

# Dissatisfied with life or with being interviewed? Happiness and the motivation to participate in a survey

Adrian Chadi<sup>1</sup> 

## Abstract

People with little motivation to participate in surveys can affect empirical research when they abstain from but also when they actually participate in interviews. This paper investigates whether happiness data are susceptible to such measurement bias. Evidence from the German Socio-Economic Panel Study (SOEP) reveals a strong relationship between self-reported life satisfaction and several indicators of respondent motivation, such as subsequent panel attrition. One explanation for this finding is that respondents on the margin of participation truly have lower life satisfaction. Alternatively, their low motivation may be the cause for an underreporting of life satisfaction. To learn more about this, an instrumental variable approach identifies future panel quitters with low motivation by using the occurrence of interviewer attrition in the year after the interview. The results of this analysis suggest that self-reported life satisfaction declines because of low respondent motivation. A discussion of the implications for analyses of happiness data underscores the potential importance of respondent motivation regardless of the explanation for why interviewees with low motivation report lower life satisfaction.

## 1 Introduction

Information on people's subjective assessments of their lives provides enormous potential for socio-economic research. Such data attract the interest of both economists who consider subjective well-being as a proxy for utility and policy-makers who look for alternatives to standard economic indicators (see e.g. Fleurbaey 2009; Stiglitz et al. 2009; Frey and Stutzer 2012). Research based on such data has provided many important insights, as discussed in numerous books written by economists (e.g. Frey and Stutzer 2002; Layard 2005; Van Praag and Ferrer-i-Carbonell 2008; Frey 2008; Weimann et al. 2015). To give an example, the large amount of

---

✉ Adrian Chadi  
[adrian.chadi@uni-konstanz.de](mailto:adrian.chadi@uni-konstanz.de)

<sup>1</sup> Department of Economics, University of Konstanz, Universitätsstraße 10, 78457 Constance, Germany

evidence on unhappiness among unemployed individuals questions the economic notion of unemployment being a primarily voluntary decision (e.g. Clark and Oswald 1994; Kassenboehmer and Haisken-DeNew 2009; Chadi 2010). Other topics in this field continue to be debated, such as the impact of income (e.g. Easterlin 1995; Stutzer 2004), the effect of becoming older (e.g. Van Landeghem 2012; Schwandt 2016), or the role of gender (e.g. Stevenson and Wolfers 2009). In all these ongoing discussions, the validity of the subjective data is receiving increased attention. Among other things, the role of framing effects (see e.g. Angelini et al. 2017) and lack of interpersonal comparability when people assess their lives on scales (see e.g. Kapteyn et al. 2010, Angelini et al. 2014) are part of this research on data validity. Undoubtedly, survey phenomena such as higher satisfaction scores when persons are present during interviews (Conti and Pudney 2011; Chadi 2013a) or lower self-reported life satisfaction when respondents have more panel experience (Chadi 2013b; Van Landeghem 2014) can affect the empirical conclusions. Accordingly, researchers increasingly consider such survey factors in the analysis of well-being to improve the accuracy of their research.<sup>1</sup>

This paper aims to shed light on the little-considered, but potentially crucial, role of people's motivation to participate in surveys for measuring their well-being. As an important aspect, the discussion acknowledges the double role that respondent motivation may play. First, people with low motivation can affect conclusions of empirical research by not responding to the requests of an interviewer. This could lead to nonresponse bias if certain types of individuals avoid participation. Second, if low-motivation types do participate, they could still pose a threat to empirical research. By providing biased responses, they could induce measurement errors.

Researchers have looked mainly at the first part of this story by examining the characteristics of survey respondents on the verge of not participating. Available evidence comes primarily from research on panel respondents who do not participate in subsequent interviews, the results of which, however, are mixed concerning the link between self-reported satisfaction with life and the likelihood of remaining in the panel.<sup>2</sup> Another potential indicator for lacking motivation to take part in surveys could be the number of attempts interviewers need to contact interviewees.<sup>3</sup>

<sup>1</sup> One example is the debate on the link between happiness and age. See e.g. Frijters and Beaton (2012); Kassenboehmer and Haisken-DeNew (2012); Wunder et al. (2013) and Baetschmann (2014). Apart from the concern about data validity, another reason for researchers to investigate and to better understand survey factors lies in the potential for causal identification strategies. Factors such as the day of the interview or the interview mode have been used as instruments for well-being (Wunder and Heineck 2013). Similarly, there are increasing numbers of quasi-experimental investigations on people's well-being exploiting interview dates in survey data (e.g. Metcalfe et al. 2011; Chadi 2015; Goebel et al. 2015; Schueller 2016).

<sup>2</sup> See e.g. the data documentation on panel attrition in the SOEP (Kroh 2009, 2010, 2012).

<sup>3</sup> Only a few researchers so far use information on the number of contacts with potential respondents to discuss the role of difficult-to-reach survey participants in data collection. For instance, Kreuter et al. (2010) investigate data on benefit recipients and look at the role played by number of calls, and Heffetz and Reeves (2019) compare the responses of easy- and difficult-to-reach respondents in various governmental surveys. Heffetz and Rabin (2013) introduce the idea of using contact attempts to discuss potential nonresponse bias in happiness data.

After controlling for other reasons why some people are harder to reach than others, the number of contact attempts could be indicative of differences in a respondent's motivation to participate. Using several such indicators gathered in a large-scale panel survey, the primary goal of this paper is to provide a comprehensive discussion of the role of respondent motivation for measuring people's well-being. The first research question aims to establish the basic relationship between these two variables, asking: "Is reported life satisfaction related to respondent motivation?" This question concerning a potential empirical relationship is important because people with low motivation may refuse to participate and, thereby, differ from motivated respondents by simply not showing up in the data. If these nonrespondents have certain characteristics, such as on average lower life satisfaction compared to actual respondents, their lack of motivation can indirectly affect conclusions of empirical research via nonresponse bias.

The second research question refers to a potential direct effect of people's motivation to complete a survey on reported life satisfaction. Going beyond studying the empirical relationship between the two variables, the question here is: "Does respondent motivation affect reported life satisfaction?" The idea is that interviewees with low levels of motivation may respond differently to survey questions, compared to participants with high levels of motivation, because of dissatisfaction with the interview situation itself. Arguably, this could affect participants' response behavior, independent from their true levels of happiness. If so, reported scores may be too low, as a possible consequence, and would therefore be measured incorrectly. Such a direct effect of respondent motivation on measures of well-being could then plague empirical analyses because of measurement bias.

While the second research question picks up a novel idea in the context of measuring subjective well-being, both types of biases could be crucial for empirical research. Selective nonresponse can affect results, just as misreporting among participants can. The following plan of the paper describes the attempt to answer both research questions and deliver supplementary evidence, aimed at underlining the potential relevance of both nonresponse bias and measurement bias for research on people's well-being.

Section 2 of this paper illustrates the empirical framework for investigating the role of respondent motivation in collecting happiness data. To clarify the understanding of respondent motivation, the discussion here defines a motivated respondent as a person who reveals information willingly, does not change response behavior during the course of the questionnaire, has no aversion of any kind towards the interview itself and has sufficient intrinsic enjoyment in doing the interview to compensate for the time spent on it. To gather indicators that could reflect variation in respondent motivation, this investigation employs the paradata of the German Socio-Economic Panel (SOEP). The SOEP is a large and annually conducted panel survey whose life satisfaction data are widely used in the research field.

Section 3 addresses research question one and presents results from investigating the link between indicators of respondent motivation and life satisfaction. The first variable of interest from the SOEP paradata is the number of contact attempts by the interviewer to conduct the interview. Arguably, people's willingness to participate in a survey could be an explanatory factor for varying reachability, as measured by

contacts between interviewer and interviewee. The results from regression analyses indeed show a negative link between higher numbers of interviewer contacts and reported life satisfaction, but only after controlling for other determinants of people's reachability. Second, the book-keeping information of the data collecting agency is further exploited to gather another indicator for respondent motivation. Available interviewer assessments of respondents' willingness to participate are strongly connected to self-reported happiness, implying that low willingness goes along with low life satisfaction scores. Given that the interviewer contact effect diminishes when interviewer assessments are considered in the regression model, it appears that both indicator variables reflect differences in respondent motivation. In the third step, the empirical analysis identifies unwillingness among participants based on attrition in a subsequent survey year. Life satisfaction responses of individuals about to leave the panel in the next year are clearly reduced, which confirms the idea of a link between reported happiness and a third indicator of respondent motivation.

Section 4 addresses research question two and investigates whether respondent motivation has a direct effect on self-reported life satisfaction. An instrumental variable (IV) analysis exploits the panel phenomenon of interviewer attrition as an instrument for interviewee attrition. More specifically, the event of an interviewer's exit in the following panel year is used to explain why a current interviewee may also leave the panel in the same subsequent year due to lack of motivation. Worrying to survey organizers, but useful in the present research context, there is a well-known attrition problem when already unwilling people are asked to continue in the panel by a different person than before. Such a request is more easily rejected when there is a change in the person asking for the interview. This necessarily happens when, for example, an interviewer stops working for the data collecting agency. Arguably, the occurrence of being asked for an interview by a different person than before is irrelevant to people who are motivated to continue in the panel survey, but it does increase panel exits among those who already have low motivation to participate. These individuals are triggered by interviewer attrition, which works as a "last straw" to finally provoke the decision to refuse participation. As shown by its strong effect on subsequent respondent attrition, interviewer attrition in the year after someone's last participation is a relevant instrument. Assuming that there is no separate link between subsequent interviewer attrition and a respondent's life circumstances in the last year of participation, the interviewer exit also serves as a valid instrument. The results of the IV analysis provide evidence of a measurement bias in the subjective data, which may be caused by people's low motivation to participate in a survey. The existence of subgroup differences in the IV results, regarding gender and employment status, constitutes *prima-facie* evidence underlining the potential importance of the findings for empirical research based on survey data.

Section 5 offers a deeper discussion of the empirical evidence with the aim of outlining possible implications. The discussion makes clear that the link between motivation and life satisfaction is important for survey design decisions *per se*, independent of whether low motivation causes measurement bias, nonresponse bias, or both. As shown in an illustrative comparison of regional life satisfaction levels, inspired by the analysis in Heffetz and Rabin (2013), survey policies can affect the comparability of well-being measures across groups of survey respondents. This

demonstrates that the potential bias issues discussed in this paper matter, regardless of the direction of causality between motivation and happiness.

Section 6 concludes the paper. The key findings are briefly summarized and related to ongoing research on the validity of subjective survey data.

## 2 Data

The SOEP is Europe's longest-running representative and ongoing panel survey of households (see Wagner et al. 2007). Year by year, thousands of participants are questioned about their lives either directly by the interviewers of the data collecting agency or via questionnaires that people can fill out on their own. The main dependent variable in this paper is obtained at the end of each annual SOEP questionnaire, when interviewees are asked to assess their general life satisfaction on a scale ranging from 0 ("completely dissatisfied") to 10 ("completely satisfied"). As this question ("How satisfied are you with your life, all things considered?") is part of the questionnaire in every survey wave, the SOEP dataset is very useful for empirical research on the determinants of happiness and has therefore been used for this purpose numerous times over recent decades.<sup>4</sup>

The following paragraphs describe the datasets used in this paper to study the role of respondent motivation in the analysis of life satisfaction data. Table 1 shows summary statistics for all datasets. Each sample allows for the comprehensive investigation of one indicator of respondent motivation using a large number of observations.

The first variable for measuring differences in respondent motivation is the number of interviewer contacts. In their efforts to maximize response rates, interviewers attempt to reach all the households the agency has assigned to them and to conduct interviews with all adult household members. All contact attempts, whether successful or not, are documented by the agency, which collects information on the number of times an interviewer tries to reach a potential household. Paradata of this kind are provided to the research community as an add-on to the standard dataset made available by the SOEP and are part of the so-called "Brutto" files. The number of interviewer contacts is available for all waves, except for the initial survey year of 1984, and for most of the participating households.<sup>5</sup> As shown in Table 2, the available information on the interviewer contacts ranges from "one contact" interviews to "nine contacts or more" interviews. Many households are reached in the second

<sup>4</sup> Exemplary happiness studies based on SOEP data are from Frijters et al. (2004), Clark et al. (2008) and Luechinger et al. (2010). Note that, in line with most of the studies in the research field, the present paper treats the terms happiness, life satisfaction and (subjective) well-being as interchangeable.

<sup>5</sup> When there is contact via mail only, so that questionnaires are filled out without interviewer presence, there is usually no data available on the contact attempts. Thus, the survey mode of "self-written questionnaires by mail" has to be left out of the analysis, although motivation can be expected to be very low. As the SOEP offers the "mail" mode as a last opportunity for reluctant individuals, it is interesting to note that life satisfaction is very low in this group (Chadi 2012), which could be seen as indicative of a link between respondent motivation and reported happiness. For an overview on the interview modes analyzed here, see Table 1.

**Table 1** Descriptive information on the SOEP samples used in the analyses

| Variable                                   | Set 1<br>Mean | Set 2<br>Mean | Set 3<br>Mean | Set 4<br>Mean | Min | Max    |
|--------------------------------------------|---------------|---------------|---------------|---------------|-----|--------|
| Life satisfaction                          | 6.99          | 6.92          | 6.97          | 7.07          | 0   | 10     |
| Female                                     | 0.52          | 0.52          | 0.52          | 0.53          | 0   | 1      |
| Age                                        | 48.50         | 48.88         | 47.27         | 49.37         | 18  | 101    |
| Number of household members                | 2.74          | 2.74          | 2.81          | 2.72          | 1   | 13     |
| Living area                                | 102.51        | 104.14        | 100.46        | 97.54         | 6   | 938    |
| Equalized real income                      | 1804.2        | 1891.1        | 1777.4        | 1685.9        | 0   | 76,172 |
| Primary education                          | 0.19          | 0.19          | 0.21          | 0.24          | 0   | 1      |
| Secondary education                        | 0.62          | 0.62          | 0.62          | 0.62          | 0   | 1      |
| Tertiary education                         | 0.19          | 0.20          | 0.18          | 0.14          | 0   | 1      |
| Full-time employment                       | 0.41          | 0.40          | 0.42          | 0.39          | 0   | 1      |
| Regular part-time employment               | 0.10          | 0.10          | 0.10          | 0.09          | 0   | 1      |
| Marginal, irregular part-time employment   | 0.04          | 0.04          | 0.03          | 0.03          | 0   | 1      |
| Other forms of employment                  | 0.03          | 0.02          | 0.03          | 0.02          | 0   | 1      |
| Out of labor force                         | 0.43          | 0.43          | 0.42          | 0.47          | 0   | 1      |
| Unemployed                                 | 0.07          | 0.07          | 0.07          | 0.07          | 0   | 1      |
| Owner of dwelling                          | 0.51          | 0.53          | 0.49          | 0.48          | 0   | 1      |
| House in a good condition                  | 0.71          | 0.74          | 0.70          | 0.71          | 0   | 1      |
| Some renovation needed                     | 0.26          | 0.24          | 0.27          | 0.26          | 0   | 1      |
| Full renovation needed                     | 0.03          | 0.02          | 0.03          | 0.04          | 0   | 1      |
| Person needing care in household           | 0.04          | 0.04          | 0.04          | 0.05          | 0   | 1      |
| Moved recently                             | 0.10          | 0.09          | 0.10          | 0.09          | 0   | 1      |
| No children in household                   | 0.68          | 0.68          | 0.65          | 0.68          | 0   | 1      |
| Married                                    | 0.64          | 0.64          | 0.65          | 0.64          | 0   | 1      |
| Married but separated                      | 0.02          | 0.02          | 0.02          | 0.02          | 0   | 1      |
| Single                                     | 0.21          | 0.21          | 0.21          | 0.19          | 0   | 1      |
| Divorced                                   | 0.07          | 0.07          | 0.06          | 0.07          | 0   | 1      |
| Widowed                                    | 0.07          | 0.07          | 0.06          | 0.09          | 0   | 1      |
| Partnership                                | 0.80          | 0.80          | 0.80          | 0.77          | 0   | 1      |
| Degree of disability                       | 7.10          | 7.25          | 6.76          | –             | 0   | 100    |
| Nights in hospital                         | 1.72          | 1.66          | 1.77          | –             | 0   | 360    |
| Doctor visits                              | 2.63          | 2.54          | 2.64          | –             | 0   | 99     |
| Years spent in panel                       | 9.50          | 9.43          | 8.59          | 8.20          | 1   | 26     |
| Oral interview with paper and pencil       | 0.34          | 0.32          | 0.39          | 0.58          | 0   | 1      |
| Oral interview with computer assistance    | 0.27          | 0.30          | 0.22          | 0.29          | 0   | 1      |
| Self-written with interviewer presence     | 0.04          | 0.04          | 0.05          | 0.07          | 0   | 1      |
| Partly oral, partly self-written interview | 0.04          | 0.05          | 0.05          | 0.07          | 0   | 1      |
| Self-written without interviewer presence  | 0.30          | 0.30          | 0.30          | –             | 0   | 1      |
| Number of interviewer contacts             | 2.81          | 2.81          | –             | –             | 1   | 9      |
| Willingness to participate                 | –             | 1.29          | –             | –             | 1   | 4      |
| Last participation of the respondent       | –             | –             | 0.07          | 0.07          | 0   | 1      |
| Subjective health status                   | –             | –             | 3.39          | –             | 1   | 5      |
| Health satisfaction                        | –             | –             | 6.66          | –             | 0   | 10     |
| Last year of the interviewer               | –             | –             | –             | 0.07          | 0   | 1      |

Table 1 Continued

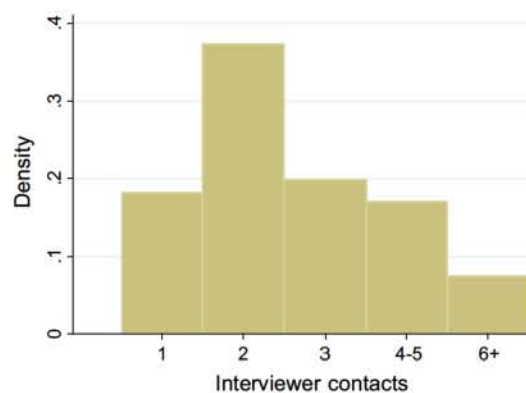
| Variable                             | Set 1   | Set 2  | Set 3   | Set 4   | Min | Max |
|--------------------------------------|---------|--------|---------|---------|-----|-----|
|                                      | Mean    | Mean   | Mean    | Mean    |     |     |
| Number of interviews per interviewer | –       | –      | –       | 66.37   | 1   | 256 |
| Number of observations               | 249,919 | 76,236 | 176,519 | 182,790 |     |     |

Figures are based on SOEP data from 1995 to 2011 (Set 1) in column 1, from 2002 to 2006 (Set 2) in column 2, from 1995 to 2006 (Set 3) in column 3, and from 1985 to 2009 (Set 4) in column 4

attempt, while a minority of cases require more than five contacts. In line with previous SOEP-based research on interviewer contacts (see Schraeppler et al. 2010), several potentially relevant factors as to why people can be “difficult to reach” are examined. The summary statistics of selected variables in Table 2 show that contact attempts decrease with age and increase with income, health, education, and the probability of being part of the labor force. While many of these factors may reflect positive aspects of the lives that hard-to-reach people lead, some of the people with higher degrees from the education system and well-paid jobs are probably very busy, which could drive contact numbers upward. This suggests that, for the purpose of revealing the role of respondent motivation in information on interviewer contacts, it is important to hold those factors constant in a regression analysis. Note that, in this respect, objective health information provides the preferred control variables, as subjective health assessments may be too closely related to subjective well-being. Limited availability of indicators for objective health in the SOEP restricts the first data sample (Set 1) to the period of 1995–2011, as can be seen in the first column of Table 1.

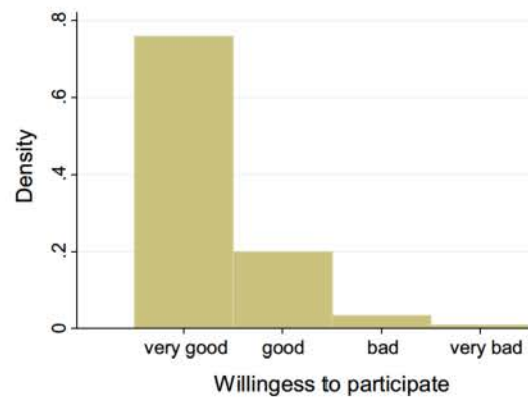
A second variable in the SOEP paradata captures the differences in respondent motivation more directly compared to interviewer contacts. For the time period from 2002 to 2006, the paradata contain information from interviewers who were asked to assess each interviewee’s willingness to participate. In contrast to the number of interviewer contacts, this variable is obtained on the individual level and has never before been used for research purposes. As shown in Table 3, there are four categories ranging from “very good” to “very bad” to allow for these assessments, which distinguish the characteristics of people being assessed as more or less willing to do the survey. Similar to the above pattern for interviewer contacts in Table 2, the people on the verge of non-participation are generally younger, on average. In contrast, however, low willingness does not clearly go along with positive features in people’s lives, like good health or high income. This suggests that using, or not using, control variables in a regression analysis may be less important. Due to the availability of this variable, the sample for this part of the analysis (Set 2) is restricted to five waves, as shown in the second column of Table 1.

Actual nonresponse in the future allows the establishment of a third indicator for respondent motivation. The idea here is that those individuals who provide evidence of their unwillingness by ultimately refusing to participate already had a lower level of motivation a year earlier. As shown in research on panel attrition and nonresponse in surveys, the most significant reason for people not participating is refusal (see e.g. Lipps 2009; Schraeppler et al. 2010). To identify panel exits in the SOEP, the two

**Table 2** Number of interviewer contacts and further descriptive information

| Interviewer contacts                           | 1                  | 2          | 3          | 4 or 5                         | 6 or more                                             |
|------------------------------------------------|--------------------|------------|------------|--------------------------------|-------------------------------------------------------|
| Observations by number of interviewer contacts | (1) 45,590         | (2) 93,435 | (3) 49,643 | (4) 27,674 (5) 14,922 = 42,596 | (6) 7220<br>(7) 4385<br>(8) 2618<br>(9) 4432 = 18,655 |
| Contact categories                             | 1                  | 2          | 3          | 4                              | 5                                                     |
| Selected variables                             | Means per category |            |            |                                |                                                       |
| Life satisfaction                              | 7.065              | 6.946      | 6.979      | 6.988                          | 6.997                                                 |
| Age                                            | 50.899             | 50.149     | 47.294     | 45.826                         | 43.652                                                |
| Household members                              | 2.567              | 2.662      | 2.831      | 2.926                          | 2.929                                                 |
| Equalized real income                          | 1736               | 1750       | 1816       | 1881                           | 2035                                                  |
| Primary education                              | 0.235              | 0.190      | 0.184      | 0.177                          | 0.162                                                 |
| Secondary education                            | 0.608              | 0.628      | 0.620      | 0.610                          | 0.612                                                 |
| Tertiary education                             | 0.157              | 0.182      | 0.196      | 0.213                          | 0.226                                                 |
| Out of labor force                             | 0.487              | 0.465      | 0.400      | 0.359                          | 0.299                                                 |
| Person needing care                            | 0.049              | 0.047      | 0.043      | 0.036                          | 0.032                                                 |
| Degree of disability                           | 8.174              | 7.855      | 6.652      | 5.880                          | 4.652                                                 |
| Nights in hospital                             | 2.000              | 1.843      | 1.659      | 1.428                          | 1.301                                                 |
| Doctor visits                                  | 2.724              | 2.711      | 2.577      | 2.480                          | 2.442                                                 |
| Years spent in panel                           | 10.550             | 9.944      | 9.050      | 8.672                          | 7.744                                                 |
| Oral interview                                 | 0.765              | 0.611      | 0.571      | 0.529                          | 0.538                                                 |
| Data loss due to missings                      | 0.134              | 0.122      | 0.143      | 0.150                          | 0.306                                                 |

The figures are based on SOEP data from 1995 to 2011 (Set 1, see Table 1). Data loss due to missings in the last row of the table reports the share of lost observations due to item nonresponse in the variables used when comparing the final sample (Set 1) to the available raw data from that same period 1995 to 2011

**Table 3** Willingness to participate and further descriptive information

| Willingness to participate           | 1 "very good"      | 2 "good" | 3 "bad" | 4 "very bad" |
|--------------------------------------|--------------------|----------|---------|--------------|
| Observations by level of willingness | 57,877             | 15,230   | 2509    | 620          |
| Selected variables                   | Means per category |          |         |              |
| Life satisfaction                    | 7.028              | 6.655    | 6.310   | 6.105        |
| Age                                  | 49.403             | 47.396   | 46.456  | 46.232       |
| Household members                    | 2.696              | 2.858    | 2.865   | 2.866        |
| Equalized real income                | 1917               | 1806     | 1834    | 1778         |
| Primary education                    | 0.171              | 0.215    | 0.226   | 0.240        |
| Secondary education                  | 0.623              | 0.621    | 0.607   | 0.619        |
| Tertiary education                   | 0.206              | 0.164    | 0.168   | 0.140        |
| Out of labor force                   | 0.435              | 0.422    | 0.411   | 0.431        |
| Person needing care                  | 0.039              | 0.053    | 0.075   | 0.076        |
| Degree of disability                 | 7.207              | 7.300    | 7.729   | 8.542        |
| Nights in hospital                   | 1.599              | 1.821    | 1.821   | 2.119        |
| Doctor visits                        | 2.554              | 2.537    | 2.337   | 2.394        |
| Years spent in panel                 | 9.539              | 9.176    | 8.557   | 8.939        |
| Oral interview                       | 0.642              | 0.561    | 0.521   | 0.447        |
| Interviewer contacts                 | 2.688              | 3.040    | 3.772   | 4.190        |
| Data loss due to missings            | 0.185              | 0.234    | 0.284   | 0.306        |

The figures are based on SOEP data from 2002 to 2006 (Set 2, see Table 1). Data loss due to missings in the last row of the table reports the share of lost observations due to item nonresponse in the variables used when comparing the final sample (Set 2) to the available raw data from that same period 2002 to 2006

waves of 2010 and 2011 are used solely for the purpose of having “future” information on people’s response behavior. Chances of a return are very low when someone refuses for two years in a row, so all respondents who stopped participating in the SOEP survey in 2009 or earlier are considered as permanent panel dropouts. For the purpose of examining not only the role of leaving in the year after but also in two,

three, and four years in the future, the data sample for this part of the analysis (Set 3) is restricted further and includes the 2006 wave as the last year of investigation (see Table 1). This corresponds to another necessary restriction connected to the availability of variables important for understanding the link between (low) willingness to participate and life satisfaction responses. According to previous research (e.g. Contoyannis et al. 2004; Jones et al. 2006), health is linked to the probability of attrition in longitudinal surveys. Panel quitters with poor health may also have a lower level of happiness, especially if they are in very bad circumstances and are about to die. Given that death, of course, also affects SOEP participants, the advice of Wagner et al. (2007) is to exploit the available “biography and life history” data for such research purposes. By doing so, Gerstorf et al. (2010) find strikingly lower life satisfaction in the time period leading up to death and a clear decline of life satisfaction as the distance-to-death time variable comes closer to zero. In the investigation here, future death indicators up to 5 years in advance of a person’s death are used to fully consider that point. This fits well with the decision to include only waves up to 2006 in the data sample (Set 3) for this part of the analysis.

Finally, a fourth data sample is used to analyze in greater depth the potential direction of causality between respondent motivation and reported life satisfaction and to thereby address the second research question. The availability of an instrument for subsequent interviewee attrition, as an indicator of respondent motivation, allows for an IV analysis, which is illustrated in detail below in Sect. 4. Column four in Table 1 illustrates this sample (Set 4), which covers as many SOEP waves as possible to ensure a strong instrument. The table provides information on additional variables relevant for this analysis, including the instrument of subsequent interviewer attrition. In line with the identification of subsequent exits by the respondents, the two waves of 2010 and 2011 are used to identify the subsequent attrition of interviewers. Thus, all interviewers who stopped doing interviews in 2009 or earlier are considered as “future” panel quitters in the last year they are observed in the data.

### 3 Regression analyses

#### 3.1 Empirical strategy

This subsection illustrates the econometric specification underlying the analyses aimed at addressing the first research question. A simple model is used to inspect the link between respondent  $i$ ’s motivation to do a survey ( $MOTIVATION_{i,t}$ ) and self-reported subjective well-being ( $SWB_{i,t}$ ), which are both captured in a given survey year  $t$ :

$$SWB_{i,t} = \alpha MOTIVATION_{i,t} + \beta X_{i,t} + u_{i,t}.$$

As respondent motivation is not directly observed, several indicators are used: number of interviewer contacts ( $CONTACTS_{i,t}$ ), the interviewer-assessed rating of the interviewee’s willingness to participate ( $RATING_{i,t}$ ), and an interviewee’s exit from the panel in a subsequent year ( $ATTRITION_{i,t+1}$ ). Subjective well-being

is captured by the interviewee's self-reported level of life satisfaction. To estimate model (1) and to consider a variety of individual-specific and potentially time-varying factors ( $X_{i,t}$ ), linear regression analyses are employed. Controlling for heterogeneity across individual observations may help in revealing empirical relationships that otherwise are not visible in the descriptive statistics.

Regarding methodology, linear regressions are a widely accepted way to deal with satisfaction scores in the SOEP and similar panel surveys (see Ferrer-i-Carbonell and Frijters 2004). To allow for a comparison of findings, it makes sense to follow the common procedure in this field, despite available alternatives to the ordinary least squares (OLS) estimator.<sup>6</sup> Given the availability of longitudinal data, many happiness researchers apply fixed-effects estimators to come closer to causal evidence on the determinants of individual satisfaction levels. Thereby, it is possible to control for unobserved but fixed individual heterogeneity, which, however, would imply omitting any fixed element in people's motivational attitude. Furthermore, it is hardly possible to fully address the question of causality between respondent motivation and self-reported life satisfaction in this way, given that both variables could also be related to time-variant unobserved heterogeneity. While this paper's identification strategy to shed light on causality relies on an instrumental-variable approach (see Sect. 4), simple regressions without consideration of fixed-effects are considered as the preferred method in the following.

### 3.2 Interviewer contacts and interviewer assessments

The number of interviewer contacts is the first variable from the SOEP paradata employed in the following analysis as an indicator of respondent motivation. Table 4 shows the results from regressions with different sets of control variables. This demonstrates how the relationship between interviewer contacts and reported life satisfaction changes as a result of considering the potential reasons for varying reachability, which, according to the above discussion of Table 2, could reflect rather positive aspects in people's lives.

The first specification in Table 4 without any covariate shows that, on average, there is no link between interviewer contacts and life satisfaction. The more the regression models control for factors relevant to people's life satisfaction, the more the effect of the contact number becomes negative. Comparing the first and second columns, controlling for employment status alone already reveals a significant link between life satisfaction and contact attempts. Using dummy variables for contact categories instead of a linear variable leads to the same finding. According to the last column, the gap in reported life satisfaction between those who are contacted in the first attempt and the hardest-to-reach interviewees amounts to approximately 0.2

---

<sup>6</sup> To tackle potential concerns regarding the ordinal character of the dependent variable, ordered probit models can be implemented to check the results' robustness. Furthermore, due to the hierarchical data structure with individual, household and interviewer levels, using mixed linear models is an option. Note that all further analyses which are discussed but not presented in the paper are available from the author upon request.

**Table 4** Interviewer contacts and life satisfaction responses

| Specification:                     | (1)                                      | (2)                  | (3)                  | (4)                  | (5)                  | (6)                  |
|------------------------------------|------------------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| Number of interviewer contacts     | −0.003<br>(0.002)                        | −0.015***<br>(0.002) | −0.019***<br>(0.002) | −0.028***<br>(0.002) | −0.028***<br>(0.002) |                      |
| Interviewer contacts: one          |                                          |                      |                      |                      |                      | Reference category   |
| Interviewer contacts: two          |                                          |                      |                      |                      |                      | −0.090***<br>(0.009) |
| Interviewer contacts: three        |                                          |                      |                      |                      |                      | −0.102***<br>(0.011) |
| Interviewer contacts: four or five |                                          |                      |                      |                      |                      | −0.136***<br>(0.011) |
| Interviewer contacts: six or more  |                                          |                      |                      |                      |                      | −0.197***<br>(0.014) |
| Employment status                  | Reference category: full-time employment |                      |                      |                      |                      |                      |
| Regular part-time employment       |                                          | 0.036***<br>(0.011)  | 0.055***<br>(0.011)  | 0.101***<br>(0.011)  | 0.051***<br>(0.012)  | 0.052***<br>(0.012)  |
| Marginal, irregular part-time      |                                          | −0.137***<br>(0.019) | −0.087***<br>(0.019) | 0.046**<br>(0.018)   | 0.063***<br>(0.018)  | 0.064***<br>(0.018)  |
| Other forms of employment          |                                          | 0.059***<br>(0.022)  | 0.228***<br>(0.022)  | 0.323***<br>(0.022)  | 0.286***<br>(0.023)  | 0.287***<br>(0.023)  |
| Out of labor force                 |                                          | −0.375***<br>(0.008) | −0.309***<br>(0.008) | −0.120***<br>(0.008) | 0.101***<br>(0.010)  | 0.102***<br>(0.010)  |
| Education                          | Reference category: secondary education  |                      |                      |                      |                      |                      |
| Primary education                  |                                          |                      | −0.164***<br>(0.010) | −0.070***<br>(0.010) | −0.093***<br>(0.009) | −0.094***<br>(0.009) |
| Tertiary education                 |                                          |                      | 0.307***<br>(0.009)  | 0.045***<br>(0.009)  | 0.115***<br>(0.009)  | 0.115***<br>(0.009)  |
| Log equalized real income          |                                          |                      |                      | 0.803***<br>(0.009)  | 0.543***<br>(0.009)  | 0.543***<br>(0.009)  |
| Additional controls                | No                                       | No                   | No                   | No                   | Yes                  | Yes                  |
| Health controls                    | No                                       | No                   | No                   | No                   | Yes                  | Yes                  |
| N (Set 1)                          | 249,919                                  | 249,919              | 249,919              | 249,919              | 249,919              | 249,919              |
| Adjusted R <sup>2</sup>            | 0.000                                    | 0.011                | 0.018                | 0.060                | 0.169                | 0.169                |

Life satisfaction is the dependent variable. Standard OLS regressions are used. Number of interviewer contacts is a linear variable ranging from 1 to 9. Additional controls are variables for unemployment, owner of dwelling, housing conditions, living area, household member in need of care, number of persons in household, no children in household, recent move, family status, partnership, age, gender, year and federal state. Health controls are variables for the degree of disability, doctor visits, and nights in hospital. Heteroscedasticity-robust standard errors are in parentheses

Levels of statistical significance are: \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$

on the scale. Importantly, the potentially negative factor of low respondent motivation, hidden in the information on interviewer contacts, is only revealed when other aspects are held constant.

Further analyses demonstrate this finding's robustness. The hierarchical data structure may play a role, as dependencies across observations within the same household or interviewer cluster can occur. So, while there is reason to consider empirical tools other than linear regressions with heteroscedasticity-robust standard errors, the negative link found appears to be robust to a whole variety of checks, whether these are carried out via clustering or hierarchical models. As soon as individual fixed effects are considered, the coefficient of the contacts variable becomes smaller, yet remains highly significant. The results of the application of panel fixed-effects regressions are shown in Appendix Table 9. This addresses the potential impact of fixed but unobserved personality traits on well-being outcomes and reveals that, for the same person, requiring more contact attempts goes along with lower self-reported happiness. Finally, the role of the interviewer contacts might also be related to other survey aspects. For instance, fewer interviewer contacts are needed for more experienced respondents who spent more years in the panel (see Table 2), which justifies the use of such additional controls. Yet, the coefficient does not change much when considering this factor or the other survey aspect of the interview mode, as can be seen in Appendix Table 9.

Table 5 presents regression results when considering the second measure for respondent motivation from the interviewers' assessments. The main finding is that the willingness to participate in the interview is strongly related to self-reported happiness, which can be seen in columns two to four. Accordingly, very high motivation is linked to higher reported life satisfaction, compared to other willingness categories, even if a large set of influencing factors is considered as controls. Note that in line with the above discussion of that variable in Sect. 2, this result does not depend on the use of the control variables. Table 3 shows that it appears already in the average values. Interestingly, the summary statistics in this table also show that the number of interviewer contacts is related to the level of willingness, as less willingness appears to go along with worse reachability. To avoid having two correlated indicators in the specification, Table 5 also presents, in the last column, results from models without the contact attempts variable. These results confirm the strong link between interviewer ratings of respondent motivation and reported life satisfaction.

Simultaneously using both potential indicators of respondent motivation in Table 5 allows the analysis to determine whether these variables measure the same aspect. This said, another finding is that the negative link between interviewer contacts and life satisfaction responses (shown in column one) becomes weaker when the more direct measure for respondent motivation is used in the regression model (column two). Supplementary results in Appendix Table 10 from applying fixed-effects regressions instead of simple OLS confirm this picture. This supports the interpretation of the negative effect of interviewer contacts as an indicator of low motivation among those who are harder to reach.

For several reasons, the exploitation of interviewer judgments about respondent motivation could be considered problematic. As shown in Table 3, there is a

**Table 5** Interviewee's willingness to participate and life satisfaction responses

| Specification:                 | (1)                           | (2)                  | (3)                  | (4)                  |
|--------------------------------|-------------------------------|----------------------|----------------------|----------------------|
| Number of interviewer contacts | −0.031***<br>(0.004)          | −0.017***<br>(0.004) | −0.012***<br>(0.004) |                      |
| Willingness to participate     | Reference category: very good |                      |                      |                      |
| Good                           |                               | −0.281***<br>(0.015) | −0.254***<br>(0.015) | −0.258***<br>(0.015) |
| Bad                            |                               | −0.544***<br>(0.037) | −0.517***<br>(0.037) | −0.528***<br>(0.037) |
| Very bad                       |                               | −0.710***<br>(0.077) | −0.651***<br>(0.078) | −0.667***<br>(0.078) |
| Standard controls              | Yes                           | Yes                  | Yes                  | Yes                  |
| Health controls                | Yes                           | Yes                  | Yes                  | Yes                  |
| Survey factors                 | No                            | No                   | Yes                  | Yes                  |
| N (Set 2)                      | 76,236                        | 76,236               | 76,236               | 76,236               |
| Adjusted R <sup>2</sup>        | 0.186                         | 0.193                | 0.207                | 0.207                |

Life satisfaction is the dependent variable. Standard OLS regressions are used. Number of interviewer contacts is a linear variable ranging from 1 to 9. Standard controls include variables for employment status, unemployment, education, income, owner of dwelling, housing conditions, living area, household member in need of care, number of persons in household, no children in household, recent move, family status, partnership, age, gender, year and federal state. Health controls are variables for the degree of disability, doctor visits, and nights in hospital. Survey factors include the year in panel and the interview mode variables. Heteroscedasticity-robust standard errors are in parentheses

Levels of statistical significance: \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$

clear link between the probability of an oral interview and the willingness indicator. Hence, the latter might be a proxy for the positivity bias associated with a face-to-face situation (see e.g. Conti and Pudney 2011). To address this potential issue, both Table 5 and Appendix Table 10 also show results for models that include survey factors (columns three and four); the effects linked to respondent motivation remain large and significant. Furthermore, the interviewer ratings might be positively biased. It could be that some interviewers tend to portray the interview and thus the result of their work rather positively. Not all cases of low willingness might be captured by such a subjective variable. In line with this concern, Table 3 shows that the distribution of interviewer ratings is heavily skewed. However, an alternative explanation for this might be that people with no motivation to do the interview may simply refuse, such that the probability of being included in the data at all goes down with lower respondent motivation. In support of the accuracy of interviewer ratings on respondent motivation, the quality of the data from those interviews that still took place is clearly worse among those interviewees assessed as less willing. Table 3 provides evidence for this, as lower willingness goes along with a larger loss of observation numbers (due to missing values) when comparing the final sample (Set 2) to the entirety of the available data from that period. As an alternative approach, it is also possible to check the potential role of bias in the interviewer ratings by adding interviewer fixed effects to the specifications in Table 5. The results

hardly change when doing so. In summary, considering all the possible problems with information based on interviewer ratings, it may be argued that it is even more impressive to obtain such strong evidence on the link between respondent motivation and self-reported happiness.

### 3.3 Subsequent interviewee attrition

The following part of the empirical analysis addresses research question one by employing a third indicator for people's possible lack of motivation. As argued in Sect. 2, the panel structure of the data allows identification and analysis of future nonrespondents, while they are still participating in their last year prior to their exits.

The first graphical illustration (a) in Fig. 1 demonstrates a clear relationship between life satisfaction and future interviewee attrition. There is a decline in reported happiness, which becomes increasingly obvious as time left in the panel heads toward zero. At the same time, there is a strong decline in people's willingness to participate (b), according to interviewer ratings. This underscores the idea of using future nonresponse as an indicator of low respondent motivation. Moreover, health seems highly relevant in this context, as it declines similarly (c, d). This aspect requires special consideration in the analysis because a participant being ill and leaving the panel due to that illness can reflect true unhappiness as a determinant of survey refusal. In contrast, the underreporting of life satisfaction before leaving the panel could be a measurement error caused by low respondent motivation. As a first check of these findings, the visual analysis of Fig. 1 can be repeated for a sample restricted to non-disabled individuals only. This shows a similar picture regarding the declines in health, respondent motivation, and life satisfaction in interviewees' final years of panel participation.

The regression output of Table 6 shows that the main insight from the graphical analysis is robust to the consideration of a broad set of control variables, including survey factors. To more closely inspect the strong link between willingness to participate and life satisfaction responses, the aspect of health is considered in multiple ways. This also concerns variables on subjective health. While such controls might actually capture some of the effect of interest and thus might lead to underestimated outcomes, the findings are even then remarkably robust. Concretely, the models include a variable for health satisfaction (measured on a 0–10 scale), dummy variables for subjective health measured in five categories (from very good to very bad), and objective health variables. The regressions also include time-to-death dummy variables, which indicate significant distress at the end of people's lives. Although the negative effects related to future nonresponse seem to become somewhat smaller as more aspects are held constant, still, the last years in the panel are linked throughout the analysis to significantly lower life satisfaction responses. Further analyses show the robustness of these findings when focusing on non-disabled individuals and when employing fixed-effects analyses.

In summary, there is strong evidence that allows answering research question one in the affirmative. There is an empirical link between respondent motivation and self-reported well-being. According to every available indicator for respondent

**Table 6** Leaving the panel and life satisfaction responses

| Specification:             | (1)                  | (2)                  | (3)                  | (4)                  | (5)                  | (6)                  |
|----------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| Last participation         | -0.236***<br>(0.017) | -0.291***<br>(0.017) | -0.257***<br>(0.017) | -0.197***<br>(0.017) | -0.152***<br>(0.016) | -0.137***<br>(0.016) |
| 2nd last participation     | -0.200***<br>(0.018) | -0.226***<br>(0.018) | -0.204***<br>(0.017) | -0.174***<br>(0.018) | -0.138***<br>(0.017) | -0.127***<br>(0.016) |
| 3rd last participation     | -0.148***<br>(0.018) | -0.163***<br>(0.018) | -0.143***<br>(0.018) | -0.115***<br>(0.018) | -0.093***<br>(0.017) | -0.089***<br>(0.017) |
| 4th last Participation     | -0.106***<br>(0.019) | -0.120***<br>(0.018) | -0.101***<br>(0.018) | -0.075***<br>(0.019) | -0.057***<br>(0.018) | -0.049***<br>(0.017) |
| Death in about one year    |                      |                      |                      | -0.508***<br>(0.061) | -0.153***<br>(0.054) | -0.088*<br>(0.053)   |
| Death in about two years   |                      |                      |                      | -0.223***<br>(0.056) | 0.060<br>(0.050)     | 0.107**<br>(0.048)   |
| Death in about three years |                      |                      |                      | -0.205***<br>(0.054) | 0.055<br>(0.048)     | 0.090**<br>(0.046)   |
| Death in about four years  |                      |                      |                      | -0.201***<br>(0.052) | -0.004<br>(0.047)    | 0.027<br>(0.046)     |
| Death in about five years  |                      |                      |                      | -0.188***<br>(0.048) | 0.017<br>(0.044)     | 0.054<br>(0.041)     |
| Standard controls          | Yes                  | Yes                  | Yes                  | Yes                  | Yes                  | Yes                  |
| Survey factors             | No                   | Yes                  | Yes                  | Yes                  | Yes                  | Yes                  |
| Health controls            | No                   | No                   | Yes                  | Yes                  | Yes                  | Yes                  |
| Subjective health          | No                   | No                   | No                   | No                   | Yes                  | Yes                  |
| Health satisfaction        | No                   | No                   | No                   | No                   | No                   | Yes                  |
| N (Set 3)                  | 176,519              | 176,519              | 176,519              | 176,519              | 176,519              | 176,519              |
| Adjusted R <sup>2</sup>    | 0.141                | 0.157                | 0.187                | 0.188                | 0.290                | 0.344                |

Life satisfaction is the dependent variable. Standard OLS regressions are used. See Table 5 for information on the control variables. Heteroscedasticity-robust standard errors are in parentheses

Levels of statistical significance: \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$

motivation, this link appears to be negative. Unclear at this point, however, is whether unhappier people are less motivated to participate in the survey (thereby potentially inducing nonresponse bias) or whether people's lack of motivation may affect their response behavior (thereby potentially inducing measurement bias), which is the second research question. To learn more about this, the last interviews of panel dropouts are looked at more closely in the following.

## 4 Instrumental variable analyses

### 4.1 Idea and empirical strategy

This part of the empirical analysis addresses the second research question concerning the possible effects of respondent motivation, approximated by future interviewee attrition, on self-reported well-being. To identify a group of future non-respondents who are selected not based on aspects related to their well-being but based solely on their low motivation, the following IV analysis exploits the event of future interviewer attrition. This incidence happens, for example, when interviewers quit their job, which usually is a coincidental event from the perspective of the interviewee. As is well-known among survey organizers, such interviewer dropouts can increase the probability that unmotivated interviewees will use this opportunity to exit the panel as well. Hence, it is possible to address the question concerning whether motivation affects reported subjective well-being ( $SWB_{i,t}$ ) by employing an expanded version of the econometric model (1). Thereby, future interviewer attrition ( $ATTRITION_{INT,t+1}$ ) serves as an instrument for future interviewee attrition ( $ATTRITION_{i,t+1}$ ), which continues to be the proxy for respondent motivation in the current survey year:

$$(2a) ATTRITION_{i,t+1} = \alpha^1 ATTRITION_{INT,t+1} + \beta^1 X_{i,t} + u_{i,t}^1.$$

$$(2b) SWB_{i,t} = \alpha^2 ATTRITION_{i,t+1} + \beta^2 X_{i,t} + u_{i,t}^2.$$

Several investigations into the determinants of panel attrition support this idea of an IV approach by providing evidence for the effect of interviewer (dis)continuity on response rates. Referring to the notion that positive rapport between interviewers and interviewees is an important aspect in response behavior, Zabel (1998) shows for U.S. panel data how the probability of interviewee attrition is reduced when keeping the same interviewer. Similarly, Laurie et al. (1999) observe higher response rates of participants in a British panel survey when the interviewer is the same over time. Arguing that rapport and the establishment of trust between interviewer and interviewee are important in this context, the authors refer to anecdotal evidence and earlier experiences with the SOEP, as reported in Pannenberg and Rendtel (1996), according to which survey respondents prefer having the same interviewer every year. Hill and Willis (2001) report a six percentage-point increase in response rates when maintaining the same interviewer in a U.S. panel survey. For data from multiple European panel surveys, Behr et al. (2005) show a positive effect of interviewer continuity on response rates, which is significant for most countries. For the same data, however, Nicoletti and Peracchi (2005) come to a less positive conclusion regarding the importance of avoiding interviewer changes. While they look at different outcomes compared to Behr et al. (2005), the results suggest that changing the interviewer may not necessarily harm response rates. This conclusion is in line with a more recent study by Lynn et al. (2014), who find that interviewer continuity reduces refusals, but not across all subgroups.

Summarizing these findings from the previous literature, there is reason to expect that the instrument of subsequent interviewer attrition satisfies the first prerequisite of being relevant for subsequent interviewee attrition. As studies have produced mixed evidence on effect strength, however, the instrumental power remains as an empirical question.

As the second prerequisite, the instrument should also be valid. Specifically, the exclusion restriction is that interviewer attrition in the following year is not directly linked to current interviewee life satisfaction. This appears likely, given that interviewers exit a panel because of their own personal life circumstances rather than those of the interviewees with whom they speak once a year. Nevertheless, the novel idea of using subsequent interviewer attrition as an instrument from the (relative) future certainly deserves a deeper discussion, which is provided next. Special attention is given to previous literature on interviewer discontinuity, which discusses potential threats to the assumption that a change of interviewer is an exogenous incidence from the perspective of the respondent.

## 4.2 Implementation and discussion of validity

The literature on interviewer discontinuity and panel attrition reports on endogenous decisions by field managers to re-assign interviewers after problematic interviews (see Zabel 1998; Hill and Willis 2001). Such intentional interviewer changes may be the consequence of interviewee behavior and thus may not be independent of it. To address this issue of the potentially endogenous replacements of one interviewer with another, this IV analysis focuses only on future interviewer changes that become a necessity due to future interviewer attrition. By identifying cases in which the interviewer is generally not doing any interviews in the years to come, the subsequent interviewer change is more likely an exogenous incidence from the perspective of the participant. In support of this, Hill and Willis (2001) point out that focusing on interviewer changes caused by interviewer exits is beneficial from the standpoint of empirical identification. Another step appears reasonable in light of these considerations, which is to exclude all cases of interviewer changes in the year of the interview from the dataset. This ensures that the data used in the IV analysis are not affected by potentially endogenous interviewer changes, which, in line with the literature, could also trigger interviewee attrition in the next year.<sup>7</sup> Nevertheless, two important concerns regarding validity of the instrument remain, both of which can be addressed empirically.

First, the occurrence of problem interviews could still be an issue for the validity of the instrument even after taking cases of endogenous interviewer re-assignments into account. One could argue that the subsequent interviewer exit may actually be the consequence of a problem interview with a respondent who also happens to have a particularly low level of life satisfaction. While this issue might be less

---

<sup>7</sup> A separate investigation of an expanded SOEP dataset reveals that the incidence of a current interviewer change is positively related to the probability of interviewee attrition in the subsequent year. Note that this incidence of a change in the person conducting the interview is also positively related to reported life satisfaction (Chadi 2013b).

important in the SOEP, as the number of interviews that one interviewer administers is typically very large (see Table 1), it makes sense to raise the annual number of interviews per interviewer in one of the various robustness checks presented in the results subsection below. Incrementally increasing this interviewee-interviewer ratio makes it less and less likely that one particular respondent could be responsible for causing the interviewer to leave the panel.

Second, some studies on the role of interviewer continuity in longitudinal surveys discuss potential endogeneity regarding regional effects. Campanelli and O'muircheartaigh (1999, 2002) argue that local economic conditions could be linked to interviewer attrition. Factors such as the regional unemployment rate may also affect the situation of the individual interviewee. To address this issue, it is possible to expand the SOEP data with additional control variables reflecting economic conditions at the regional level. For this purpose, annual data from the German Federal Statistical Office as well as the Federal Employment Agency are merged with the SOEP data at the federal state level.<sup>8</sup> As for all other control variables used in the IV analysis, the regional data reflect the situation in the year of the interview. Since the timing of control variables can be important for proper identification in IV models (see e.g. Deuchert and Huber 2017), it seems to be best practice to always use data from the year before the treatment incidence of interviewer attrition. Still, it is also possible to examine the role of regional conditions in the subsequent year, when the interviewer attrition actually takes place, as an additional robustness check.

For the following IV analysis, the dataset is expanded to the earlier waves of the SOEP data, starting in 1985, which is the first wave that includes interviewer identification numbers. This goes along with a small change in the set of additional control variables, as some health variables are not consistently available and thus have to be left out here (see Sect. 2). The idea is that sample size is in principle more important in the context of IV approaches than the availability of control variables. Another reason for this enlargement relates to a significant loss of data because the interview mode of a self-written questionnaire without interviewer presence has to be dropped from this final part of the analysis. Logically, interviewer exit is not a meaningful trigger for interviewee dropout when there has been no personal contact and thus no rapport between the respondent and the interviewer. A further restriction of the data is the exclusion of interviewer changes in the year of the interview, as discussed above. All these measures lead to the largest possible sample (Set 4), as illustrated above in Table 1.

### 4.3 Results and robustness checks

Table 7 presents the results from the IV analysis. The first stage results (panel A) show a significant effect of subsequent interviewer attrition on subsequent

---

<sup>8</sup> Note that the SOEP data used in this paper do not distinguish between observations in the two federal states of Saarland and Rhineland-Palatinate until the wave of 1999, making it necessary to edit all original regional data prior to merging the datasets.

**Table 7** Leaving the panel and life satisfaction responses (IV analyses)

|                                      |                                      |                       |                      |                     |                      |
|--------------------------------------|--------------------------------------|-----------------------|----------------------|---------------------|----------------------|
| A. First stage results               |                                      |                       |                      |                     |                      |
| Data restriction                     | All                                  | Out of<br>Labor force | Working              | Female              | Male                 |
| Dependent variable                   | Last participation of the respondent |                       |                      |                     |                      |
| Last year of the interviewer         | 0.093***<br>(0.003)                  | 0.107***<br>(0.005)   | 0.081***<br>(0.004)  | 0.097***<br>(0.005) | 0.088***<br>(0.005)  |
| Standard controls                    | Yes                                  | Yes                   | Yes                  | Yes                 | Yes                  |
| Survey factors                       | Yes                                  | Yes                   | Yes                  | Yes                 | Yes                  |
| N (Set 4)                            | 182,790                              | 86,025                | 96,765               | 96,360              | 86,430               |
| Adjusted R <sup>2</sup>              | 0.040                                | 0.039                 | 0.042                | 0.041               | 0.038                |
| B. Second stage results              |                                      |                       |                      |                     |                      |
| Dependent variable                   | Life satisfaction                    |                       |                      |                     |                      |
| Last participation of the respondent | -0.534***<br>(0.169)                 | -0.346<br>(0.234)     | -0.766***<br>(0.246) | -0.347<br>(0.227)   | -0.763***<br>(0.256) |
| Standard controls                    | Yes                                  | Yes                   | Yes                  | Yes                 | Yes                  |
| Survey factors                       | Yes                                  | Yes                   | Yes                  | Yes                 | Yes                  |
| N (Set 4)                            | 182,790                              | 86,025                | 96,765               | 96,360              | 86,430               |
| Adjusted R <sup>2</sup>              | 0.161                                | 0.189                 | 0.100                | 0.157               | 0.167                |
| Wald F statistic                     | 778.568                              | 426.074               | 353.765              | 449.622             | 329.915              |

See Table 5 for information on the control variables. Heteroscedasticity-robust standard errors are in parentheses. The instrument for a respondent's last year in the panel is the last year of the interviewer in the panel

Levels of statistical significance: \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$

interviewee attrition. This goes along with high F statistics and the observation that the instrument is relevant and strong. The main result of the second stage (column one, panel B) is a significantly negative effect on reported life satisfaction when last-time interviewees are randomly identified as persons with low motivation to continue participating.

Numerous checks help to assess the robustness of this finding, the first of which refers to the discussion in the previous subsection on the potential role of regional variations in economic conditions. Appendix Table 11 shows the results from re-running the main specification in column one of Table 7 with additional control variables at the state level. The findings are unaffected when considering unemployment rates, GDP per capita, and gross wages per employee from the year of the interview. A further check based on regional data from the subsequent year, that is, the year of interviewer attrition, also fails to lead to any other result.

Second, Appendix Table 12 picks up the point that interviewers usually deal with dozens of households per year. As argued above, it is unlikely that an interviewer's decision to quit might somehow be affected by a single interviewee and a possible problem interview. Nevertheless, to tackle such concerns, further analyses use data limited to interviewers with at least five interviews per year (column two).

Despite an increasing loss of observations, the finding holds when this minimum requirement is raised even further to 10, 15, or 20 interviews (columns three to five). Further analyses also support the idea that a subsequent interviewer exit is not the consequence of a problematic interview with a particularly unwilling respondent. In fact, in additional estimations with subsequent interviewer attrition as the dependent variable, there is no effect of a potential problem interview, independent of its definition. This is true for interviews conducted with interviewees whose interviewer-assessed willingness to participate was very low and for those who reported very low life satisfaction. The same non-finding emerges when analyzing spousal death as a particularly detrimental incidence, triggering both low motivation to participate and low life satisfaction of the affected respondent.

A third check addresses the potential role of rapport between the interviewer and respondent. Accordingly, interviewer exits may not matter much to interviewees in their first interviews who have not yet become familiar with the person asking for the interview. This suggests using a sample that requires people to have participated in the panel for a certain period of time. The results in panel B (respectively, panel C) of Appendix Table 12 show an increase in the magnitude of the measured effect, compared to panel A, when first-year (respectively first-year and second-year) observations are dropped from the dataset.

Fourth, the possible role of interviewee health is considered in some additional checks. Recall that objective health data are not consistently available in the SOEP throughout the entire period of investigation. Thus, using information on disability for the sake of focusing solely on non-disabled individuals goes along with a large loss of data, but the finding remains the same. Even if health satisfaction (which is consistently available) is added as a control variable to the analysis of Table 7 to check the role of interviewee health based on the same dataset, it does not alter the finding.

Fifth, some final checks address the hypothetical concern that interviewers who drop out in the future may affect their respondents' life satisfaction in the interviewer's last year. While reduced-form estimations show the expected significant effect of future interviewer dropout on respondents' self-reported life satisfaction, this negative effect becomes insignificant when subsequent interviewee dropout is included in the model. This suggests that the only channel through which subsequent interviewer attrition matters for current life satisfaction is low respondent motivation reflected by subsequent interviewee attrition, while future interviewer attrition itself has no separate effect on self-reported life satisfaction. In another check, the effect of subsequent interviewer attrition is also insignificant when the data are restricted to those respondents who remain in the panel, thereby excluding subsequent interviewee exits from the dataset. This again confirms that there is no separate effect in reported life satisfaction triggered by future interviewer dropout.

In summary, interviewer attrition appears to be irrelevant for interviewees in general, while it seems to be an effective trigger of panel exits among those respondents with low motivation. This group of future nonrespondents with low motivation to continue participating reports strongly reduced life satisfaction before leaving the panel, in line with the notion of measurement bias. Supported by numerous

robustness checks, the results from the IV analysis suggest answering research question two in the affirmative, according to which respondent motivation may indeed affect reported life satisfaction.

Columns two to five in Table 7 complement the IV analysis by showing results from an inspection of effect heterogeneity. This serves two purposes, as subgroup differences in IV results may provide insights on the nature of the compliers responding to the treatment and on potential implications for research on happiness determinants. Columns two and three of Table 7 show substantial differences in the effect of low respondent motivation on self-reported life satisfaction when people out of the labor force are separated from those who in one way or another are working (i.e. full-time, part-time, vocational training, and other forms of employment). For both subgroups the instrument remains strong, as F statistics are routinely above the common thresholds that are used for checking instrumental power. However, only for the subgroup of working people is the second-stage effect statistically significant. One interpretation is that less busy individuals with generally more available time also take the chance to leave a panel survey when their former interviewer has left the panel, but the negative bias in reported happiness scores is more prevalent among those with a busy time schedule due to being part of the labor force.<sup>9</sup> Another subgroup differentiation considers potential gender differences. The final results in columns four and five of Table 7 suggest that males drive the negative low-motivation effect in happiness scores. Considering the fact that females in the SOEP data are comparatively happy, this finding could offer a new perspective on the topic of interpreting gender differences in subjective well-being.

## 5 Discussion

### 5.1 Interpretation of main results

This section provides a discussion of the findings and their relevance for users of subjective data. The subgroup analysis of the last section suggests that employed and hence probably busy individuals are more likely to report lower life satisfaction because of less motivation to be interviewed. The argument of Hill and Willis (2001), who discuss the opportunity costs of participating in a survey, fits in well with these considerations, as busier people are likely to be more skeptical toward investing their time in surveys. Thus, regarding research on the life satisfaction gap between employed and unemployed people, this study suggests that the happiness premium of having a job may be underestimated when employees underreport their true state of well-being.

---

<sup>9</sup> The idea that busy, but not necessarily unhappy, types of people drive the findings receives additional support in further subgroup analyses. While there is no effect heterogeneity across income levels, the effects are stronger for younger people than for people above the median age. Furthermore, the effects are driven by married people living together with their spouses as well as by respondents with at least one child in the household.

A potential caveat of the IV analysis emerges out of these considerations. The crucial assumption is that the instrument identifies low-motivation interviewees independent of their actual level of life satisfaction. One could argue that the people on the margin of participation not only report lower life satisfaction, but that they are actually unhappy with their lives. This idea is hard to disprove though there is also no evidence suggesting that only the unhappy among the unmotivated exploit the occurrence of an interviewer exit to also leave the panel. Regardless of the potential direct effect of respondent motivation on self-reported life satisfaction, it is important to note that the evidence for such an effect does not reject the alternative interpretation, according to which unhappy people may be underrepresented in panel surveys. In that case, the interpretation would be that there is a threat to empirical research in the form of nonresponse bias, rather than in the form of measurement bias by misreporting happiness. Since instruments always identify a local effect, the conclusion here is that both explanations for the link between motivation and happiness have to be taken seriously.

Notably, the existence of a link, as investigated in Sect. 3, is in itself important, independent of the direction of causality between life satisfaction and respondent motivation. Survey design choices may affect the estimates and thus the comparability of results, which has been shown by Heffetz and Rabin (2013) for data on people's happiness. They employ information on contact attempts as a measure for respondent reachability. By including or excluding difficult-to-reach respondents who were almost not part of the data collection process, it is possible to simulate hypothetical situations that could result from more or less restrictive survey policies regarding the number of contact attempts. The findings in this paper suggest that simply having more people with low willingness in a survey could reduce overall happiness outcomes, regardless of whether those people are truly unhappy or whether they just self-report more negatively. To empirically illustrate this, the following analysis shows results from a comparison of life satisfaction in different regions in Germany using the available SOEP data.

## 5.2 Illustrative example

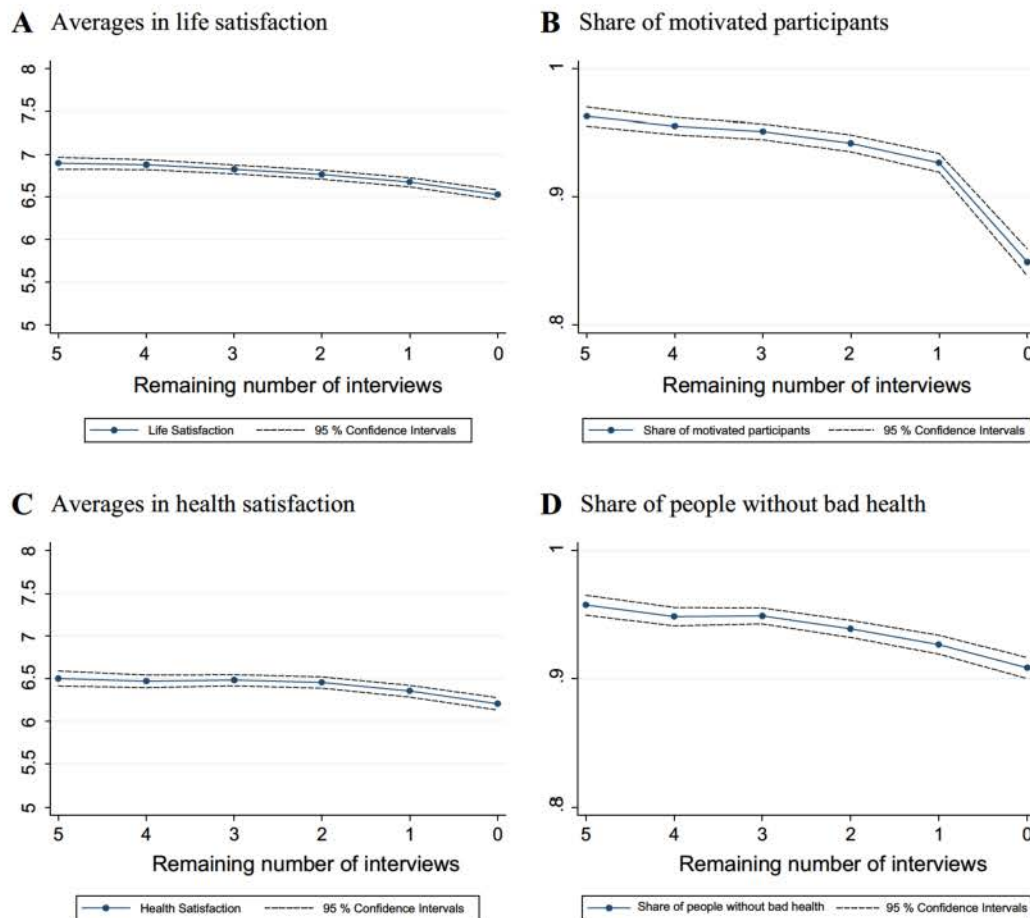
To simulate a hypothetical survey policy that is more restrictive than the SOEP regarding collecting happiness data, the number of contact attempts is limited in the following analysis. In a second step, interviewer assessments of interviewees' willingness to participate are used for simulating a survey design decision, according to which data collection is limited to those interviews for which interviewers have assessed respondent motivation as very good. Third, the survey policy of maintaining the same interviewer can be simulated by focusing on data without interviewer changes. The SOEP wave of 2006 is the most recent wave that includes all the variables needed for this analysis.

Table 8 shows the ranking of Germany's federal states according to the average happiness of people living in each region in 2006. The results from analyzing the data without any further restrictions puts Hamburg at the top with a score of approximately 7.4 (column one), outpacing all other states, as shown by the ranking in

**Table 8** Comparison of average life satisfaction in 2006 across federal states

| Region                     | Mean                  | Mean                  | Mean                  | Mean                  | Mean                  |
|----------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| Hamburg                    | [1] 7.411<br>(0.109)  | [3] 7.259<br>(0.133)  | [1] 7.486<br>(0.103)  | [1] 7.409<br>(0.111)  | [4] 7.366<br>(0.127)  |
| Bremen                     | [2] 7.265<br>(0.196)  | [2] 7.309<br>(0.201)  | [3] 7.356<br>(0.204)  | [2] 7.330<br>(0.204)  | [1] 7.481<br>(0.210)  |
| Schleswig–Holstein         | [3] 7.225<br>(0.080)  | [1] 7.386<br>(0.086)  | [2] 7.382<br>(0.083)  | [4] 7.125<br>(0.096)  | [2] 7.442<br>(0.106)  |
| Lower Saxony               | [4] 7.159<br>(0.049)  | [4] 7.229<br>(0.056)  | [4] 7.302<br>(0.051)  | [3] 7.189<br>(0.050)  | [3] 7.374<br>(0.060)  |
| Hesse                      | [5] 7.043<br>(0.058)  | [7] 7.015<br>(0.072)  | [6] 7.210<br>(0.067)  | [6] 7.022<br>(0.059)  | [7] 7.141<br>(0.084)  |
| Bavaria                    | [6] 7.039<br>(0.040)  | [8] 7.015<br>(0.048)  | [9] 7.120<br>(0.044)  | [5] 7.044<br>(0.044)  | [8] 7.089<br>(0.054)  |
| Rhineland-Palatinate       | [7] 7.038<br>(0.067)  | [6] 7.034<br>(0.072)  | [5] 7.275<br>(0.072)  | [8] 6.977<br>(0.071)  | [6] 7.196<br>(0.087)  |
| North Rhine-Westphalia     | [8] 7.000<br>(0.032)  | [5] 7.060<br>(0.037)  | [8] 7.154<br>(0.035)  | [7] 7.002<br>(0.032)  | [5] 7.211<br>(0.039)  |
| Saarland                   | [9] 6.995<br>(0.121)  | [10] 6.901<br>(0.136) | [7] 7.169<br>(0.116)  | [10] 6.799<br>(0.147) | [10] 6.982<br>(0.154) |
| Baden-Wuerttemberg         | [10] 6.976<br>(0.040) | [9] 6.932<br>(0.045)  | [10] 7.060<br>(0.045) | [9] 6.977<br>(0.042)  | [9] 7.029<br>(0.053)  |
| Mecklenburg-Vorpomm.       | [11] 6.706<br>(0.087) | [12] 6.621<br>(0.100) | [11] 6.794<br>(0.089) | [12] 6.711<br>(0.087) | [12] 6.714<br>(0.100) |
| Berlin                     | [12] 6.689<br>(0.083) | [11] 6.682<br>(0.094) | [12] 6.704<br>(0.086) | [11] 6.719<br>(0.082) | [11] 6.716<br>(0.096) |
| Saxony                     | [13] 6.475<br>(0.055) | [13] 6.441<br>(0.062) | [13] 6.603<br>(0.062) | [13] 6.467<br>(0.058) | [13] 6.588<br>(0.074) |
| Saxony-Anhalt              | [14] 6.348<br>(0.070) | [14] 6.308<br>(0.097) | [15] 6.345<br>(0.078) | [14] 6.328<br>(0.075) | [14] 6.236<br>(0.118) |
| Brandenburg                | [15] 6.278<br>(0.072) | [15] 6.211<br>(0.082) | [14] 6.382<br>(0.077) | [15] 6.266<br>(0.075) | [16] 6.188<br>(0.090) |
| Thuringia                  | [16] 6.223<br>(0.070) | [10] 6.129<br>(0.077) | [16] 6.297<br>(0.094) | [16] 6.220<br>(0.071) | [15] 6.231<br>(0.104) |
| Germany (all 16 states)    | 6.885<br>(0.015)      | 6.884<br>(0.017)      | 7.006<br>(0.016)      | 6.876<br>(0.015)      | 6.992<br>(0.019)      |
| Data restrictions:         |                       | (1)                   | (2)                   | (3)                   | (1–3)                 |
| Interviewer contacts       |                       | ≤3                    | –                     | –                     | ≤3                    |
| Willingness to participate |                       | –                     | 1                     | –                     | 1                     |
| Change of interviewer      |                       | –                     | –                     | No                    | No                    |
| Number of observations     | 14,338                | 10,676                | 11,131                | 13,171                | 7,827                 |

Averages in life satisfaction are based on SOEP data from 2006 (using Set 2, see Table 1). Life satisfaction is measured on a scale ranging from 0 (“completely dissatisfied”) to 10 (“completely satisfied”). Ranks based on comparisons of regional life satisfaction averages are in brackets. Standard errors are in parentheses. Data restriction (1) excludes data collected after more than three interviewer contacts. Data restriction (2) restricts the data to interviews with an interviewer assessment of “very good” willingness to participate. Data restriction (3) excludes data from interviews with a change of interviewer



**Fig. 1** The last years in the panel. *Notes:* The illustrations are based on SOEP data from 1995 to 2006 (Set 3, see Table 1). Life and health satisfaction are both measured on a scale ranging from 0 (“completely dissatisfied”) to 10 (“completely satisfied”). Motivated participants are those who are considered as having either “good” or “very good” willingness to participate in the survey, according to interviewers’ judgments. Bad health is a binary variable that is derived from a subjective health variable with five categories

column one. Excluding data gathered after four or more interviewer contacts moves Schleswig–Holstein from the number three spot to the top (column two). Focusing only on interviews with respondents assessed as very willing increases most of the average scores (column three), thereby making Germany in total appear happier than in the data without this restriction. Bavaria, for example, does not benefit from this increase as much as other states and falls from position six to nine. As another hypothetical survey policy, excluding interviewer changes also modifies average happiness scores, implying, for example, that Schleswig–Holstein falls down to the fourth position (column four). Implementing all survey policies simultaneously changes the ranking again with Bremen becoming the new number one with the highest mean happiness (column five). Meanwhile, the state of Thuringia leaves the bottom of the pile. Whether the people of Brandenburg were the unhappiest in Germany during

2006, however, is a question that the results of Table 8 certainly cannot prove or disprove. Adding to the possible implications of country-wide survey policies, imagine, for the sake of the argument, how the results might look if certain survey design decisions were implemented in one state but not in another. The evidence shows that survey design choices can affect cross-group comparisons within a survey but also across surveys. The analysis also demonstrates that it does not matter much whether those excluded from the analysis are truly unhappy or whether they just misreport their true state of happiness. In both cases, the results in such group comparisons can change easily.

## 6 Conclusion

The first empirical finding of this paper is a strong link between people's self-reported life satisfaction and several indicators reflecting their motivation to participate in a survey. By controlling for reasons why some people are harder to reach than others, a significantly negative relationship between interviewer contact attempts and happiness scores appears in the data. Interviewer-assessed interviewee unwillingness and subsequent interviewee attrition, as further proxies for low respondent motivation, reveal the same finding, even without using control variables. This robust link between respondent motivation and reported happiness gives rise to the second research question regarding causality. By exploiting the incidence of interviewer attrition, the panel data offer an opportunity to implement an IV approach. Accordingly, for some of the most unmotivated participants, the exit of the interviewer is the proverbial straw that breaks the camel's back, triggering the attrition of respondents with particularly low levels of motivation. As the second finding of the paper, the IV analysis shows a significantly negative effect in reported life satisfaction when using interviewer attrition in the following year to identify respondents with low motivation in the current survey year. This suggests a measurement error in life satisfaction data, caused by low motivation of individuals who report being dissatisfied with their lives, but, possibly, they are just dissatisfied with being interviewed. However, it is important to stress that evidence in favor of such measurement bias does not reject the alternative explanation for the empirical link revealed in the first part of the analysis, according to which the unmotivated people actually are less happy with their lives. If those unhappy people with low motivation are less likely to participate in surveys, this could lead to nonresponse bias.

In discussing the implications, it becomes clear that the link between respondent motivation and reported life satisfaction is important for surveys, regardless of the direction of causality. A regional comparison of life satisfaction averages shows that survey design choices can affect the outcomes of analyses on people's happiness by changing the number of participants with potentially low motivation. Thus, for two exemplary panels, one with easy policies and the other with a tough policy for obtaining as much data as possible, there is a problem with the comparability of the well-being data. Survey organizers can affect, knowingly or unknowingly, via their design choices, what happiness measures reveal about a region, a cultural group, or

whatever else. This must be considered when comparing data from different surveys, and it should encourage both survey organizers and data users to focus more on survey design choices when drawing conclusions from analyzing happiness data.

Several avenues for further research emerge out of these considerations. While this paper focuses on one direction of causality to examine the link between respondent motivation and reported well-being, a question that the analysis here does not fully address concerns the other direction of causality. This relates to the possible role of (un)happiness as a determinant of (low) respondent motivation and possible consequences, such as panel exits. Using alternative approaches for identification, researchers may wish to study the issue of sample selection more deeply. One option for determining what is missed in the data is to compare an attrition-plagued sample with some type of “gold standard” sample that represents how the data should look in the absence of attrition. Researchers typically use administrative data for that purpose and compare it to survey data, which is feasible for objective information, such as income (e.g. Meyer and Sullivan 2003; Abowd and Stinson 2013; Golsteyn and Hirsch 2019), but not for subjective survey data. An alternative approach to analyzing attrition bias in panels relies on using refreshment samples as the gold standard, which could be an option for subjective data (see Chadi 2016).

The determination of why respondent motivation may affect self-reported happiness is another question that may be useful to better understand the second major finding of this study. Some discussions in the literature may help in this respect. For instance, Olson (2013) points out that people who are “eventually convinced to participate” may show a low commitment to answering questions accurately, which suggests that the act of persuading individuals to do a survey may come at the price of measurement error. Other researchers discuss the significance of maintaining a pleasant atmosphere (e.g. Nederhof 1987; Pickery et al. 2001) or even recommend keeping respondents “happy” when being interviewed (Olsen 2005). In this context, one could ask whether, in particular, participants of lengthy household surveys may experience the interview as a burden (see Meyer et al. 2015). Since employed people in particular seem to underreport their life satisfaction, as a possible result of being busier than other respondents, it appears promising to study subgroup differences in potential reporting errors using other survey datasets with longer or shorter questionnaires.

Finally, ongoing research on the validity of happiness data attempts to assess the economic value of this type of information. From an economic perspective, the determinants of happiness may not be considered relevant if there is no relationship to actual choices and behavior. Various efforts in this direction have been conducted (e.g. Benjamin et al. 2012, 2014), the evidence from which, however, often appears as mixed. The findings in this paper may contribute to this discussion by pointing out the important but little recognized role of respondent motivation in collecting happiness data. Arguably, if the data are not accurately obtained and suffer from measurement and selection issues, the validation becomes much more difficult, even if happiness does indeed relate to people’s choices and behavior.

**Acknowledgements** I am grateful to Peter Krause and Martin Kroh for very helpful advice. For comments and discussions, I thank Daniel Arnold, Olivier Bargain, Ori Heffetz, Martin Huber, Joachim

Winter, Jinan Zeidan, two anonymous referees as well as participants of the International SOEP User Conference, Berlin 2012; the CISEPS International Conference on Public Happiness, Rome 2013; the TIBER Conference, Tilburg 2013; the IAAEU seminar, Trier 2013; the Statistical Week of the German Statistical Society, Berlin 2013; the LAGV Public Economics Conference, Aix-en-Provence 2014; the Quality of Life Conference, Berlin 2014; and the Annual Meetings of the International Association for Applied Econometrics, London 2014; the Austrian Economic Association, Vienna 2014; and the German Economic Association, Hamburg 2014.

## Appendix

See Tables 9, 10, 11, 12.

**Table 9** Interviewer contacts and life satisfaction responses

| Method                         | Standard regressions |                      |                      | Fixed effects regressions |                      |                      |
|--------------------------------|----------------------|----------------------|----------------------|---------------------------|----------------------|----------------------|
|                                | (1)                  | (2)                  | (3)                  | (1)                       | (2)                  | (3)                  |
| Specification                  |                      |                      |                      |                           |                      |                      |
| Number of interviewer contacts | −0.028***<br>(0.002) | −0.037***<br>(0.002) | −0.026***<br>(0.002) | −0.012***<br>(0.002)      | −0.014***<br>(0.002) | −0.010***<br>(0.002) |
| Standard controls              | Yes                  | Yes                  | Yes                  | Yes                       | Yes                  | Yes                  |
| Health controls                | Yes                  | Yes                  | Yes                  | Yes                       | Yes                  | Yes                  |
| Year in panel                  | No                   | Yes                  | Yes                  | No                        | Yes                  | Yes                  |
| Interview mode                 | No                   | No                   | Yes                  | No                        | No                   | Yes                  |
| N (Set 1)                      | 249,919              | 249,919              | 249,919              | 249,919                   | 249,919              | 249,919              |
| Adjusted R <sup>2</sup>        | 0.169                | 0.176                | 0.184                | 0.055                     | 0.057                | 0.061                |

Life satisfaction is the dependent variable. Number of interviewer contacts is a linear variable ranging from 1 to 9. Standard controls include variables for employment status, unemployment, education, income, owner of dwelling, housing conditions, living area, household member in need of care, number of persons in household, no children in household, recent move, family status, partnership, age, gender (used only in the standard regressions), year and federal state. Health controls are variables for the degree of disability, doctor visits, and nights in hospital. Year in panel variables include dummies for the first survey participations. Interview mode variables are in line with the categories in Table 1. Heteroscedasticity-robust standard errors are in parentheses

Levels of statistical significance: \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$

**Table 10** Interviewee's willingness to participate and life satisfaction responses

| Method                         | Fixed effects regressions     |                       |                       |                       |
|--------------------------------|-------------------------------|-----------------------|-----------------------|-----------------------|
| Specification                  | (1)                           | (2)                   | (3)                   | (4)                   |
| Number of interviewer contacts | – 0.012***<br>(0.004)         | – 0.008*<br>(0.004)   | – 0.005<br>(0.004)    |                       |
| Willingness to participate     | Reference category: very good |                       |                       |                       |
| Good                           |                               | – 0.126***<br>(0.018) | – 0.118***<br>(0.018) | – 0.119***<br>(0.017) |
| Bad                            |                               | – 0.277***<br>(0.039) | – 0.263***<br>(0.039) | – 0.266***<br>(0.039) |
| Very bad                       |                               | – 0.334***<br>(0.089) | – 0.315***<br>(0.090) | – 0.322***<br>(0.090) |
| Standard controls              | Yes                           | Yes                   | Yes                   | Yes                   |
| Health controls                | Yes                           | Yes                   | Yes                   | Yes                   |
| Survey factors                 | No                            | No                    | Yes                   | Yes                   |
| N (Set 2)                      | 76,236                        | 76,236                | 76,236                | 76,236                |
| Adjusted R <sup>2</sup>        | 0.040                         | 0.042                 | 0.046                 | 0.046                 |

Life satisfaction is the dependent variable. Number of interviewer contacts is a linear variable ranging from 1 to 9. See Appendix Table 9 for information on the control variables. Survey factors include the year in panel and the interview mode variables. Heteroscedasticity-robust standard errors are in parentheses

Levels of statistical significance: \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$

**Table 11** Leaving the panel and life satisfaction responses (IV analyses) – Robustness check I

| A. First stage results               |                                      |                       |                       |                       |                       |
|--------------------------------------|--------------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| Dependent variable                   | Last participation of the respondent |                       |                       |                       |                       |
| Last year of the interviewer         | 0.093***<br>(0.003)                  | 0.093***<br>(0.003)   | 0.093***<br>(0.003)   | 0.093***<br>(0.003)   | 0.093***<br>(0.003)   |
| Unemployment rate                    |                                      | 0.000<br>(0.001)      |                       |                       | 0.000<br>(0.001)      |
| GDP per capita                       |                                      |                       | – 0.000<br>(0.001)    |                       | – 0.001<br>(0.001)    |
| Average gross wages                  |                                      |                       |                       | 0.001<br>(0.001)      | 0.002<br>(0.002)      |
| N (Set 4)                            | 182,790                              | 182,790               | 182,790               | 182,790               | 182,790               |
| Adjusted R <sup>2</sup>              | 0.040                                | 0.040                 | 0.040                 | 0.040                 | 0.040                 |
| B. Second stage results              |                                      |                       |                       |                       |                       |
| Dependent variable                   | Life satisfaction                    |                       |                       |                       |                       |
| Last participation of the respondent | – 0.534***<br>(0.169)                | – 0.564***<br>(0.170) | – 0.533***<br>(0.169) | – 0.514***<br>(0.169) | – 0.542***<br>(0.170) |
| Unemployment rate                    |                                      | – 0.008**<br>(0.004)  |                       |                       | – 0.005<br>(0.004)    |
| GDP per capita                       |                                      |                       | 0.023***<br>(0.004)   |                       | 0.015**<br>(0.007)    |
| Average gross wages                  |                                      |                       |                       | 0.036***<br>(0.008)   | 0.016<br>(0.011)      |
| N (Set 4)                            | 182,790                              | 182,790               | 182,790               | 182,790               | 182,790               |
| Adjusted R <sup>2</sup>              | 0.161                                | 0.160                 | 0.161                 | 0.161                 | 0.161                 |
| Wald F statistic                     | 778.568                              | 773.397               | 778.532               | 778.191               | 773.079               |

All specifications include standard controls and survey factors (see Table 5 for more information). Unemployment rate is the number of unemployed people divided by the working population. GDP per capita is the inflation-adjusted gross domestic product divided by the total population. Average gross wages are the inflation-adjusted gross wages divided by the number of employees. Each macroeconomic variable is determined at the federal state level and is for the year of the interview. Heteroscedasticity-robust standard errors are in parentheses. The instrument for a respondent's last year in the panel is the last year of the interviewer in the panel

Levels of statistical significance: \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$

**Table 12** Leaving the panel and life satisfaction responses (IV analyses)—Robustness check II

| Dependent variable:                                                                            | Life satisfaction     |                       |                       |                       |                       |
|------------------------------------------------------------------------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| A. Second stage results—no additional data restriction                                         |                       |                       |                       |                       |                       |
| Last participation of the respondent                                                           | − 0.534***<br>(0.169) | − 0.543***<br>(0.175) | − 0.520***<br>(0.187) | − 0.604***<br>(0.201) | − 0.528**<br>(0.211)  |
| Data restriction: Minimum # of interviews per interviewer                                      | —                     | 5                     | 10                    | 15                    | 20                    |
| N (Set 4)                                                                                      | 182,790               | 180,964               | 174,998               | 164,392               | 153,002               |
| Adjusted R <sup>2</sup>                                                                        | 0.161                 | 0.161                 | 0.161                 | 0.162                 | 0.164                 |
| Wald F statistic                                                                               | 778.568               | 735.912               | 653.757               | 565.025               | 499.409               |
| B. Second stage results—data restriction: only respondents in at least their 2nd year in panel |                       |                       |                       |                       |                       |
| Last participation of the respondent                                                           | − 0.645***<br>(0.184) | − 0.621***<br>(0.189) | − 0.633***<br>(0.199) | − 0.743***<br>(0.208) | − 0.682***<br>(0.217) |
| Data restriction: Minimum # of interviews per interviewer                                      | —                     | 5                     | 10                    | 15                    | 20                    |
| N (Set 4)                                                                                      | 164,483               | 163,087               | 158,047               | 148,895               | 139,238               |
| Adjusted R <sup>2</sup>                                                                        | 0.160                 | 0.160                 | 0.160                 | 0.160                 | 0.163                 |
| Wald F statistic                                                                               | 700.748               | 668.570               | 613.920               | 557.468               | 494.805               |
| C. Second stage results: data restriction—only respondents in at least their 3rd year in panel |                       |                       |                       |                       |                       |
| Last participation of the respondent                                                           | − 0.769***<br>(0.196) | − 0.749***<br>(0.201) | − 0.771***<br>(0.210) | − 0.883***<br>(0.220) | − 0.855***<br>(0.231) |
| Data restriction: Minimum # of interviews per interviewer                                      | —                     | 5                     | 10                    | 15                    | 20                    |
| N (Set 4)                                                                                      | 151,844               | 150,629               | 146,190               | 138,023               | 129,341               |

**Table 12** (continued)

C. Second stage results: data restriction—only respondents in at least their 3rd year in panel

|                         |         |         |         |         |         |
|-------------------------|---------|---------|---------|---------|---------|
| Adjusted R <sup>2</sup> | 0.158   | 0.159   | 0.159   | 0.158   | 0.161   |
| Wald F statistic        | 640.287 | 613.822 | 564.340 | 512.647 | 453.152 |

All specifications include standard controls and survey factors (see Table 5 for more information). Heteroscedasticity-robust standard errors are in parentheses. The instrument for a respondent's last year in the panel is the last year of the interviewer in the panel

Levels of statistical significance: \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$

## References

- Abowd JM, Stinson MH (2013) Estimating measurement error in annual job earnings: a comparison of survey and administrative data. *Rev Econ Stat* 95:1451–1467
- Angelini V, Cavapozzi D, Corazzini L, Paccagnella O (2014) Do danes and Italians rate life satisfaction in the same way? Using Vignettes to correct for individual-specific scale biases. *Oxf Bull Econ Stat* 76:643–666
- Angelini V, Bertoni M, Corazzini L (2017) Unpacking the determinants of life satisfaction: a survey experiment. *J R Stat Soc Ser A* 180:225–246
- Baetschmann G (2014) Heterogeneity in the relationship between happiness and age: evidence from the German Socio-economic panel. *Ger Econ Rev* 15:393–410
- Behr A, Bellgardt E, Rendtel U (2005) Extent and determinants of panel attrition in the european community household panel. *Eur Sociol Rev* 21:489–512
- Benjamin DJ, Heffetz O, Kimball MS, Rees-Jones A (2012) What do you think would make you happier? What do you think you would choose? *Am Econ Rev* 102:2083–2110
- Benjamin DJ, Heffetz O, Kimball MS, Rees-Jones A (2014) Can marginal rates of substitution be inferred from happiness data? Evidence from residency choices. *Am Econ Rev* 104:3498–3528
- Campanelli P, O'muircheartaigh C (1999) Interviewers, interviewer continuity, and panel survey nonresponse. *Qual Quant* 33:59–76
- Campanelli P, O'muircheartaigh C (2002) The importance of experimental control in testing the impact of interviewer continuity on panel survey nonresponse. *Qual Quant* 36:129–144
- Chadi A (2010) How to distinguish voluntary from involuntary unemployment: on the relationship between the willingness to work and unemployment-induced unhappiness. *Kyklos* 63:317–329
- Chadi A (2012) I would really love to participate in your survey! bias problems in the measurement of well-being. *Econ Bull* 32:3111–3119
- Chadi A (2013a) Third person effects in interview responses on life satisfaction. *J Appl Soc Sci Stud* 133:323–333
- Chadi A (2013b) the role of interviewer encounters in panel responses on life satisfaction. *Econ Lett* 121:550–554
- Chadi A (2015) Concerns about the Euro and happiness in Germany during times of crisis. *Eur J Polit Econ* 40:126–146
- Chadi A (2016) Identification of Attrition Bias Using Different Types of Panel Refreshments" IAAEU Discussion Paper 02/2016, Trier
- Clark AE, Oswald AJ (1994) Unhappiness and unemployment. *Econ J* 104:648–659
- Clark AE, Diener E, Georgellis Y, Lucas RE (2008) Lags and leads in life satisfaction: a test of the baseline hypothesis. *Econ J* 118:F222–F243
- Conti G, Pudney S (2011) Survey design and the analysis of satisfaction. *Rev Econ Stat* 93:1087–1093
- Contoyannis P, Jones AM, Rice N (2004) The dynamics of health in the british household panel survey. *J Appl Econom* 19:473–503
- Deuchert E, Huber M (2017) A cautionary tale about control variables in IV estimation. *Oxf Bull Econ Stat* 79:411–425
- Easterlin RA (1995) Will raising the income of all increase the happiness of all? *J Econ Behav Organ* 27:35–47
- Ferrer-i-Carbonell A, Frijters P (2004) How important is methodology for the estimates of the determinants of happiness? *Econ J* 114:641–659
- Fleurbay M (2009) Beyond GDP: the quest for a measure of social welfare. *J Econ Lit* 47:1029–1075
- Frey BS (2008) Happiness: A revolution in economics. MIT Press, London
- Frey BS, Stutzer A (2002) Happiness and economics. How the economy and institutions affect well-being. Princeton University Press, Princeton
- Frey BS, Stutzer A (2012) The use of happiness research for public policy. *Soc Choice Welf* 38:659–674
- Frijters P, Beaton T (2012) The mystery of the U-shaped relationship between happiness and age. *J Econ Behav Organ* 82:525–542
- Frijters P, Haisken-DeNew JP, Shields MA (2004) Investigating the patterns and determinants of life satisfaction in germany following reunification. *J Hum Res* 34:649–676
- Gerstorf D, Ram N, Mayraz G, Hidajat M, Lindenberger U, Wagner GG, Schupp J (2010) Late-life decline in well-being across adulthood in Germany, the United Kingdom, and the United States: something is seriously wrong at the end of life. *Psychol Aging* 25:477–485

- Goebel J, Krekel C, Tiefenbach T, Ziebarth NR (2015) Natural disaster, policy action, and mental well-being: the case of how natural disasters can affect environmental concerns, risk aversion, and even politics: evidence from Fukushima and three European countries. *J Popul Econ* 28:1137–1180
- Golsteyn BH, Hirsch S (2019) Are estimates of intergenerational mobility biased by non-response? Evidence from the Netherlands. *Soc Choice Welf* 52:29–63
- Heffetz O, Rabin M (2013) Conclusions regarding cross-group differences in happiness depend on difficulty of reaching respondents. *Am Econ Rev* 103:3001–3021
- Heffetz O, Reeves DB (2019) Difficulty to reach respondents and nonresponse bias: evidence from large government surveys. *Rev Econ Stat* 101:176–191
- Hill DH, Willis RJ (2001) Reducing panel attrition: a search for effective policy instruments. *J Hum Res* 36:416–438
- Jones AM, Koolman X, Rice N (2006) Health-related non-response in the British household panel survey and european community household panel: using inverse-probability-weighted estimators in non-linear models. *J R Stat Soc Ser A* 169:543–569
- Kapteyn A, Smith JP, van Soest A (2010) Life satisfaction. In: Diener E, Kahneman D, Helliwell J (eds) *International differences in well-being*, pp 70–107. Oxford University Press, Oxford
- Kassenboehmer SC, Haiken-DeNew JP (2009) You're fired! The causal negative effect of entry unemployment on life satisfaction. *Econ J* 119:448–462
- Kassenboehmer SC, Haiken-DeNew JP (2012) Heresy or enlightenment? The well-being age u-shape effect is flat. *Econ Lett* 117:235–238
- Kreuter F, Mueller G, Trappmann M (2010) Nonresponse and Measurement error in employment research: making use of administrative data. *Pub Opin Q* 74:880–906
- Kroh M (2009) Documentation of Sample Sizes and Panel Attrition in the German Socio Economic Panel (SOEP) (1984 until 2008), DIW Data Documentation No. 47
- Kroh M (2010) Documentation of Sample Sizes and Panel Attrition in the German Socio Economic Panel (SOEP) (1984 until 2009), DIW Data Documentation No. 50
- Kroh M (2012) Documentation of sample sizes and panel attrition in the German Socio Economic Panel (SOEP) (1984 until 2011), DIW Data Documentation No. 66
- Laurie H, Smith RA, Scott L (1999) Strategies for reducing nonresponse in a longitudinal panel survey. *J Off Stat* 15:269–282
- Layard R (2005) *Happiness. Lessons from a New Science*. Allen Lane, London
- Lipps O (2009) Attrition of Households and Individuals in Panel Surveys, DIW SOEPpapers on Multidisciplinary Panel Data Research 164
- Luechinger S, Meier S, Stutzer A (2010) Why does unemployment hurt the employed? Evidence from the life satisfaction gap between the public and the private sector. *J Hum Res* 45:998–1045
- Lynn P, Kaminska O, Goldstein H (2014) Panel attrition: how important is interviewer continuity? *J Off Stat* 30:443–457
- Metcalf R, Powdthavee N, Dolan P (2011) Destruction and distress: using a quasi-experiment to show the effects of the september 11 attacks on mental well-being in the United Kingdom. *Econ J* 121:F81–F103
- Meyer BD, Sullivan JX (2003) Measuring the wellbeing of the poor using income and consumption. *J Hum Res* 38:1180–1220
- Meyer BD, Mok WKC, Sullivan JX (2015) Household surveys in crisis. *J Econ Perspect* 29:199–226
- Nederhof A (1987) When neutrality is negative: pleasantness of most recent survey experience and non-response. *Qual Quant* 21:425–432
- Nicoletti C, Peracchi F (2005) Survey response and survey characteristics: microlevel evidence from the european community household panel. *J R Stat Soc Ser A* 168:763–781
- Olsen RJ (2005) The problem of respondent attrition: survey methodology is key. *Mon Labor Rev* 128:63–70
- Olson K (2013) Do non-response follow-ups improve or reduce data quality?: a review of the existing literature. *J R Stat Soc Ser A* 176:129–145
- Pannenberg M, Rendtel U (1996) Documentation of Sample Sizes and Panel Attrition in the German Socio Economic Panel (SOEP), Discussion Paper No. 137, DIW Berlin
- Pickery J, Loosveldt G, Carton A (2001) The effects of interviewer and respondent characteristics on response behavior in panel surveys: a multilevel approach. *Sociol Methods Res* 29:509–523
- Schueller S (2016) The effects of 9/11 on attitudes toward immigration and the moderating role of education. *Kyklos* 69:604–632

- Schraeppler JP, Schupp J, Wagner GG (2010) Individual and Neighborhood Determinants of Survey Non-response, DIW SOEPpapers on Multidisciplinary Panel Data Research 288
- Schwandt H (2016) Unmet aspirations as an explanation for the age U-shape in wellbeing. *J Econ Behav Organ* 122:75–87
- Stevenson B, Wolfers J (2009) The paradox of declining female happiness. *Am Econ J Econ Policy* 1:190–225
- Stiglitz JE, Sen A, Fitoussi JP (2009) Report of the Commission on the Measurement of Economic Performance and Social Progress, CMEPSP ([www.stiglitz-sen-toussi.fr](http://www.stiglitz-sen-toussi.fr))
- Stutzer A (2004) The role of income aspirations in individual happiness. *J Econ Behav Organ* 54:89–109
- Van Landeghem B (2012) A test for the convexity of human well-being over the life cycle: longitudinal evidence from a 20-year panel. *J Econ Behav Organ* 81:571–582
- Van Landeghem B (2014) A test based on panel refreshment for panel conditioning in stated utility measures. *Econ Lett* 124:236–238
- Van Praag B, Ferrer-i-Carbonell A (2008). *Happiness quantified: a satisfaction calculus approach*, (Revised edition), Oxford, New York
- Wagner GG, Frick JR, Schupp J (2007) The German socio-economic panel study (soep): scope, evolution and enhancements. *J Appl Soc Sci Stud* 127:139–169
- Weimann J, Knabe A, Schöb R (2015) *Measuring happiness: the economics of well-being*. MIT Press, Cambridge
- Wunder C, Heineck G (2013) Working time preferences, hours mismatch and well-being of couples: are there spillovers? *Labour Econ* 24:244–252
- Wunder C, Wiencierz A, Schwarze J, Kuechenhoff H (2013) Well-being over the life span: semiparametric evidence from British and German longitudinal data. *Rev Econ Stat* 95:154–167
- Zabel JE (1998) An analysis of attrition in the panel study of income dynamics and the survey of income and program participation with an application to a model of labor market behavior. *J Hum Res* 33:479–506