Keeping the oldest old – A framework for survey adaptations to improve panel retention in the Survey of Health, Ageing and Retirement in Europe (SHARE)

Michael Bergmann, Johanna Bristle, and Annette Scherpenzeel

Munich Center for the Economics of Aging (MEA), Max Planck Institute for Social Law and Social Policy, Technical University of Munich

Email: bergmann@mea.mpisoc.mpg.de, bristle@mea.mpisoc.mpg.de, scherpenzeel@mea.mpisoc.mpg.de

Working paper for the 30th International Workshop on Household Survey Nonresponse
25.-28. August 2020 (virtual meeting)

Introduction

Panel attrition is and remains a major challenge for maintaining long-running panel surveys, especially for vulnerable groups. The oldest old are such a vulnerable group and at the same time are the fastest growing group of the population (Andersen-Ranberg, Petersen, Robine, & Christensen, 2005; Eurostat, 2019), which makes them an important group of investigation in the Survey of Health, Ageing, and Retirement (SHARE). SHARE is a panel study, conducted bi-annually across 27 European countries plus Israel that studies ageing societies. The survey explicitly looks at socio-economic conditions, health, and social networks in older age. Therefore, it is of great importance to maintain good representation of the elderly and to avoid underrepresentation or sample selectivity by selective dropout due to old age. In part (i.e. besides death), this could be prevented by improving the survey's questionnaire and fieldwork procedures.

With this project, we want to detect the most prominent determinants, events, and changes in old age that lead to panel attrition, besides the physical death of respondents, and identify potential areas for intervention throughout the whole survey lifecycle. In case of a higher risk for dropout for the oldest old, we would then aim on developing survey adaptations that make SHARE more fitting for the oldest old and, by this, tackle preventable dropout in a panel study. The project consists of three consecutive steps: (1) Identification of underlying causes of attrition regarding the oldest old; (2) Identification of potential intervention areas; and (3) Implementation of field adaptions. Step 2 and especially step 3 aim at specific implementations that will be based on the results of step 1. Preliminary results of this first step are the focus of this contribution.

In this paper, we analyze the composition of the oldest old (80+) sample in SHARE and their likelihood to stay in the panel based on individual characteristics and changes in prior waves. Therefore, we employ SHARE data from the first seven waves to predict dropout across all waves using logistic regression to answer the following questions: Who are the oldest old that participate in SHARE and what are the most relevant factors for their attrition compared to the younger elderly (i.e. 50-79 year-olds)?

From the literature on "hard-to-survey" populations (Tourangeau, 2014) we know that some groups are more at-risk for attrition than others. This refers to the unemployed (e.g. N. Watson & Wooden, 2009), minority groups (e.g. Aglipay, Wylie, & Jolly, 2015), ethnic minorities (e.g. Kapteyn, Michaud, Smith, & Van Soest, 2006;

Lipps, 2009), but also to old age (e.g. Hardy, Allore, & Studenski, 2009; Lillard & Panis, 1998; Lipps, 2009; McHenry et al., 2015). In addition, shocks and life changes can increase the risk for attrition, for example with regard to health or cognitive decline (Lugtig, 2014; N. Watson & Wooden, 2009), moving (Lillard & Panis, 1998), or instability in marital status (Fitzgerald, Gottschalk, & Moffitt, 1998; D. Watson, 2003; N. Watson & Wooden, 2009). In this respect, we specifically look at barriers introduced during the observation period and look at attrition as a process (Lugtig, 2014; Voorpostel & Lipps, 2011). Further, we base our conceptualization of oldest old on differentiating between working status, activity, and risk for diseases. Consequently, we will work with three age categories in this paper. First, the 50-64 year olds, who are in general characterized by still being part of the working force – the "occupationally active". The second age group ranges from 65-79 years, which can be characterized by being in retirement but still having a sociable, active life – the "young retirees". Third, the age group of the "oldest old", which refers to persons aged 80 and above and a life span generally marked by a higher risk for physical decline and diseases.

Data and methods

Data come from Wave 1 to Wave 7 of SHARE and cover the period from 2004 to 2017. In this period, 17 countries (Austria, Germany, Sweden, Netherlands until Wave 5, Spain, Italy, France, Denmark, Greece, Switzerland, Belgium, Israel, Czech Republic, Poland, Luxembourg, Slovenia, Estonia) took part in SHARE at least twice, which is a necessary prerequisite to be included in the analyses to measure changes. The average participation time is 3.9 (SD: 1.6). The multivariate analyses are based on a sample of 42.217 respondents.

To explore how oldest old behave in comparison the other age groups in SHARE, we distinguish between different participation patterns across Wave 1 to Wave 7 (see Voorpostel & Lipps, 2011 for a similar proceeding): "always in", "ever out", "dropped out", and "died". The "always in" category refers to respondents participating in every wave, in which the respondent is part of the eligible sample.[1] "Ever out" refers to an interrupted participation pattern with one or more waves of non-response, before a continued participation. "Dropped out" refers to attrition, i.e. leaving the panel permanently. Finally, "died" refers to a death during the observation period, which is reported or verified by mortality registers (where available).[2]

Based on this differentiation, we measure at-risk characteristics at the last interview before dropping out (for the first time or permanently).[3] New barriers/changes are measured by the difference between the last and the baseline interview before dropping out. We standardized this difference by the number of participations to consider different time lengths between the two measurement points. By this, we can take into account differential implications of short-term versus long-term changes. The covariates include survey data on socio-demographics (sex, age, education, born abroad, urban area, working status), living conditions (living alone, living in a nursing home, financial well-being), health and health behavior (chronic diseases, limitations in (instrumental) activities of daily living, unable to do physical measurements, visual/hearing problems, degree of moderate activities), cognition (word recall, fluency, need for clarifications, help with showcards), and social interaction and support (social activities, personal care-giving, proximity of children, availability of a proxy). We enhance the survey data by paradata measuring

[1] In SHARE, countries entered the study at different point in times. In addition, refreshment samples have been drawn on a regular basis, which leads to variation in time of entering the sample.
[2] Persons who dropped out and have an unknown vital status are still included in the" dropped out" category.
[3] In the case of „always in", we randomly selected an interview, excluding the baseline interview (see Voorpostel & Lipps, 2011).

characteristics of the specific interview (willingness to respond as observed by the interviewer, interview duration measured by keystrokes, item nonresponse).

Results

Who are the oldest old?

Table 1 shows the means of all used covariates by participation pattern for the oldest old. On a bivariate basis, respondents aged 80 years and older who dropped out (in contrast to participating in each wave; always in) are more likely to be older, born abroad, limited in their (instrumental) activities of daily living (ADL, IADL), unable to perform a grip strength measurement, and physically not very active. In addition, oldest olds who drop out more often live in urban areas, have poor eyesight, show limitations in cognitive functioning tests (recall of a number of displayed words, naming as many animals as possible), and ask for clarifications. In contrast, they are less likely to be low educated and socially active, and have fewer chronic diseases. Further, they less frequently have children living nearby or receive help from a proxy (most often a close relative). With the exception of chronic diseases, all these differences are highly significant on a 5% level.

Table 1: Descriptives of oldest old (80+) in SHARE by participation pattern, Wave 1-7

Characteristics	Always in	Ever out	Dropped out	Died
Socio-demographics				
Sex: female	0.67	0.67	0.67	0.61***
Age	84.85	83.74***	85.34***	86.90***
Education: low	0.45	0.61***	0.37***	0.65***
Education: high	0.13	0.09*	0.13	0.07***
Born abroad	0.10	0.09	0.14***	0.05***
Urban area	0.36	0.39	0.46***	0.40*
Working status: retired	0.84	0.79	0.84	0.78***
Living conditions				
Living alone	0.54	0.49	0.55	0.47***
Living in a nursing home	0.05	0.03	0.06	0.10***
Easy to make end meets	0.64	0.62	0.68	0.54***
Health & health behavior				
Chronic diseases	0.81	0.73**	0.78*	0.81
Limitations in ADL	0.31	0.26	0.38***	0.57***
Limitations in IADL	0.45	0.37**	0.52***	0.71***
Unable to do grip strength	0.07	0.11*	0.16***	0.31***
Poor eyesight	0.13	0.17	0.18***	0.29***
Poor hearing	0.11	0.10	0.13	0.21***
Little moderate activity	0.38	0.35	0.49***	0.71***
Cognition				
Limitations: word recall	0.70	0.72	0.73***	0.80***
Limitations: verbal fluency	0.85	0.87***	0.87***	0.89***
Asked for clarifications	0.12	0.16	0.18***	0.21***
Help needed reading showcards	0.24	0.23	0.26	0.41***
Social interaction & support				
Any social activity	0.21	0.24	0.13***	0.06***
Giving personal care	0.08	0.10	0.07	0.08
Children <25km away	0.73	0.62***	0.68***	0.77**
Proxy available	0.13	0.09**	0.08***	0.30***
Interview characteristics				
Unwilling to respond	0.17	0.26**	0.25***	0.36***
Income item nonresponse	0.09	0.11	0.18***	0.14***
Interview length	70.06	60.19***	64.65***	55.58***
N	5170	525	2137	3186

Data: SHARE Release 7-0-0 (weighted). Reference category for significance tests: always in.

How do oldest old respond?

Table 2 shows the participation patterns by age groups. Oldest old show a significantly lower share of participating in each wave ("always in") or having an interrupted participation pattern ("ever out") in contrast to occupationally active respondents. Not surprisingly, older olds die more often (29% deceased during the observation period in contrast to 2% of the occupationally active). However, with regard to attrition ("dropped out"), we do not observe significant differences across age groups, which might lead us to the conclusion that oldest olds do not require special attention of survey practitioners with regard to attrition.

Table 2: Participation pattern by age groups

Participation pattern	Age categories			Total
	Occupationally active (50-64 years)	Young retirees (65-79 years)	Oldest old (80+)	
Always in	63.06	64.11	46.92***	43939
Ever out	16.36	10.22***	4.76***	8695
Dropped out	18.39	17.67	19.40	13142
Died	2.19	8.00***	28.92***	6232
Total	31533	29457	11018	72008

Data: SHARE Release 7-0-0 (weighted). Reference category for significance tests: occupationally active.

However, there is not much we as survey practitioners can do in preventing death of old respondents. In contrast, addressing attrition, given that a respondent is still alive, opens room for (potentially successful) interventions. When we focus on respondents who are alive during the whole observation period (survivors) and exclude deceased respondents, the picture changes quite a bit and now shows a substantial difference in attrition rates. Table 3 shows that about 27% of the oldest old dropped out in contrast to 19% of the occupationally active (and young retirees as well). This is an indication that it might be worthwhile to put special focus on the oldest old. Further, keeping loyal respondents ("always in") seems to work equally well for the oldest old as for the other age groups. Young retirees are even more loyal than occupationally active respondents, perhaps due to having more time at home. Overall, however, the panel retention measures already applied in SHARE seem to work rather well to keep respondents in the panel, largely independent of their age.

Table 3: Participation pattern of survivors by age groups

Participation pattern	Age categories			Total
	Occupationally active (50-64 years)	Young retirees (65-79 years)	Oldest old (80+)	
Always in	64.47	69.68***	66.01	43939
Ever out	16.73	11.11***	6.70***	8695
Dropped out	18.80	19.21	27.29***	13142
Total	30843	27101	7832	65776

Data: SHARE Release 7-0-0 (weighted). Reference category for significance tests: occupationally active.

Which factors can help to explain attrition of oldest old (relative to younger respondents)?

To explore the role of certain indicators on potentially avoidable attrition, we run multivariate logistic regression models on dropout with respondents always taking part in SHARE as reference. Deceased persons are excluded from these analyses. Further, we interact all used covariates in the model with a dummy variable, indicating whether a respondent is 80+ or not at the observation before s/he dropped out. By this, we are able to investigate explicitly how oldest old differ in their attrition behavior from younger age groups.

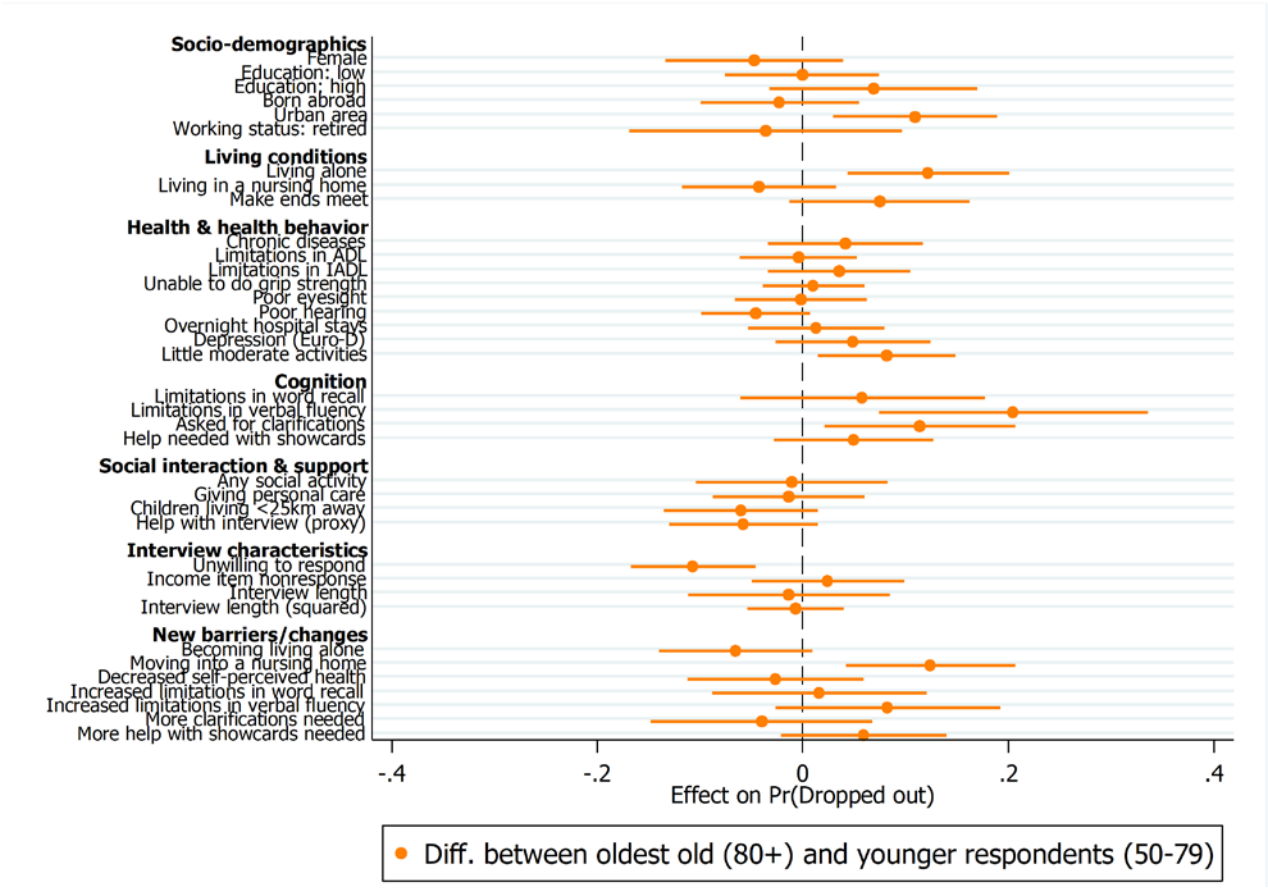

Figure 1: Indicators for dropout for oldest old

Figure 1 shows the difference between oldest old and younger respondents regarding the probability on dropping out in the following wave for the set of indicators presented before. A significant coefficient tells us that the respective indicator has a significant influence on drop out, unique for the oldest old. What we cannot see in the graph is that several indicators have an important overall effect on drop out. For example, social activity has a strong negative effect for young and old respondents, i.e. such activities prevent respondents (rather independent from age) from dropping out. In contrast, our analyses here show, which factors are important for keeping or loosing oldest olds and on which indicators we should focus, when we want to take an extra step for keeping the oldest old in SHARE. This depiction hence serves us as a starting point for potential further interventions in a next step.

Against this background, Figure 2, which only presents significant differences but otherwise is identical to Figure 1, reveals that an extra effort with respect to keeping the oldest old should be directed to those living alone and in urban areas. In addition, special attention should be given to the cognitive performance of oldest old. According to our findings, cognitive limitations as well as assistance from interviewers during the interview is significantly more important for oldest old compared to younger respondents regarding a subsequent dropout. Interestingly, unwillingness to respond has a significant negative effect. This means that the (un)willingness to respond of oldest old actually plays a significant less important role regarding attrition compared to younger respondents. This, in turn, can be interpreted in a way that more effort in this respect should be directed to younger but not to older respondents. In addition, new barriers or changes, such as cognitive decline do not add much to this picture. The amount of explained variance increases only slightly by 0.2 percent to 11 percent overall. However, the coefficient of moving into a nursing home is significantly larger for oldest old compared to younger age groups. Such a far-reaching change in the living situation is stronger correlated with attrition for oldest old than for younger respondents.

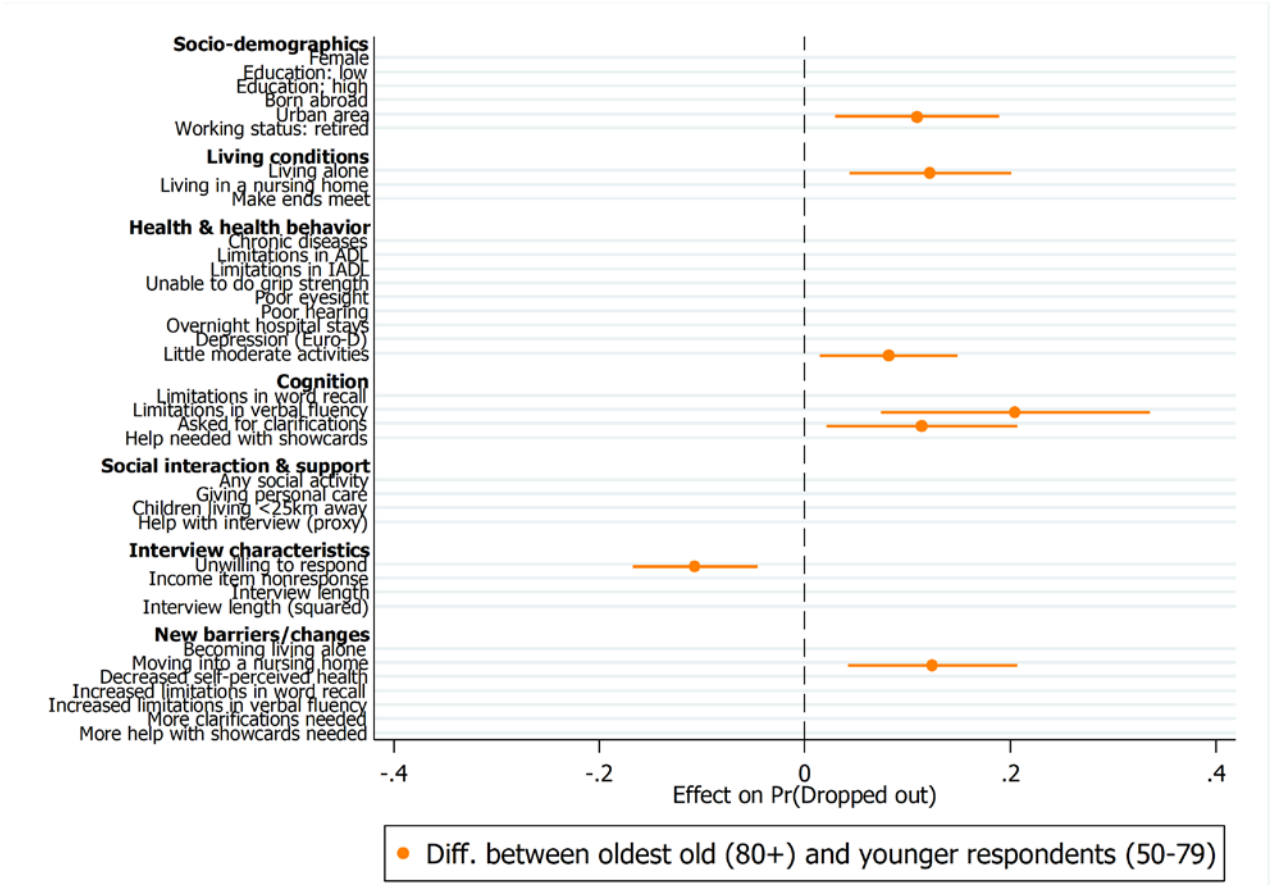

Figure 2: Indicators for dropout for oldest old (only significant coefficients)

Discussion points

The presented results regarding the first step of our study based on available SHARE data seem to support our expectation that focusing on oldest old might be worthwhile. When leaving natural mortality aside, oldest old drop out of SHARE significantly more often than younger respondents do. We further observe that there are several indicators, unique for oldest old, which can help to explain a subsequent dropout in the next wave. In this respect, our results show that oldest old are more likely to drop out when they live alone and/or in urban areas. This points out that the contact phase might be more important for oldest old than for younger respondents. In addition, an increased assistance during the interview might also be helpful as limitations in cognition play a larger role for oldest old compared to younger respondents regarding subsequent attrition.

Based on these findings, we already have some points of references where we can move on. In the second step of the project, we intend to go beyond the quantitative analysis of SHARE data and investigate what oldest old would need to continue participation by taking their individual perspective stronger into account. For this, we want to utilize qualitative or expert interviews that address the different possible causes for an individual dropout, such as health related factors or the burden of a long and complex interview. Additional suggestions or different ideas, also regarding the implementation of specific survey adaptations (e.g. experiences with proxy respondents/modules or a shortened interview version for particularly affected respondents) from the workshop participants are very much appreciated.

References

Aglipay, M., Wylie, J. L., & Jolly, A. M. (2015). Health research among hard-to-reach people: six degrees of sampling. *Canadian Medical Association Journal, 187*(15), 1145-1149. doi: 10.1503/cmaj.141076.

Andersen-Ranberg, K., Petersen, I., Robine, J.-M., & Christensen, K. (2005). Who are the oldest old? In A. Börsch-Supan, A. Brugiavini, H. Jürges, J. Mackenbach, J. Siegrist & G. Weber (Eds.), *Health, ageing and retirement in Europe - first results from the Survey of Health, Ageing and Retirement in Europe* (pp. 35-40). Mannheim: MEA.

Eurostat. (2019). Population structure and ageing Retrieved 10.08.2020, from https://ec.europa.eu/eurostat/statistics-explained/index.php/Population_structure_and_ageing#The_share_of_elderly_people_continues_to_increase

Fitzgerald, J., Gottschalk, P., & Moffitt, R. (1998). An analysis of sample attrition in panel data: The Michigan Panel Study of Income Dynamics. *Journal of Human Resources, 33*(2), 251-299. doi: 10.2307/146433.

Hardy, S. E., Allore, H., & Studenski, S. A. (2009). Missing data: A special challenge in aging research. *Journal of the American Geriatrics Society, 57*(4), 722-729. doi: 10.1111/j.1532-5415.2008.02168.x.

Kapteyn, A., Michaud, P.-C., Smith, J. P., & Van Soest, A. (2006). *Effects of attrition and non-response in the Health and Retirement Study*. IZA Discussion Papers, No. 2246. Bonn: IZA – Institute of Labor Economics.

Lillard, L. A., & Panis, C. W. A. (1998). Panel attrition from the Panel Study of Income Dynamics: Household income, marital status, and mortality. *Journal of Human Resources, 33*(2), 437-457. doi: 10.2307/146436.

Lipps, O. (2009). *Attrition of households and individuals in panel surveys*. SOEPpapers on Multidisciplinary Panel Data Research, No. 164. Berlin: DIW Berlin.

Lugtig, P. (2014). Panel Attrition: Separating stayers, fast attriters, gradual attriters, and lurkers. *Sociological Methods & Research, 43*(4), 699-723. doi: 10.1177/0049124113520305.

McHenry, J. C., Insel, K. C., Einstein, G. O., Vidrine, A. N., Koerner, K. M., & Morrow, D. G. (2015). Recruitment of older adults: Success may be in the details. *Gerontologist, 55*(5), 845-853. doi: 10.1093/geront/gns079.

Tourangeau, R. (2014). Defining hard-to-survey populations. In R. Tourangeau, B. Edwards, T. Johnson, K. Wolter & N. Bates (Eds.), *Hard-to-Survey Populations* (pp. 3-20). Cambridge: Cambridge University Press.

Voorpostel, M., & Lipps, O. (2011). Attrition in the Swiss Household Panel: Is change associated with drop-out? *Journal of Official Statistics, 27*(2), 301-318.

Watson, D. (2003). Sample attrition between Waves 1 and 5 in the European Community Household Panel. *European Sociological Review, 19*(4), 361-378. doi: 10.1093/esr/19.4.361.

Watson, N., & Wooden, M. (2009). Identifying factors affecting longitudinal survey response. In P. Lynn (Ed.), *Methodology of longitudinal surveys* (pp. 157-182). Chichester: John Wiley & Sons.