

## Item non-response on income and wealth questions

Regina T. Riphahn<sup>1,2</sup>, Oliver Serfling<sup>1</sup>

<sup>1</sup> University of Basel, P.O. Box 517, 4003, Basel, Switzerland  
(e-mail: regina.riphahn@unibas.ch)

<sup>2</sup> IZA Bonn, DIW, Berlin

First version received: June 2003/Final version received: June 2004

**Abstract.** This study investigates the mechanisms determining item nonresponse focusing on three issues: First, is there significant heterogeneity in item non-response across financial questions and in the association of covariates with item non-response across outcomes? Second, can the informational value of surveys be improved by matching interviewers and respondents based on their characteristics? Third, how does offering a “don’t know” answer option affect respondent behavior? The questions are answered based on detailed survey and interviewer data from the German Socioeconomic Panel using a broad set of income and wealth outcomes. We find considerable heterogeneity in non-response across financial items, little explanatory power of interviewer-respondent matches and strong evidence that “don’t know” answers result from mechanisms that differ from those yielding valid responses and outright refusals to respond.

**Key words:** Item non-response, survey quality, interviewer effects, income and wealth

**JEL classification:** C81, J30, I32

*“... the subject of item nonresponse’ is badly in need of investigation.”  
Ferber (1966, p.415)*

---

We thank Jörg-Peter SchrPäpler for generous support regarding the interviewer data used in this study and two anonymous referees, numerous participants of the GSOEP 2002 conference, of the 2003 meeting of the Ausschuss für Pökonometrie / Verein für Socialpolitik, and George Sheldon for very helpful comments.

## 1. Introduction

Survey data form the basis of much empirical economic research. Accordingly its quality and the various determinants thereof, as well as the implications of data deficiencies deserve the attention of researchers.

Within the range of data problems and quality concerns some garnered more attention in the social sciences than others: The disciplines of sociology and psychology, where interest often focuses on subjective statements, are mainly concerned with effects of the interview situation or interviewer influences: <sup>1</sup> If respondents seek interviewers' respect, their answers may deviate from the truth. Unit non-response and sample representativeness are issues raised in the economic literature (cf., Hill and Willis 2001; Horowitz and Manski 1998). Here also measurement error and recall bias find attention.<sup>2</sup> In contrast, the problem of item non-response is largely neglected. This is astounding as the loss of information due to item non-response could be even more problematic than respondent dropout from a survey.

Given the typically high rates of item non-response on sensitive issues such as income and wealth, it is important to learn about the determinants of non-response behavior. An understanding of the mechanisms driving item non-response may permit the development of techniques to reduce it and thus to substantively increase the value of interviews. It might improve researchers' ability to rigorously deal with non-response in their analyses, and finally, it may yield important insights to improve imputation procedures for missing data.

This study investigates such mechanisms and addresses the following questions: First, does the matching of interviewers to respondents affect respondents' willingness to provide information? This could be the case if 'observational closeness' between the interview partners helps to build up the level of trust required to reveal private information. If this were the case, survey administrators might be able to improve data quality by carefully pairing interviewers and respondents. Second, we analyze whether offering the option of "don't know" answers in questionnaires helps increase the amount of information provided, and, third, we investigate whether there is measurable heterogeneity in the response propensity for different types of financial questions. Little evidence exists on these last two issues. If item non-response propensities depend on the way the question is posed and differ for different types of financial questions it might be possible to utilize the findings to optimize survey strategies and to improve informational outcomes.<sup>3</sup>

The interdisciplinary literature on respondent behavior in surveys uses two prominent theoretical frameworks, the cognitive and the rational choice model. The cognitive model distinguishes several stages in the process of answering a question always focusing on cognitive challenges they pose for the respondent. The rational choice approach (Esser 1984) suggests that the respondent evaluates behavioral alternatives such as answering or refusing to answer based on their expected consequences and chooses what maximizes expected utility. One further issue in models of respondent behavior is that of

---

<sup>1</sup> For careful discussions of these problems see Esser (1984) or Reinecke (1991).

<sup>2</sup> See the special issues of the *Journal of Human Resources* (1998.2, 2001.3) and sources cited there.

<sup>3</sup> For a survey of possible procedures see Juster and Smith (1997).

trust in the interview situation. If a respondent distrusts an interviewer he is less ready to invest effort to recall or reveal information.

The empirical literature on item non-response yields four main findings:<sup>4</sup> (i) item non-response on income questions is concentrated in the tails of the income distribution; (ii) there seems to be little systematic variation in item non-response behavior; (iii) the predictions based on the cognitive model and the “trust” framework find support and interviewer-respondent matching affect survey success; (iv) the cognitive requirement and sensitivity of an issue reduce respondents’ willingness to answer.

One limitation of this literature is its focus on the income question: Heterogeneity in the level of cognitive challenges and item-specific sensitivities across financial outcomes have been neglected so far.<sup>5</sup> Also, the literature does not investigate the role of framing: If individuals show different responses depending on how a question is formulated, this information is relevant for the design of future surveys.<sup>6</sup>

We add to this literature in various ways. We exploit excellent data from the German Socioeconomic Panel (GSOEP), which contains information on respondent and interviewer characteristics, thus permitting research on the relevance of interviewer-respondent matches. We consider item non-response for a variety of income and wealth outcomes and investigate the characteristics of the don’t know answer option.

We find significant heterogeneity in item non-response across financial questions. Our data reveal that there is not much to be gained for the informational value of surveys by matching interviewers and respondents, once age and gender effects are controlled for. Our third result with respect to “don’t know” answer options is that the observed and unobserved characteristics of “don’t know” respondents are neither close to those providing informative answers nor to those refusing to respond. Therefore the “don’t know” responses should be considered as a separate group. Simple statements as to whether offering a “don’t know” answer option takes away from valid answers or from non-responses are not feasible. This also implies that missing data imputations should use different procedures depending on whether missing values derive from “don’t know” answers or non-responses.

## 2. Hypotheses and empirical approach

### 2.1. Detailed hypotheses

The theoretical frameworks describe individual item non-response behavior as determined by the relationship between respondent and interviewer, as well

---

<sup>4</sup> For prior studies see Lillard et al. (1986), Zweimüller (1992), Juster and Smith (1997), Knäuper et al. (1997), Souza-Poza and Henneberger (2000), Biewen (2001), Schräpler (2001), or Hill and Willis (2001).

<sup>5</sup> Exceptions are Schräpler (2001) who also investigates subjective concerns and Loosveldt et al. (1999) who look at political preferences.

<sup>6</sup> Framing is much discussed in the literature on attitude surveys. Trometer (1996) summarizes the evidence there which suggests that offering respondents a “don’t know” answer affects responses in important ways.

as by the costs and benefits of providing an answer. The literature operationalized these aspects by interpreting respondent and interviewer characteristics in the light of their effects on trust, and cost-benefit considerations.

If trust affects response behavior, sending a new interviewer to a given household should generate fewer informative responses than sending a well known interviewer. Thus an interviewer change in a panel survey is hypothesized to increase item non-response. Further, we expect that a match in relevant characteristics of interviewer and respondent increases response propensities, because it may affect respondents' perception of the interviewers' trust-worthiness.

In general, we assume that non-response propensities increase with the cost and decline with perceived benefits of answering. Costs and benefits vary with the type of question, the characteristics of the interview partners and the interview situation. As a respondent's cognitive ability may determine the effort required to answer a question, and as the cognitive ability may be correlated to education, we expect a negative correlation between high education and non-response. A factor that might be correlated with the perceived benefit derived from survey participation is the appreciation of public service. Existing studies suggest that this is particularly high among public sector employees, who in turn seem to be more ready to participate and provide information in surveys (e.g., Biewen 2001).

The costs and benefits of an interview might also be affected by the characteristics of the interview situation. One might e.g. take the size of a person's town of residence as an indicator of a general attitude of openness and trust. This is based on evidence showing that individuals refuse to participate in surveys because of fear of crimes and that larger cities often entail a sense of anonymity, where the limits of privacy are guarded more carefully than in rural areas.<sup>7</sup> Similarly, it might be easier for individuals to communicate with an interviewer if they are used to such exchanges. Therefore, residents in large households might be more at ease answering questions and providing information. Another characteristic of the interview situation is whether a respondent answers a questionnaire partly in writing as opposed to an oral interview. As it should be easier and less costly to refuse an answer if this does not have to be communicated to the interviewer it is plausible to expect higher item non-response in this situation.

## 2.2. Empirical approach

In our model of response behavior, we follow a rational choice framework and consider factors discussed in the literature. When asked a survey question individual  $i$  may respond in  $J$  different ways (e.g., provide a valid answer, not respond at all, or - if possible - answer "don't know").<sup>8</sup> In the framework of a random utility model, we can describe utility  $u$  resulting for individual  $i$  from behavior option  $j$  as follows:

<sup>7</sup> De Maio (1980) found significantly more survey cooperation among rural than among urban dwellers.

<sup>8</sup> We consider the event of the interview, the selection of the respondent, and the fact that the individual is in principle willing to respond to the survey as being exogenously given.

$$u_{ij} = c_{ij}\alpha_{1j} + b_{ij}\alpha_{2j} + X_i\beta_{1j} + W_m\beta_{2j} + (X_i^*W_m)\beta_{3j} + \mu_{ij} \quad (1)$$

where  $c_{ij}$  and  $b_{ij}$  represent the costs and benefits of answer option  $j$ ,  $\alpha$  and  $\beta$  are coefficients, and  $\mu$  is random noise. Also, respondent ( $i$ ) and interviewer ( $m$ ) characteristics ( $X, W$ ), and their interactions may affect the utility connected to a given behavioral response. Summarizing the right hand side variables in vector  $z$  and the coefficients in vector  $\gamma$ , our random utility model is

$$u_{ij} = z_{ij}\gamma_j + \mu_{ij} \quad (2)$$

The probability that individual  $i$  chooses option  $j$  is then

$$\begin{aligned} \Pr(\text{option } j \text{ is chosen} | c, b, X, W) &= \Pr(u_{ij} > u_{ik} | c, b, X, W) \\ &= \Pr(z_{ij}\gamma_j - z_{ik}\gamma_k > \mu_{ik} - \mu_{ij} | c, b, X, W) \\ &\text{for all } k \neq j, k = 1, 2, \dots, J, \end{aligned} \quad (3)$$

which must hold jointly for all  $J-1$  alternative options  $k$ . Assuming a distribution for  $\mu_{ik} - \mu_{ij}$  the resulting cumulative distribution function can be estimated by maximum likelihood.

Within this framework we investigate three issues: First, we describe whether item non-response rates differ across outcomes, and study whether such differences are associated with observable and unobservable determinants of item non-response.<sup>9</sup> For an intuitive indication of outcome-specific heterogeneity we pool item non-response outcomes across questions and test the significance of question specific covariate effects in addition to question specific fixed effects.

Second, we investigate whether the match between interviewer and respondent affects response behavior, by controlling for interactions between interviewer and respondent characteristics. We hypothesize that individuals feel more confident reporting financial information to someone of their own characteristics as suggested by with Sousa-Poza and Henneberger: "One potential source of nonresponse is the existence of a mis-match between the characteristics of the interviewer and the characteristics of the respondent."x (2000, p.83)

Finally, response probabilities might be affected by the way questions are posed (for evidence see e.g., Trometer 1996 and sources cited there). However, these studies typically do not focus on financial measures, and an analysis of the effect of alternative answer options for financial questions is missing in the literature.<sup>10</sup> Our data contains some questions with the option of answering "don't know", and others without this option. We first describe item non-response rates for both. Then, we perform two tests to find out

<sup>9</sup> If e.g. wealth items are considered a more private issue than income, the cost of revealing wealth may exceed that of income and we expect higher non-response for wealth. Similarly, if information on wealth is less familiar and difficult to obtain, we expect differences in response based on cognitive ease.

<sup>10</sup> Juster and Smith (1997) concentrate on financial measures but focus on the impact of adding follow-up bracket answer options for "don't know" and non-response outcomes and for imputation results.

whether response processes yielding “don’t know” answers differ from those yielding non-responses or informative answers.

Both tests are applied within the framework of the multinomial logit estimator. The first tests the assumption underlying this estimator that the disturbances of alternative (answer-) outcomes are uncorrelated. This ‘independence of irrelevant alternatives’ (IIA) property states that the set of outcome alternatives is correctly specified only if the estimates do not vary with the set of considered outcomes (Hausman and McFadden 1984).<sup>11</sup> This will be investigated applying a Hausman test. If the Hausman test yields that the *unobservable determinants* of the three outcomes – valid response, “don’t know” answer, and item non-response – are uncorrelated, this provides a first piece of evidence for the independence of the “don’t know” answer alternative. If uncorrelatedness is rejected, “don’t know” answers are not truly independent outcomes. In that case further investigation into the similarity to response or non-response alternatives is required.

The second test looks at whether the *observable determinants* of the three possible outcomes are correlated. The test was suggested by Cramer and Ridder (1991) but had also been performed before by Hill (1983). Cramer and Ridder (1991) describe the condition under which a subset of multinomial logit outcomes may be treated as a single state. They assume uncorrelated unobservables and describe a criterion by which one may choose the most parsimonious set of outcomes: If the slope coefficients of two outcome options do not differ significantly, the two options may be combined.<sup>12</sup> If the mechanisms determining the choice of a “don’t know” answer do not differ significantly from those determining valid answers, these processes are very similar, and it might well be that offering a “don’t know” answer option takes away from valid answers. Similarly, we can test whether the mechanisms leading to “don’t know” answers and non-responses are close. These tests provide a second indicator as to whether the availability of a “don’t know” answer option takes away from valid answers, or whether this reply is a substitute for non-responses. In the former case offering “don’t know” answers reduces the informative value of the survey.

### 3. Data description

#### 3.1. Dataset and sample

Our data are taken from the German Socioeconomic Panel (GSOEP). The GSOEP gathers information on German households and individuals periodically adding topical modules to the survey (SOEP Group 2001). Since the

---

<sup>11</sup> The classic illustration of the IIA property looks at alternative means of public transport. While taxi, train, and bus constitute valid alternatives, a split between red and blue buses most likely violates the IIA assumption: One would assume that the unobservable determinants of the choice between red and blue buses are correlated. We test whether don’t know answers are a “red bus” as opposed to being an independent alternative such as the train.

<sup>12</sup> Hill (1983) investigated whether females consider the decision to enter the labor force as an employee as being distinct from the choice to enter the labor force as a family worker. Similar tests were performed by Flinn and Heckman (1983), and Riphahn (1997).

1988 module was devoted to household wealth we evaluate item non-response for that year, when 4,814 households with 10,023 individuals were interviewed. Our data are taken from three questionnaires. The individual survey was administered to everybody aged 16 or older, whereas the household and wealth questionnaires were answered by heads of households.<sup>13</sup> We also take advantage of data describing GSOEP interviewers (cf., Schräpler and Wagner 2001), which is matched to respondent records.

The GSOEP applies various interview methods: Individuals can answer questions orally, they can fill in the questionnaire themselves with or without interviewer support, questionnaires may be sent out by mail, or answered via telephone. Generally interviewers are required to perform oral interviews, but they may use different formats depending on the situation.

To circumvent language problems, we select German respondents from the GSOEP's nationally representative subsample. We disregard observations where the survey was administered other than by meeting the interviewer in person as we are interested in the interaction between interviewer and respondent. This induces a loss of between 35 and 25 % of observations.<sup>14</sup> We also drop observations where information on the interviewer is missing.

### 3.2. *Dependent and explanatory variables*

*Dependent variables.* The financial variables of interest are taken from the individual, household, and wealth questionnaires. Table 1 describes measures gathered in the individual and household survey. Due to filtering mechanisms in the questionnaire the sample sizes vary by question.<sup>15</sup> The last column of Table 1 describes the item non-response rate for each measure. The rates vary markedly between 15 % for income from self-employment and less than 3 % for the "Christmas bonus", a common employment benefit in Germany. Averaging across all outcomes, we obtain a non-response rate of 5.0 % for individual income variables.

Based on cognitive ease, one might assume that providing last month's earnings should require less effort than last year's average monthly income. However, item non-response on the former is about twice that for the latter. If it is the sense of privacy that determines the cost of reporting earnings, this outcome may indicate that current earnings are more sensitive than those of the past. Generally regulated payments, such as vacation or retirement transfers seem to involve lower reporting costs - possibly because they are considered as less private - than those that entail information on labor market success (e.g. unemployment benefits, or earnings).

---

<sup>13</sup> The GSOEP has no strict definition of the "head of household". Instead it surveys a knowledgeable person for every household and tries to re-interview that same person in subsequent surveys (Hanefeld 1987).

<sup>14</sup> Preliminary results (not presented) confirm prior studies which show that the presence of an interviewer strongly affects item non-response (cf., Lillard et al. 1986; Schräpler 2001). For details on GSOEP interview methods and the sample selection see Tables 1 and 2 in Riphahn and Serfling (2002).

<sup>15</sup> E.g. only those who had indicated employment were asked about labor incomes, or those who were retired could indicate retirement benefits.

**Table 1.** Item non-response rates for individual and household income questions

Type of income	Number of cases <sup>1</sup>	Non-responses	
		N	Share
Questions from individual questionnaire			
Income from self employment <sup>2</sup>	274	42	15.3%
Bonus / profit sharing <sup>2</sup>	106	13	12.3%
Other benefits <sup>2</sup>	30	3	10.0%
Unemployment transfers <sup>2</sup>	196	18	9.2%
Gross earnings last month	2,546	211	8.3%
Earnings from other employment <sup>2</sup>	140	9	6.4%
Net earnings last month	2,546	135	5.3%
End of year payment: 13. & 14. monthly salary <sup>2</sup>	1,031	26	2.5%
Gross wage last calendar year <sup>2</sup>	2,454	91	3.7%
Vacation benefits <sup>2</sup>	1,501	47	3.1%
Christmas bonus <sup>2</sup>	1,149	33	2.9%
Retirement benefits <sup>2</sup>	983	26	2.6%
Total all individual questions:	12,956	654	5.0%
Questions from household questionnaire			
Interest and dividend income (last year)	2,149	312	14.5%
Monthly household net income (amount)	2,769	84	3.0%
Total all household questions: <sup>3</sup>	14,733	716	4.9%

Notes: <sup>1</sup> Number of cases indicating receipt of income or transfer.

<sup>2</sup> Average gross monthly amount in the last calendar year. If the respondent was unable to provide exact figure the questionnaire prompted for an approximation.

<sup>3</sup> Ten income and transfer categories are not listed separately to save space. For details see Riphahn and Serfling (2002, Table 4).

Source: Own calculations based on GSOEP.

The wealth questionnaire typically asked whether the household holds a given asset and, if so, at which value. If the respondent indicated possession of a given item but could not provide the exact amount, the person was first asked to guess, and if that failed in most cases answer categories or a “don’t know” reply were offered. Column 2 in Table 2 shows that the number of cases for each measure varies. The rates of non-response and “don’t know” answers differ strongly across items. The highest refusal rates of about 30 % are observed for questions on stock, bond, and equity ownership, which agrees with the findings of Juster and Smith (1997). The “don’t know” responses are distributed differently: The highest rates appear for equity (15 %) and inheritances (16.6 %). As the value of these items seems difficult to determine, “don’t know” likely reflects lack of knowledge. This seems less plausible in the case of monthly life insurance payments, where the respondent should be familiar with a figure showing up regularly on bank statements. Here a 10.9 % “don’t know” rate seems high.

While the non-response rates in Table 2 do not differ markedly from those in Table 1, the joint share of non-response and don’t know answers more than doubles these rates. Two factors might explain this difference: Either, offering an answer option “don’t know” induces individuals who may have otherwise provided an answer to indicate ignorance. Alternatively, wealth is either more sensitive than income or it is more difficult to know the correct answer.



**Table 2.** Item non-response rates for household wealth questions

Type of asset <sup>1</sup>	Number of cases	Non-response		"Don't know"		Total	
		N	Share	N	Share	N	Share
		Equity in a business	164	43	26.2%	25	15.2%
Stocks and bonds	636	217	34.1%	27	4.2%	244	38.4%
Home loan savings certificates ( <i>Bausparvertrag</i> )	1,001	150	15.0%	82	8.2%	232	23.2%
Inheritances since 1960	392	23	5.9%	65	16.6%	88	22.4%
Life Insurance	2,660	48	1.8%	290	10.9%	338	12.7%
Property other than occupied flat or home	306	6	2.0%	20	6.5%	26	8.5%
Savings account	2,064	70	3.4%	97	4.7%	167	8.1%
Owned home: Market value	1,065	8	0.8%	74	6.9%	82	7.7%
Total household wealth	2,427	32	1.3%	124	5.1%	156	6.4%
Household debt	771	7	0.9%	25	3.2%	32	4.2%
Total all:	11,486	604	5.3%	829	7.2%	1,433	12.5%

*Notes:* The survey first posed yes / no questions as to whether the household owns a given asset. Then the respondent was prompted for the exact amount held in this type of asset, or for an estimate. If that was not provided, response categories including the "don't know" option were provided. Non-response is coded if the asset type is available, but the amount was not provided. "Don't know" is coded if the asset type is available, but the amount was not provided. "Don't know" is coded if the first yes / no answer was positive and the respondent replied that the exact amount is unknown.

*Source:* Own calculations based on GSOEP.

*Explanatory variables.* Equation 1 (see Sect. 2.2) describes individual response behavior as determined by the costs and benefits of providing an answer, the characteristics of respondent and interviewer, as well as by their interactions. Clearly, it is not possible to actually measure individual costs and benefits in answering a given question. Therefore, the characteristics of respondents and interviewers are interpreted in the light of their effect on cost and benefit considerations.

In our item non-response model we control for characteristics of the respondent-interviewer match, such as equal labor market status and schooling, for the age difference, as well as for the gender combination between interviewer and respondent. We also measure whether a household's interviewer has changed since the last survey, which should increase item non-response. The remaining covariates were chosen as indicators of relevant costs and benefits in an interview situation. Education, as indicator of a respondent's cognitive ability, is measured using three categorical indicators.<sup>16</sup> We consider an indicator for whether respondents work in the public sector, control for household size, the size of a person's town of residence and for whether a respondent answers a questionnaire partly in writing as opposed to orally.

## 4. Results and Discussion

### 4.1. Heterogeneity in item non-response behavior and its determinants

*Item non-response and its determinants across outcomes.* To describe the relevance of alternative explanatory mechanisms behind item non-response and to investigate outcome-specific heterogeneity in item non-response we pool the outcome data described in Tables 1 and 2 and consider outcome specific fixed effects in the specification of our logit model.<sup>17</sup> The fixed effect controls are jointly highly significant (see Table 3) and reflect heterogeneity across outcomes, even after controlling for covariates. Adding outcome specific fixed effects to the model increases the pseudo (McFadden)  $R^2$  from about 0.018 (not presented) to almost 0.14.<sup>18</sup> The magnitudes of the estimated item-specific fixed effects (which we do not present to save space) reflect the descriptive evidence in Tables 1 and 2.

The estimation yields a number of statistically significant coefficients. The first group of variables describes the gender combination of respondent and interviewer with two males as the reference. The coefficients indicate positive associations between female interviewers and item non-response. If we assume that it is easier to avoid an answer in front of a female than a male the pattern fits the rational choice model's predictions.

<sup>16</sup> Low schooling is coded for mandatory schooling, medium schooling for the German *Realschule*, and high schooling for degrees preparing for academic studies.

<sup>17</sup> To render the bivariate non-response outcome measure of the income variables comparable to the multivariate outcome measure of the wealth indicators we dropped those wealth observations with "don't know" answers from the sample. The results presented below justify this procedure.

<sup>18</sup> This result holds in smaller subsamples as well, when we pool outcomes at the individual, household, or wealth level only (not presented to save space).

Table 3 Logit estimates on pooled outcomes

Variable	Fixed effects		Fixed effects with wealth interactions			
	Coeff.	t	Main effects		Interaction	
			Coeff.	t	Coeff.	t
Respondent female interviewer male	-0.007	-0.08	-0.010	-0.08	0.000	0.00
Respondent male interviewer female	0.140	1.72	0.149	1.30	0.004	0.02
Respondent female interviewer female	0.357	4.16	0.196	1.51	0.291	1.67
Respondent part time employed	0.422	4.59	0.520	4.34	-0.230	-1.19
Respondent not employed	0.353	4.19	0.405	2.54	0.020	0.10
Interviewer part time employed	-0.215	-2.20	-0.242	-1.68	0.036	0.18
Interviewer not employed	-0.071	-0.92	-0.001	-0.01	-0.137	-0.85
Same employment status	-0.054	-0.91	-0.035	-0.36	-0.035	-0.28
Respondent medium level schooling	-0.126	-1.79	0.261	2.68	-0.754	-5.29
Respondent high schooling	-0.192	-2.33	0.259	2.25	-0.880	-5.32
Interviewer medium level schooling	-0.235	-3.45	-0.411	-4.34	0.331	2.42
Interviewer high schooling	-0.041	-0.50	-0.056	-0.49	0.018	0.11
Same schooling	0.094	1.47	0.086	0.98	0.045	0.35
Respondent age	0.009	2.30	0.016	3.00	-0.013	-1.77
Age difference: respond. - interviewer	-0.004	-1.30	-0.000	-0.05	-0.008	-1.33
Change of interviewer	0.068	0.82	0.302	2.68	-0.457	-2.75
Public sector employee	-0.341	-4.20	-0.632	-5.47	0.582	3.53
Self administered survey	0.109	1.32	0.137	1.33	-0.037	-0.21
Lives in small town	0.081	-1.45	0.099	1.20	-0.365	-3.22
Household size	-0.126	-5.16	-0.016	-0.46	-0.240	-4.87
Respondent schooling missing	-0.036	-0.13	0.659	2.10	-1.958	-2.85
Significance test of fixed effects ( $\chi^2$ , p)	1390.8	-0.00		1153.5	0.00	
Log likelihood		-5,444.8			-5,376.8	

Table 3 (Continued)

Variable	Fixed effects		Fixed effects with wealth interactions			
			Main effects		Interaction	
	Coeff.	t	Coeff.	t	Coeff.	t
Pseudo R <sup>2</sup> (McFadden)		0.138			0.149	
Number of observations		28,531			28,531	

*Notes:*

- 1) The estimations combine the following outcome measures: All individual and household income categories listed separately in Table 1 and all measures from Table 2.
- 2) Since income from interest and dividends (reported at the household level) is an indicator of wealth we considered this outcome as a wealth outcome.
- 3) Fixed effect coefficients are not presented to save space.
- 4) The figures in the row on fixed effect significance tests provide the test statistic of a chi<sup>2</sup> test with 26 degrees of freedom. The p-value is given behind the test statistics.

*Source:* Own calculations based on GSOEP.

The next set of indicators describes the employment status of the participants. Overall, there seems to be a weak tendency for non full time employed respondents to refuse an answer. Again, the finding can be explained within the rational choice model: If the earnings of part-time workers are comparatively low, and these respondents prefer to indicate personal labor market success to the interviewer they may choose non-response. Alternatively, the intertemporal variance in part-time workers' earnings might be much higher compared to full-time employees'. In that case, the higher number of refused answers might be due to the higher cognitive challenge in providing an ever changing figure for an answer. The evidence on the role of the interviewers' employment status suggests that data quality is lowest when interviewers are full time employed. We find no indication of matching effects based on interviewer and respondent employment status.

Similarly, having respondents and interviewers with equal schooling does not significantly affect the results. The estimates indicate a positive correlation between both low respondent and interviewer schooling and non-responses.

Older respondents seem to be slightly more prone to item non-response than younger individuals.<sup>19</sup> We also find some evidence that having interviewers who are younger than the respondents reduces non-response. In contrast to the literature we find no significant non-response effect of an interviewer change. Possibly the change of an interviewer has strong effects on *unit* non-response (cf., Rendtel 1995) such that item non-response cannot even be observed. Overall, public sector employees seem to be significantly less likely to refuse an answer.

Based on the lower disutility involved, one might expect more non-responses among those who completed the questionnaire without an interviewer. This is confirmed only by an imprecisely estimated coefficient. Rural residence similarly has no statistically significant correlation with response behavior, and the household size effect yields as expected that individuals living in larger households have significantly lower non-response rates.<sup>20</sup>

Since non-response rates differ between wealth and income outcomes, we investigate in a second step whether this is a level effect or whether the covariate effects differ across the two outcome groups. We reestimate the fixed effects model, now adding a full set of interaction terms (I) which indicate whether a wealth or income measure is observed.<sup>21</sup> The model is thus:

$$u_{ij} = c_{ij}\alpha_1 + b_{ij}\alpha_2 + X_i\beta_1 + W_m\beta_2 + (X_i^*W_m)\beta_3 \\ + c_{ij}^*I_j\alpha'_1 + b_{ij}^*I_j\alpha'_2 + X_i^*I_j\beta'_1 + W_m^*I_j\beta'_2 + (X_i^*W_m^*)I_j\beta'_3 + \mu_{ij}$$

<sup>19</sup> This corresponds to the predictions of the cognitive model and corroborates the findings of Knäuper et al. (1997) who showed that differences in cognitive ability may lead to differential response patterns particularly among older respondents.

<sup>20</sup> In preliminary analyses we estimated these specifications separately on a wide variety of individual outcome measures. Whereas there was some heterogeneity in the correlation patterns across outcomes the overall outcomes correspond to those described for the pooled model (for details see Riphahn and Serfling 2002).

<sup>21</sup> In these estimations we treat "income from interest and dividends" as an indicator of wealth holdings and group it with the outcomes listed in Table 2.

The explanatory power of the model increases significantly with the full set of interaction terms added, of which a number are statistically significant (see last columns of Table 3).

The results suggest that the increase in non-response for female interviewers is somewhat more pronounced for wealth outcomes. A clear pattern appears for the schooling indicators: Whereas item non-response on income measures increases with higher respondent education, we find the opposite result for wealth outcomes. The differences are significant and difficult to interpret. If education is correlated with a respondent's level of information about wealth, then the high response propensity might be explained by cognitive ability. However, given that the same individuals should also be well informed on their income one can only speculate that they consider income as more private information.

The age effects seem to differ between income and wealth outcomes. The non-response probability on income measures increases with respondent age. The effect disappears for wealth questions. The negative correlation between the respondent-interviewer age difference and item non-response - pointed out above - seems to be based mostly on wealth outcomes.

Differences in covariate associations with non-response probabilities by outcome are observable also for the remaining variables: While the change of an interviewer increased non-responses for incomes, it reduces them for wealth. The beneficial effect of public sector employment on the propensity to provide financial information seems to be limited to incomes: Since the earnings of public sector workers in Germany typically follow publicly available pay scales, it is possible that these workers are more open about their income, as these may be public knowledge anyway. When it comes to wealth, however, their privacy protection instincts seem to be the same as for anyone else. Living in a small town is correlated with significantly lower non-response on wealth while the effect on income is insignificant. Also the negative effect of household size differs significantly for the two outcomes. Thus, non-response is heterogeneous in frequency and correlation patterns across outcomes.<sup>22</sup>

#### *4.2. A closer look at "don't know" answers*

In this section we investigate whether answering "don't know" is an independent outcome, or whether this response can be grouped with valid responses or with non-responses. As described above we first apply a Hausman test to determine whether the unobserved determinants of item non-response are correlated with those of valid answers or outright non-responses. We start with a sample that pools all of the outcomes presented in Table 2, combining the 11,486 observations of the "total" row. Then, we consider

<sup>22</sup> Since the combination of outcomes considered in the sample used in Table 3 is somewhat arbitrary, we performed robustness tests by reestimating the same models for alternative outcome subsets. There most coefficients have the same sign, but their statistical significance is not always robust to modifications of the sample.

**Table 4.** Summary of Hausman test results

Dependent variable	Number of obs.	Test-statistic	p-value
Pooled wealth measures	11,486	0.29	1.00
Stocks and bonds	636	-0.00	<sup>1</sup>
Home loan savings	1,001	-0.00	<sup>1</sup>
Savings account	2,064	-0.00	<sup>1</sup>
Owned home: Market value	1,065	0.00	1.00
Total household wealth	2,427	-3.23	<sup>1</sup>

*Note:* <sup>1</sup> The test statistic takes on a negative value, which can be interpreted as strong evidence against rejecting the null hypothesis that the IIA assumption holds (Hausman and McFadden 1984, p. 1226 Footnote 4, or Stata 7 Manual volume 2 p. 13).

*Source:* Own calculations based on GSOEP.

some of the wealth outcomes separately to determine whether the results for the pooled sample are robust.

The evidence presented in Table 4 seems to be strong and clear: The null hypothesis that the IIA assumption holds cannot be rejected in any of the tests. The unobservables do not seem to be correlated and therefore “don’t know” answers are valid and independent alternatives to informative responses and to non-responses in the IIA sense.

Next, we perform the (Hill-) Cramer-Ridder test of whether the observable covariates have significantly different effects on the propensity to provide a “don’t know” answer relative to either valid answers or to outright non-response. The test compares the multinomial logit slope coefficients to those of a model where the slope parameters for don’t know and one alternative answer option are restricted to be identical. Again, we perform it first for the sample of pooled wealth measures and then for some of the wealth items separately.

The evidence (see Table 5) is clear: In all cases we can reject the hypotheses that the coefficients of the “don’t know” answer are identical to those of valid responses at high levels of statistical significance. For all outcomes but ‘market value of an owned home’ we reject that the coefficients of the “don’t know” answer are identical to those of valid responses mostly at the one percent level. We read this evidence as indicative of the independence of “don’t know” answers: Typically, observable determinants of response behavior have significantly different impacts on the three considered outcomes.

Jointly the two tests suggest that neither by their observable nor by their unobservable determinants “don’t know” answers are correlated with - and therefore likely substitutes of - valid answers or complete item non-responses. These findings can be misleading if they are completely due to the limited statistical power of the tests. Thus, it would be interesting to complement the statistical argument presented here using evidence from experiments. Unfortunately that is not possible with the GSOEP data. We conclude that “don’t know” answers can be viewed as independent outcomes in their own right. Missing values due to “don’t know” replies cannot simply be mixed

**Table 5.** Summary of Hill-Cramer-Ridder test results

Dependent variable	No. of obs.	$H_0: \beta^{\text{Don't Know}} = \beta^{\text{Response}}$ LR <sup>1</sup> (DF <sup>2</sup> ; p-value)	$H_0: \beta^{\text{Don't Know}} = \beta^{\text{Item-Nonresponse}}$ LR <sup>1</sup> (DF <sup>2</sup> ; p-value)
Pooled wealth measures	11,486	215.34 (21; 0.000)	162.52 (21; 0.000)
Stocks and bonds	636	39.51 (20; 0.006)	42.18 (20; 0.003)
Home loan savings	1,001	60.81 (20; 0.000)	100.98 (20; 0.000)
Savings account	2,064	65.75 (20; 0.000)	45.87 (20; 0.001)
Owned home: market value	1,065	65.31 (17; 0.000)	21.44 (17; 0.207)
Total household wealth	2,427	87.74 (21; 0.000)	31.24 (21; 0.070)

Notes: <sup>1</sup> LR represents the value of the likelihood ratio test statistic.

<sup>2</sup> The degrees of freedom differ across wealth measures, since due to collinearity and small number of cases the full model (see Table 2) could not be estimated. The full model was estimated for the pooled wealth measures and for total household wealth. For testing Savings, Home Loan Savings and Stocks and Bonds, the indicator of missing respondent schooling was omitted. In the case of ownership of occupied flat or home, the indicators of self administered survey and higher respondent schooling were also dropped from the econometric model.

Source: Own calculations based on GSOEP.

with item non-responses. The test results show that the two processes are determined by different observable and unobservable mechanisms.<sup>23</sup> “Therefore, our results yield additional support to the conclusion of Juster and Smith (1997) in their analysis of responses to follow-up bracket questions in surveys (p.1272): “This marked contrast in the behavior of DK and REF responses suggests that the two need to be handled separately when imputations are being done”, where DK represents don’t know and REF refusal to respond.

## 5. Conclusions

Even though item non-response affects every analysis using survey data, it has found little attention as a behavioral phenomenon in its own right.<sup>24</sup> In this study we present a number of results that are new to the literature. The empirical literature on item non-response is limited and generally focuses on

<sup>23</sup> This confirms Juster and Smith (1997), who showed that “don’t know” respondents and non-respondents differ in their willingness to provide responses to “bracket” questions (i.e. follow-up questions asked when initially no valid response is received): Whereas almost 80 % of initial “don’t know” respondents provided complete bracket data, the share among non-respondents reached only 40 %.

<sup>24</sup> Certainly a vast statistical literature has developed following Rubin’s influential work on missing data imputation (Rubin 1987). However the issue there is to find the best possible correction given that the data is missing. Our interest is to explain at least in part why it is missing in order to improve data collection efforts.



measures of labor income. We address this limitation by investigating the frequency and determinants of item non-response for a variety of financial outcomes. We find significant heterogeneity in non-response intensities. This conclusion from descriptive statistics is confirmed in regressions of non-response behavior where much explanatory power derives from the consideration of outcome-specific fixed effects. We confirm several results of the literature regarding correlates of item non-response. An investigation of the homogeneity of non-response determinants across outcomes yields new insights: estimating a fully interacted model shows clearly that a number of the established correlates of item non-response behavior vary depending on the specific item under consideration.

We investigate whether the match of interviewer and respondent characteristics affects the quality of survey outcomes. Robust findings on this matter would be valuable to reduce the cost and to increase the quality of information gathered from social surveys. The analysis yields that non-response rates tend to be higher if the interviewer is female in particular when the respondent is female as well. Having a respondent and an interviewer with the same employment status or the same educational level does not significantly affect non-response outcomes. However, our measures of employment and educational attainment may be too rough to reflect the impact of potential matching effects on non-response behavior. With respect to age differences there is some evidence that matching a younger interviewer to an older respondent may increase response propensities particularly with respect to wealth outcomes. Interestingly, the personal acquaintance of the respondent with the interviewer is beneficial for wealth but not for income items.

Our third research question concerns “don’t know” answers in questionnaires. A Hausman test of the independence of irrelevant alternatives assumption and the (Hill-)Cramer-Ridder test suggest that “don’t know” responses cannot be viewed as a subcategory of valid answers nor as comparable to item non-response. Therefore, simple statements as to how offering “don’t know” answer options affects the set of valid answers are not possible.

In the end, researchers have to acknowledge that the group of respondents who refuse to answer a survey question is not a random draw from the population, that the group varies depending on the question looked at, that those answering “don’t know” differ from non-respondents, and that simply omitting these individuals from the analysis may well bias results. Much attention has been devoted to developing appropriate imputation mechanisms when data is missing. Our results suggest that imputation procedures should differentiate between missing values due to “don’t know” answers and due to outright non-response.

## References

- Biewen M (2001) Item non-response and inequality measurement: Evidence from the German earnings distribution. *Allgemeines Statistisches Archiv* 85: 409–425
- Cramer JS, Ridder G (1991) Pooling states in the multinomial logit model. *Journal of Econometrics* 47: 267–272
- De Maio TJ (1980) Refusals: Who, where and why. *Public Opinion Quarterly* 44: 223–233
- Esser H (1984) Determinanten des Interviewer- und Befragtenverhaltens: Probleme der theoretischen Erklärung und empirischen Untersuchung von Interviewereffekten (Determinants of interviewer- and respondent behavior: Problems of theoretical explanations and

- empirical analyses of interviewer effects). In: Mayer K, Schmidt P (eds) *Allgemeine Bevölkerungsumfrage der Sozialwissenschaften*, 26–71, Frankfurt
- Ferber R (1966) Item nonresponse in a consumer survey. *Public Opinion Quarterly* 30: 399–415
- Flinn CJ, Heckman JJ (1983) Are unemployment and out of the labor force behaviorally distinct labor force states? *Journal of Labor Economics* 1(1): 28–42
- Hanefeld U (1987) *Das sozio-oekonomische Panel: Grundlagen und Konzeption*. Campus Verlag, Frankfurt / Main
- Hausman JA, McFadden D (1984) Specification tests for the multinomial logit model. *Econometrica* 52(5): 1219–1240
- Hill MA (1983) Female labor force participation in developing and developed countries - consideration of the informal sector. *Review of Economics and Statistics* 6(3): 459–468
- Hill D, Willis RJ (2001) Reducing panel attrition: A search for effective policy instruments. *Journal of Human Resources* 36(3): 416–438
- Horowitz JL, Manski CF (1998) Censoring of outcomes and regressors due to survey nonresponse: Identification and estimation using weights and imputations. *Journal of Econometrics* 84: 37–58
- Journal of Human Resources (1998) Special issue: Attrition in longitudinal surveys. *Journal of Human Resources* 33(2)
- Journal of Human Resources (2001) Special issue on data quality. *Journal of Human Resources* 36(3)
- Juster FT, Smith JP (1997) Improving the quality of economic data: Lessons from the HRS and AHEAD. *Journal of the American Statistical Association* 92(440): 1268–1278
- Knäuper B, Belli RF, Hill DH, Herzog AR (1997) Question difficulty and respondent's cognitive ability: The effect on data quality. *Journal of Official Statistics* 13(2): 181–199
- Lillard L, Smith JP, Welch F (1986) What do we really know about wages? The importance of nonreporting and census imputation. *Journal of Political Economy* 94(3): 489–506
- Loosveldt G, Pickery J, Billiet J (1999) Item non-response as a predictor of unit non-response in a panel survey. Paper presented at the International Conference on Survey Non-response, Portland Oregon, USA
- Reinecke J (1991) *Interviewer- und Befragtenverhalten. Theoretische Ansätze und methodische Konzepte*. Westdeutscher Verlag, Opladen
- Rendtel U (1995) *Panelausfälle und Panelrepräsentativität*. Campus Verlag, Frankfurt, New York
- Riphahn RT (1997) Disability retirement and unemployment - substitute pathways for labor force exit? *Applied Economics* 29(5): 551–561
- Riphahn RT, Serfling O (2002) Item non-response on income and wealth questions. IZA Discussion Paper No. 573, Bonn
- Rubin DB (1987) *Multiple imputation for non-response in surveys*. Wiley, New York
- Schräpler J-P, Wagner GG (2001) *Das Verhalten von Interviewern - Darstellung und ausgewählte Analysen am Beispiel des "Interviewerpanels" des Sozio-ökonomischen Panels*. *Allgemeines Statistisches Archiv* 85: 45–66
- Schräpler J-P (2001) Respondent behavior in panel studies. A case study of the German socio-economic panel (GSOEP). DIW Discussion Papers No. 244. DIW-Berlin
- SOEP Group (2001) The German socio-economic panel (GSOEP) after more than 15 years - Overview. In: Holst E, Lillard DR, DiPrete TA (eds) *Proceedings of the 2000 Fourth International Conference of German Socio-Economic Panel Study Users (GSOEP2000)*. *Vierteljahrshefte zur Wirtschaftsforschung (Quarterly Journal of Economic Research)* 70(1): 7–14
- Sousa-Poza A, Henneberger F (2000) Wage data collected by telephone interviews: an empirical analysis of the item nonresponse problem and its implications for the estimation of wage functions. *Schweizerische Zeitschrift für Volkswirtschaft und Statistik* 136(1): 79–98
- Trometer R (1996) *Warum sind Befragte "meinungslos"? Kognitive und kommunikative Prozesse im Interview. (Why do respondents not have an opinion? Cognitive and communicative processes in interviews)* Ph.D. Dissertation University of Mannheim, mimeo
- Zweimüller J (1992) Survey non-response and biases in wage regressions. *Economics Letters* 39: 105–109