series=SI.POV.GINI&country=#

World Health Organization. (2008). *WHO Study on global AGEing and adult health* (*SAGE*). Retrieved from http://www.who.int/healthinfo/sage/en/

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.