Journal of Economic Behavior & Organization 96 (2013) 65–84

Contents lists available at ScienceDirect

Journal of Economic Behavior & Organization

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

Survey response in probabilistic questions and its impact on inference☆

Jochem de Bresser[a,*], Arthur van Soest[b]

[a] Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands
[b] Tilburg University and Netspar, P.O. Box 90153, 5000 LE Tilburg, The Netherlands

ARTICLE INFO

Article history:
Received 30 July 2012
Received in revised form 4 September 2013
Accepted 23 September 2013
Available online 2 October 2013

JEL classification:
C81
D84
J14

Keywords:
Expectations
Survey response
Rounding
Recall error
Unobserved heterogeneity
Aging

ABSTRACT

We develop a panel data model of expectations of a continuous outcome variable elicited on a percentage-chance scale. The model explains the location and dispersion of the subjective distributions by socio-economic covariates and unobserved factors. Moreover, it accounts explicitly for non-response, non-informative focal answers, and recall and rounding errors. We apply the model to the expected retirement income replacement rate of Dutch wage workers. We find that incorporating these features of the answering process increases the size and significance of relationships with covariates. The estimates indicate substantial rounding but few focal answers. Respondents tend to stick to a certain answering strategy: non-response, rounding and especially non-informative focal answers are characterized by substantial unobserved heterogeneity across individuals.

1. Introduction

The idea of forward-looking agents who take account of the future consequences of their decisions has long been commonplace among economists. The behavior of such agents is driven by their preferences and expectations, neither of which are usually directly observed. Before the 1990s almost all researchers pinned down expectations by assuming rationality, so that expectations coincide with actual stochastic processes that can be estimated. From the early 1990s economists have started to ask survey respondents about their expectations, allowing rationality assumptions to be verified and relaxed. Since then questions on subjective expectations have been included in several large socio-economic surveys like the Michigan Survey of Consumers, the Health and Retirement Study and the Survey of Economic Expectations, see Dominitz and Manski (2006). To simplify interpretation and interpersonal comparisons, expectations are often measured by means of items that ask respondents to think in terms of probabilities. The quality of this kind of data has been questioned; as noted by Manski

☆ We would like to thank Luc Bissonnette, an anonymous referee and participants at the 2011 Netspar workshop in Amsterdam and at the 2011 ESEM conference in Oslo for helpful comments on an earlier draft of the paper.

* Corresponding author at: Present address: University of Groningen, Faculty of Economics and Business, P.O. Box 800, 9700 AV Groningen, The Netherlands. Tel.: +31 50 363 7018.
E-mail addresses: j.debresser@gmail.com (J. de Bresser), avas@uvt.nl (A. van Soest).

0167-2681/$ – see front matter © 2013 Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.jebo.2013.09.012

(2004), the mere fact that surveys come up with relatively high response rates on probabilistic items does not imply that the responses indeed reflect respondents' true beliefs.

In this paper, we analyze the validity of survey items used to measure expectations by simultaneously modeling expectations and various aspects of survey response. We draw on a representative, longitudinal sample of Dutch employees aged 25–64 who were asked about their expected replacement rate of income at retirement. Our measure of the replacement rate compares real expected pension income with real current income. We adjust the model of Kleinjans and Van Soest (forthcoming) for probabilities of binary events to a situation in which several probabilistic items are used to fit subjective log-normal distributions of a continuous outcome along the lines of Dominitz and Manski (1997). Location and dispersion measures are allowed to depend upon observed and unobserved respondent characteristics. In our data, we have questions on the probability that the retirement replacement rate is less than 50%, 60%, 70%, 80%, 90% and 100%, for five waves of panel data (2006–2010). We model non-response, non-informative answers and rounding, issues which have been identified as potentially important aspects of the answering process of subjective probability questions (Bruine de Bruin et al., 2000; Dominitz and Manski, 1997). By comparing with models ignoring these aspects, we also analyze the impact of such processes on the conclusions regarding the location and dispersion of the fitted distributions.

Retirement income in the Netherlands is organized along three pillars: pay-as-you-go public pensions, fully funded occupational pensions and private savings (Bovenberg and Meijdam, 2001). The first two together provide 80 percent of income during retirement for the average employee. The first pillar pay-as-you-go public pension provides retirees with a subsistence level income; it neither reflects individual contributions nor previously earned income. The second pillar consists of fully funded occupational pensions that cover about 90 percent of Dutch employees. These industry- or company-specific provisions are usually defined benefit schemes, though defined contribution arrangements are becoming more common.[1] In contrast to the public pension, the level of occupational benefits is based on previous wages through average or final wage arrangements.

The third pillar of retirement income consists of private savings, but due to the comprehensive level of the first two pillars, private arrangements are a relatively minor part of pension income for most employees. The first and second pillar pensions are probably easier to predict than defined contribution pensions or third pillar pension savings (often invested in the stock market), making the Dutch pension system especially suitable for research on retirement expectations. In our survey, pension income is defined to include the first two pillars, discarding potential income from the third pillar.

Our results indicate that there is statistically significant and quantitatively important unobserved heterogeneity at the level of the individual in all parts of the model. Especially, there is a small group of respondents that consistently give non-informative focal point (fifty-fifty) answers. Moreover, the individual effects of the various aspects of survey response are correlated among each other and with the unobserved heterogeneity terms in the parameters of the location and dispersion of the subjective distributions. Though we also find some statistically significant sequence effects (effects common to the series of six questions answered by a given respondent in one specific wave), these effects are small compared to the individual effects that are persistent over time. We find rounding is important, with almost 50 percent of the answers rounded to a multiple of 10. Non-informative fifty-fifty answers, on the other hand, make up only 1.3 percent of the reported probabilities.

A full specification including covariates indicates that non-response, rounding, and reporting error vary significantly with socio-economic respondent characteristics, as well as the location and dispersion of the underlying subjective probability distributions. Finally, the magnitude and significance of the relationships between the location and dispersion of the subjective distributions and the covariates is larger in the joint model of the answering process and expectations than in a linear random effects model that ignores rounding, non-response, and focal answers. This suggests it is useful to take the answering process into account when analyzing the association between subjective expectations and respondent characteristics.

The structure of this paper is as follows. Section 2 summarizes related literature, both with regard to subjective (pension) expectations and the econometric model applied. Section 3 introduces the data and provides descriptive statistics. Section 4 describes the model in detail. Estimation results are presented in Section 5. Section 6 concludes.

2. Literature

This paper fits in with the literature on probabilistic expectations in general and pension expectations in particular. Expectations play an important role in economic models of intertemporal choice. Rational, forward-looking agents not only consider the present consequences of their actions, but also (their perception of) their impact on future utility. For instance, Delavande and Rohwedder (2011) show that individuals who are uncertain about their future level of social security benefits tend to hold a smaller share of their portfolio in stocks, reducing the riskiness of their retirement provisions. Direct measurement of expectations is especially important to separately identify expectations and preferences – different combinations of preferences and expectations may be behaviorally equivalent, but may yield different policy implications (Manski, 2002).

Spurred by a growing body of evidence in favor of survey respondents' willingness and ability to answer questions in a probabilistic format, measuring expectations directly through probabilistic questions gained momentum during the late 1990s and early 2000s and has become the standard elicitation method (see Manski, 2004, and Hurd, 2009, for

[1] Compensation for inflation (indexation) is not guaranteed so that only nominal benefits are guaranteed.