Journal of Health Economics 27 (2008) 1275–1284

Contents lists available at ScienceDirect

Journal of Health Economics

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

Optimal recall length in survey design

Philip M. Clarke[a,*], Denzil G. Fiebig[b], Ulf-G. Gerdtham[c,d,e]

[a] School of Public Health, University of Sydney, Edward Ford Building (A27), The University of Sydney, NSW 2006, Australia
[b] School of Economics, University of New South Wales, Australia
[c] Department of Economics, University of Aberdeen, Scotland
[d] Health Economics Research Unit (HERU), University of Aberdeen, Scotland
[e] Department of Clinical Sciences, Lund University, Malmö, Sweden

ARTICLE INFO

Article history:
Received 5 October 2005
Received in revised form 18 April 2007
Accepted 14 May 2008
Available online 26 June 2008

JEL classification:
C42
I1

Keywords:
Survey methods
Health care
Errors in variables
Recall error

ABSTRACT

Self-reported data collected via surveys are a key input into a wide range of research conducted by economists. It is well known that such data are subject to measurement error that arises when respondents are asked to recall past utilisation. Survey designers must determine the length of the recall period and face a trade-off as increasing the recall period provides more information, but increases the likelihood of recall error. A statistical framework is used to explore this trade-off. Finally we illustrate how optimal recall periods can be estimated using hospital use data from Sweden's Survey of Living Conditions.

1. Introduction

Self-reported data collected via surveys are a key input into a wide range of research conducted by economists and other social scientists. For example, in health economics, economic evaluations of health care interventions are often based on self-reported health care data that is collected during the course of a study. It is also routinely collected in national health surveys such as the Health Survey for England, Australia's National Health Survey or the Medical Expenditure Survey in the US. While these surveys involve interviews conducted over the course of a year, there is considerable variation in the period over which subjects are asked to recall their previous health care use. This has been highlighted by a recent OCED study comparing inequity in access to health care across 21 developed countries based on health surveys in different countries. The period of recall for primary care and aspects of hospital use ranged from 2 weeks in the Australian National Health Survey to one year in the European Community Household Panel (Health Equity Research Group, 2004).

It has been widely recognized that there is an inverse relationship between the length of time over which subjects are asked to recall prior use and the accuracy of the reported estimates. The longer the period of recall the greater the likelihood of error. An early study by Sudman and Bradburn (1973) on the impact of the length of the recall period on response suggested two types of memory error may arise: (i) a respondent may forget an episode of use entirely; and (ii) they may remember the episode but incorrectly recall when it occurred. In regard to the latter, a particular form of error is *telescoping* which arises

* Corresponding author. Tel.: +61 2 9351 5424; fax: +61 2 9351 7420.
E-mail address: philipc@health.usyd.edu.au (P.M. Clarke).

0167-6296/$ – see front matter © 2008 Published by Elsevier B.V.
doi:10.1016/j.jhealeco.2008.05.012

when the respondent recalls events that occurred before the period in question. For example, a person my report they have been in hospital in the last three months, even though the actual event was five months prior to the date of interview. While these errors can work in opposite directions over the last three decades validation studies of self-reported health care use have consistently shown that under-reporting is more common than over-reporting especially of primary care visits (e.g. see Cannell et al., 1977 and for a more recent review Bhandari and Wagner, 2006). Underreporting over longer recall periods may also characterise other types of data that are a focus of health economic research such as reported use of illicit drugs (see Bachman and O'Malley, 1981).

While recall error is undesirable, a survey with a short length of recall provides very little information about an individual's normal health care use. As Deaton (1997, p. 24) has noted in the context of measuring general household consumption expenditure:

> "...if the object of the exercise is to estimate average consumption over a year, one extreme is to approach a sample of households on January 1 and ask each to recall expenditures for the last year. The other extreme is to divide the sample over the days of the year, and to ask each to report consumption for the previous day. The first method would yield a good picture of each household's consumption, but runs the risk of measurement error... [t]he second method... will give a good estimate of the mean consumption over all households... [but] will yield estimates of individual expenditure that... are only weakly related to normal expenditures..."

Suppose self-reported consumption or utilisation data are to be collected by a survey. At the survey design stage there is a choice of the length of the recall window. Even if the variable of interest is consumption over a particular period, say a year, errors induced by having long recall windows equal to the period of interest suggest that one could opt for a shorter window or sub-period as the basis of the survey question. While this provides a less error ridden measure it comes at the cost of providing less information. Hence, the basic research question is: "What is the optimal recall window over which to ask the utilisation question when recall error increases with window length but a longer window is likely to provide more (albeit imperfect) information directly relevant to the variable of interest?"

We investigate these tradeoffs in order to determine what factors impact on the optimal recall window over which to ask consumption and utilisation questions. Our analysis distinguishes between two types of recall error. Firstly, error due to forgetting (or telescoping) that may impact on the average reported health care use in one direction leading to under (or over) reporting. This can be regarded as bias in the reported use. Another potential source of bias is rounding which is explored by Huttenlocher et al. (1990) who note respondents tend to round off large values. For example, a physician when responding to a survey on particular patients use over a year: "may use raw values up to roughly 5, then round by 5s to roughly 20 and round by 10s past 20" (p. 211). They argue that rounding is likely to impart a downward bias on the reported value. Secondly, reporting errors that are more random in nature. This may not result in a biased measure, but will increase the degree of variation in the reported measures.

The rest of this study is structured as follows. The next section outlines a simple framework that is used to capture the key dimensions of the research question. The third section derives statistical criteria for examining the consequences of choosing different recall periods. The fourth section applies these to illustrate how optimal recall periods can be estimated for a question on hospital usage using data from Statistics Sweden's Survey of Living Conditions. Section five discusses the implications of these results and avenues for future research. The final section draws some conclusions.

2. A framework for analysis

Denote the variable of interest for the ith individual by Y_i ($i = 1, \ldots, N$). If this could be recalled without error then there would be no problem. Such a situation represents a benchmark for comparisons as specific forms of recall errors are introduced. Suppose the period over which the utilisation variable is required, call this the *target period*, can be divided into S sub-periods where the division is determined so that individuals can accurately recall their utilisation over the most recent sub-period. If we were interested in how much someone spent on medications over a year but the longest window over which they can recall last period's drug expenditure without error is say a month then S is set to 12. We denote the *recall period* (i.e. the period of time over which we ascertain previous use) as w and let $w = 1, 2, 3, \ldots, S$. While by definition the shortest recall period ($w = 1$) is without error, we may want to consider longer periods as we are ultimately interested in annual consumption and a window of less than a year implies that an imputation process must be undertaken in order to estimate annual use.

Now denote actual utilisation over the recall window w by Y_i^w. However, when asking the survey question the error ridden response will be given by:

$$X_i^w = Y_i^w + v_i^w \tag{1}$$

where v_i^w represents the measurement error. By assumption, $X_i^1 \equiv Y_i^1$ but for $w > 1$ measurement error can occur because of incomplete or inaccurate recall. There are other potential causes for measurement error. In particular, respondents could act strategically or simply provide false answers to sensitive questions. In such cases the form of the measurement error is likely to be different from that associated with recall problems as are the likely remedies; see for example Carson et al. (1999) and Philipson and Malani (1999). Other motivations for survey responses are discussed in Tourangeau et al. (2000).