Economics Letters 121 (2013) 163–166

Contents lists available at ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

# Can you trust survey responses? Evidence using objective health measures

Agne Suziedelyte[a,*], Meliyanni Johar[b]

[a] School of Economics, University of New South Wales, Australia
[b] Economics Discipline Group, University of Technology Sydney, Australia

## HIGHLIGHTS

- We investigate measurement error in objective survey variables.
- Major medical operations are used as a case study.
- Cross-checking self-reports with administrative data we find large reporting errors.
- These errors are not explained by observed individual characteristics.
- Hence coefficients are not biased when the survey variables are used as outcomes.

## ARTICLE INFO

*Article history:*
Received 17 May 2013
Received in revised form
29 July 2013
Accepted 29 July 2013
Available online 6 August 2013

*JEL classification:*
C1
I1
I2

*Keywords:*
Measurement error
Health
Socio economic status

## ABSTRACT

We investigate the common assumption in applied research that reporting errors are negligible in variables where there is no clear incentive for misreporting. Using major medical operations, we find high misreporting rates, but the coefficients of their predictors remain unbiased.

© 2013 Elsevier B.V. All rights reserved.

## 1. Introduction

The ever-growing acknowledgement of non-random measurement error in self-reported data has motivated researchers to investigate its impact (e.g., Johnston et al., 2009, Akee, 2011, Erickson and Whited, 2000, Forbes, 2000, Dwyer and Mitchell, 1999) and to improve techniques to deal with it in empirical models (Bound et al., 2001; Black et al., 2000; Brownstone and Valletta, 1996). The measurement error literature has focused on variables where misreporting may be perceived to provide an advantage, real or psychological, for the respondent. For example, studies have examined measurement errors in self-reported earnings (Akee, 2011; Brownstone and Valletta, 1996); disability status (Kreider and Pepper, 2007); grades, class ranks and test scores (Kuncel et al., 2005); height, weight and body mass index (Gorber et al., 2007); sexual behaviour (Tennekoon and Rosenman, 2013); and subjective health or wellbeing (Bago d'Uva et al., 2008; Butler et al., 1987; Dwyer and Mitchell, 1999). Labour market outcomes, wealth and income are often misreported to gain a tax or subsidy advantage; principals and teachers may inflate their school's achievement scores in order to improve reputation; due to stigma associated with undesirable body shape, survey respondents often misreport their weight and height; subjective variables such as self-rated general health or well-being may not reflect the true health and well-being conditions, because the answers to these questions are affected by adaptation bias and cultural norm. For other variables, there is no clear incentive for the survey respondents to be untruthful and/or they are deemed to be objective. These variables may also be reported with errors, but they are often

---

[*] Correspondence to: Australian School of Business Building, Level 4, 447, UNSW Sydney NSW 2052, Australia. Tel.: +61 293852660.
*E-mail address:* agne@unsw.edu.au (A. Suziedelyte).

0165-1765/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.econlet.2013.07.027

assumed to be harmless and random (conditional on basic demographics). The purpose of this study is to confirm such assumption, using the case of life-changing medical operations.

In contrast to previous studies, we investigate reporting errors in variables that might be expected to have small, most likely random errors. If they are found to be truly random and therefore harmless to subsequent analysis then we can have confidence in relying on the self-reports of such variables. We use the case of major medical operations that are likely to be memorable. We cross-check the self-reports of operations with contemporaneous hospital administrative records. To examine the consequence of ignoring a reporting error, we compare the socio economic status (SES)-health gradient estimated based on the self-reported and administrative data.

Our study is related to Johnston et al. (2009) who show that reliance on self-reported chronic conditions can lead to underestimation of income-related inequalities in health because low income individuals are more likely than high income individuals to under-report bad health. As a result, the size of the income-health gradient using measured health is much stronger than that implied by self-reported health. Their study uses hypertension as a measure of health and finds that over 85% of individuals measured by a trained nurse to have hypertension do not indicate in the survey to have hypertension. One issue with the study of Johnston et al. (2009) is that disparity in hypertension rates may reflect infrequent health checks, rather than reporting error. Moreover, the respondents may have been prescribed hypertension medication and believe that their blood pressure has returned to normal. Our use of major, life-changing operations does not suffer from these drawbacks.

## 2. Data and method

Our analysis sample is based on 241,138 non-institutionalised individuals who participated in the 45 and Up Study[1] in the state of New South Wales (NSW), Australia, fielded in stages during 2006–2009; NSW has a population of about 7.3 million with 39% aged over 45. In this survey, respondents were prompted with a list of operations and asked whether they had any of these operations, and for each of them, the age at the most recent operation. The 45 and Up Study is linked to the Admitted Patient Data Collection (APDC) consisting of all admissions in NSW hospitals during 2000–2009.[2] This data linkage allows us to cross-check the consistency of the self-reported history of operations with the hospital records during this period. From the operations listed in the survey, we select four operations that are performed in a hospital setting and specifically defined so that matching them in the administrative data is straightforward. For example, it is difficult to match "heart operation", because we cannot tell from the survey question which procedures should be included in the definition of "heart operation" in the administrative data. The four operations are knee replacement, removal of gall bladder, removal of prostate, and hysterectomy.

So that all respondents have an equal cross-check window, we focus on the hospital records in the last 5 years from the survey date. To be consistent, using the survey data, we create binary variables that indicate whether or not a respondent reported an operation in the last 5 years. Imposing the 5 years window, we define four possible cases: (1) an operation found both in the survey and hospital data (true positive); (2) an operation not found in the survey and hospital data (true negative); (3) an operation found in the administrative data but not in the survey data (false negative or under-reporting); and (4) an operation found in the survey data but not in the administrative data (false positive or over-reporting).

Our analysis has three stages. In the first stage for each of the above mentioned operations, we present the under-reporting rates, calculated as the proportion of operations in the hospital records that are not found in the survey data, and the over-reporting rates, calculated as the proportion of operations in the survey data that are not found in the hospital records. The under-reporting rates are due to respondents either: (i) not reporting the operation at all or (ii) misreporting the timing of the operation. Since we are only looking at operations in the last 5 years, it is unlikely that respondents would forget having them, especially since the selected operations are life-changing, occurring only once in a lifetime and/or involving major interventions. But there could be other reasons why respondents did not report an operation, including the question being considered too sensitive and lack of attention or medical knowledge (comprehension) in filling in the survey. On the other hand, the over-reporting rates may be also explained by the limited coverage of the administrative data. For instance, operations overseas will not be captured in the data, although we can expect them to be rare given that Australian public hospitals provide these operations for free. Operations performed in other Australian states are also outside the data range, but studies suggests that incentives to travel inter-state are small (e.g., Johar et al., 2012).

In the second stage, we investigate whether the propensity of misreporting (over or under) can be explained by individual characteristics, capturing variation in memory capabilities, preferences for health, and medical knowledge. *R*-squared statistics from OLS regressions are used as measures of explanatory power. We reveal the proportion of the variation in reporting error that is explained by individual characteristics such as age, sex, ethnicity, marital status, education, income, employment, housing, neighbourhood characteristics, and family health history. We also investigate whether own health variables have additional explanatory power over and above the other individual background variables.

The last stage assesses the implications of reliance on survey data. We run probit regressions of the propensity of having each of these operations using first survey then hospital data. We pay specific attention to the slope coefficients of SES measures: education, income and employment. If the measurement error is random, we would find a consistent SES-health gradient, regardless of the source of data.

In addition, we also look at hypertension, which unlike operations, may be harder for an individual to notice, because it is often asymptomatic. Previous literature finds that misreporting in such hard-to-observe variables correlates with SES (Johnston et al., 2009). We verify self-reported diagnosis of hypertension using administrative data on claims for anti-hypertensive prescription drugs and hypertension diagnoses in hospital admission records in the next 12 months from the survey date.

## 3. Results

The first two columns of Table 1 present the incidence of the selected operations and hypertension by each data source. The incidence of most of the operations is higher in the survey data than in the hospital records, suggesting that respondents are more likely to over-report than under-report the operations. On the other hand, hypertension tends to be slightly over-reported. Columns 3 and 4 show the extent of the under- and over-reporting. As an example, there are 1294 self-reports of knee replacements that are not found in the hospital data, giving an over-reporting

---

[1] For details see http://www.saxinstitute.org.au/. Participants were randomly selected from the central database, which contains everyone who has ever used health services in Australia, for this population age group. The survey was done in stages, but the bulk of it (about 80%) was collected in 2008. This variation in survey year is due to sampling process rather than the choice of respondents.

[2] The data linkage is undertaken by the Centre for Health Records Linkage and the linked, de-identified data is released under ethics approval.

Download English Version:

https://daneshyari.com/en/article/5059090

Download Persian Version:

https://daneshyari.com/article/5059090

Daneshyari.com