Journal of Official Statistics, Vol. 26, No. 2, 2010, pp. 233–269

Changing from PAPI to CAPI: Introducing CAPI in a Longitudinal Study

Jörg-Peter Schräpler[1,3]*, Jürgen Schupp*[2,3]*, and Gert G. Wagner*[3,4]

This article examines the implications of moving to Computer-Assisted Personal Interviewing (CAPI) for data quality by analyzing the transition from Paper-and-Pencil (PAPI) to Computer-Assisted Personal Interviewing (CAPI) on a subsample of the German Socio-Economic Panel (SOEP) conducted using an "experimental design" in Wave 1. The 2,000 addresses for the sample E of SOEP were split into two subsamples with the same structure, E1 and E2, using twin-sample points. Each of the 125 sample points contained 16 addresses (8 for E1 and 8 for E2) and with each interviewer alternating in the first wave between PAPI and CAPI mode. In the subsequent waves, the PAPI mode was partly replaced by CAPI. With this design we are able to control for possible interviewer effects in the analysis of mode effects in Wave 1. The article assesses whether any mode effects or mode-related compound effects are apparent for the response rate. Within the data, we examine monetary dimensions such as gross income and item and unit nonresponse rates.

Key words: CAPI; mode effects; data quality; interviewer effects; JEL classification: C81.

1. Introduction

This article assesses the effect of moving from the traditional Paper-and-Pencil Interviewing (PAPI) method to Computer-Assisted Personal Interviewing (CAPI) within an ongoing panel study by means of an "experimental design." This study was conducted on a subsample of the German Socio-Economic Panel (SOEP) starting in 1998.

The SOEP is a longitudinal representative survey containing socioeconomic information on private households in the Federal Republic of Germany (cf. Wagner et al. 2007). It is similar to the US Panel Study of Income Dynamics (PSID). DIW Berlin (German Institute for Economic Research) manages the SOEP study. The first wave of data, collected in 1984 in the Federal Republic of Germany, prior to German reunification, contains 5,921 households. The original sample was supplemented by a sample of East German residents in 1990 (2,179 households) and a sample of immigrants in 1994–1995 (522 households). Additional refreshment samples were added in 1998

[1] Ruhr-University Bochum and IT. NRW, Gebäude GB1/129, Universitätsstr. 150, D-44780 Bochum, Germany. Email: joerg-peter.schraepler@rub.de
[2] FU Berlin, Berlin, Germany.
[3] DIW Berlin, Berlin, Germany.
[4] Berlin University of Technology and Max Planck Institute for Human Development, Berlin, Germany.
Acknowledgments: We would like to thank four anonymous reviewers for helpful comments. The usual disclaimer applies.

(1,056 households), 2000 (6,052 households) and 2006 (1,506 households). In 2002 a subsample of oversampled high income earner was introduced (1,224 households). In the SOEP, all household members age 16 and older are interviewed. We focus our analysis on the refreshment sample of 1998, when Computer-Assisted Personal Interviewing (CAPI) was used in the SOEP for the first time.

CAPI is an increasingly viable alternative for data collection in survey research. In CAPI, interviewers visit respondents with a portable computer and conduct a face-to-face interview using the computer. After the interview, the data are transmitted to a central computer (cf. Biemer and Lyberg 2003). CAPI was first tested in Europe by Statistics Sweden (Danielsson and Maarstad 1982) in 1982 and by Statistics Netherlands in 1984 (Bemelmans-Spork and Sikkel 1985). The first national household survey that used CAPI for all of its data collection was the Netherlands Labor Force Survey in 1987 (van Bastelaer et al. 1988). In the same year the first U.S. national household survey, the Nationwide Food Consumption Survey, was conducted by national analysts, using CAPI for at least part of the data collection (Rothschild and Wilson 1988). Since that time the use of CAPI has grown rapidly. The further history and development of the implementation process worldwide has been described in detail by Couper and Nicholls (1998).

An important challenge is the transition from PAPI to CAPI within an ongoing panel study. The British Household Panel Study (BHPS) took this step, but without the means to check the potential influence of the new mode on results (cf. Laurie 2003). The problem here is that even if CAPI is capable of delivering better data quality than PAPI, an impact resulting from a change in interviewing mode is undesirable since it could potentially create artificial longitudinal results. For the second major household panel study in Europe to introduce CAPI, the German Socio-Economic Panel (SOEP), the new mode was introduced in a more controlled manner. In this article we analyze the results of the transition from PAPI to CAPI in the SOEP.

2. Previous Experiences of the Use of CAPI in Surveys

De Leeuw and Nicholls (1996) point out that "the question of whether or not computer assisted data collection methods (CADAC) should be used for survey data collection is no longer an issue. In fact, most professional research organizations. . . are adopting these new methods with enthusiasm." The main potential advantages and disadvantages of CAPI as well as the effects of the transition from PAPI to CAPI have been discussed by Weeks (1992), Martin and Manners (1995) and Nicholls et al. (1997). Often anticipated benefits in moving from PAPI to CAPI are cost savings and a reduction in the time elapsed between fieldwork and the availability of the data for analysis. For academic studies like BHPS and SOEP, the potential improvement in data quality is the most important benefit. The quality improvement need not only be due to CAPI itself, but can partly be the result of a self-selection process among interviewers: if the professional interviewers want to work with CAPI, the quality of surveys conducted using PAPI decline, due to negative self-selection into the group of remaining interviewers.

Nevertheless, cost savings do not seem to be a common outcome when PAPI interview methods are replaced by CAPI (Couper and Nicholls 1998). CAPI requires sizeable investments in hardware and more front-end design and development work than PAPI.

These costs have to be balanced against the avoided costs of producing and handling paper questionnaires, data keying and office editing at the back-end (Martin et al. 1993). Since these back-end costs are variable and the front-end costs are largely fixed, CAPI is economically attractive for large and ongoing surveys (Weeks 1992). This means that CAPI remains the domain of a few large survey organizations (Couper 2005; Biemer and Lyberg 2003).

A main advantage of CAPI is that at the end of the interview the CAPI data are in electronic form and ready to be combined into a simple raw data set. Hence the time elapsed between the fieldwork and the availability of the data for analysis is reduced (Martin et al. 1993). Nicholls and De Leeuw (1996) found that saving time was the most frequently cited reason for adopting computer-assisted interviewing for specific studies mentioned in the literature.

Another important aspect is the abundance of empirical evidence that computer assisted interviewing (CAI) and especially CAPI improves data quality. The literature reports fewer instances of missing data (Sebestik et al. 1988; Olsen 1992), mostly because interviewers cannot make routing errors.[5] Nicholls et al. (1997) report that "one of the most consistent conclusions of the CAI literature is that CAI can eliminate virtually all respondent and interviewer omissions of application items, but provide little or no reduction in rates of explicit refusals."

In the case of unit nonresponse, there was some concern that this new technology would further reduce willingness to participate in surveys (Couper 2005) or that CAPI respondents would object to having their information stored on a computer. But the studies that compared the refusal rates in CAPI with those in PAPI have found no significant differences (cf. Baker et al. 1995) or only slightly higher rates of unit nonresponse with computer-assisted data collection than with conventional paper methods (Tourangeau et al. 1997). Often these small differences are attributed to the inevitable hardware difficulties when a new technology is introduced, but they may also reflect some resistance to the computers on the part of interviewers or respondents.

The willingness of respondents and interviewers to accept the new technology may also affect the data quality. Baker (1992) reports that most respondents find CAPI interesting and amusing, and attribute a greater degree of professionalism to CAPI. Generally speaking, most reactions are either neutral or positive, while only a steady minority of 5% tends to prefer paper-and-pencil versions of the interview (De Leeuw et al. 1995). De Leeuw et al. (1995) report that when explicitly asked about the data privacy, 47% have more trust in the privacy of computer-collected data, 5% have more trust in traditionally collected data, and 48% see no difference. Respondents' positive reactions to the new data collection methods are in line with the findings of some studies that compare PAPI and CAPI and report slightly less social desirability bias with CAPI (Baker and Bradburn 1992; Martin et al. 1993). Baker et al. (1995) reported a greater respondent willingness to

[5] Besides these control features, CAPI is also easier to implement in longitudinal study designs like "dependent interviewing." Such surveys may attempt to update information collected previously by presenting sample members with the prior information and asking them to confirm whether or not their circumstances have changed (dependent interviewing), rather than simply asking them to state their current circumstances (independent interviewing). Further details can be found in Lynn et al. (2006).

disclose sensitive information. But overall these differences seem to be rather small (De Leeuw et al. 1995).

De Leeuw et al. (1995) and Martin et al. (1993) also describe broad interviewer acceptance of CAPI. Once trained, most interviewers preferred to use CAPI. The only important complaint raised by interviewers was the difficulty of grasping the overall structure of the questionnaire (Riede and Dorn 1991), and some complained about the weight of the laptop (Edwards et al. 1993, cited in De Leeuw et al. 1995).

3. The "Experimental Design" of Sample E, Wave 1, in the SOEP

The German Socio-Economic Panel (SOEP) study (cf. Schupp and Wagner 2002) was extended by a refreshment sample E in 1998 (Infratest 1998; 2000). The aim in using this new sample is fourfold: 1. stabilization of the number of cases, 2. in-depth analysis of potential panel effects, 3. extended analysis of SOEP's overall representativeness and 4. analysis of the transition from PAPI to CAPI using an experimental design.

3.1. Design of Sample E

All samples of SOEP are multi-stage random samples which are regionally clustered and the respondents (households) are selected by random walk.[6] The SOEP is conducted by a "method mix," which is now generally recommended by a number of survey researchers (cf. Voogt and Saris 2005, for an overview see De Leeuw 2005). The preferred procedure for performing the survey is PAPI-based face-to-face interviews. Respondents may also complete the questionnaire themselves in the presence of the interviewer (self-completed) and receive help from the interviewer if needed. Sometimes a single interview combines both procedures (mixed). In Waves 2 and later, interviews were conducted by mail in cases where respondents would otherwise probably not cooperate.

Wave 1 of subsample E was done in a rarely performed "textbook version" of a random walk. The listing of the addresses was separated from the interviewing process, so the interviewers had fixed addresses (like register addresses). Sample E contains 2,000 German households that were split into two identical subsamples with the same structure, E1 and E2, using twin sample points. Each of the 125 sample points contains 16 addresses (8 for E1 and 8 for E2), which had to be surveyed in the first wave by a single interviewer, alternating between PAPI-based methods and CAPI modes. For each address, it was defined in advance if the interviewer had to use PAPI or CAPI. Nevertheless, to achieve the intended sample size and to prevent refusals and nonparticipation, some exceptions were allowed:

- In some CAPI households with many respondents, some individuals were allowed to use PAPI as well. In the case of large households the PAPI method is slightly more flexible than CAPI because the other respondents can complete their questionnaire

[6] The "guestworker sample" (Subsample B) was the only exception: it was surveyed by means of register data. In a random walk, sample points are selected randomly by means of a multi-stage stratified sampling procedure. Here, the interviewer selected the households within the selected constituency according to a random-route procedure. Working from a given random start address, the interviewer had to select every seventh household as a target household (cf. Haisken-DeNew and Frick 2005).

with paper and pencil at the same time, with the interviewer present in the household (von Rosenbladt and Stutz 1998).

- In the last stage of the fieldwork, there were very few good PAPI interviewers who worked also in the CAPI subsample and used paper and pencil.
- In a few cases in both subsamples the interview was carried out by mail to prevent refusals.

Question presentation and wording: For comparability, the question wording for the PAPI and the CAPI mode was exactly the same. This was also true for the "Don't know" categories. In both versions there were no explicit categories for "refusals." In the case of CAPI respondents, the answer options were provided in the form of lists, especially to visualize questions like Likert scales. This had the advantage that the interviewer could concentrate on the laptop entries. In the case of PAPI, the interviewer showed respondents the questionnaire to visualize scale-type questions.

Table 1 shows the response rate in Wave 1 of subsample E1 and E2. Because of 3.4% neutral losses (uninhabited apartments) in E1 and 2.6% in E2, 52 reserve addresses were used. After that the remaining gross sample consisted of 994 addresses in E1 and 998 addresses in E2. The systematic losses in E1 were caused by 40.1% refusals (41.2% in E2) and 4.0% noncontacts (5.2% in E2) as well as 0.5% nonutilizable interviews (0.6% in E2). The resulting total response rate was 54.1% of sample E1 and 51.9% of sample E2. In addition 23 household interviews were detected as having been fabricated by two interviewers (12 in E1 and 11 in E2).

Table 2 shows the effective data collection methods in both subsamples E1 and E2 on the household and the individual level in the first wave. Overall from the 2,000 household addresses (E1 + E2) 1,056 household interviews were realized. According to the survey plan over 80% of the household interviews in E1 were in fact collected via PAPI and over 76% of the household interviews in E2 were in fact collected via CAPI. On the individual level, the corresponding proportions are slightly lower. Hence we can conclude that the intended method split was not performed completely in order to avoid unit nonresponse, but the partial segregation of data collection methods and interviewer clusters does allow us to analyze these components separately.

Table 1. Response rate in Wave 1 of Subsamples E1 and E2

	E1 (PAPI)		E2 (CAPI)	
	N	%	*N*	%
Household addresses	1,000		1,000	
Neutral losses[a]	34		26	
Reserve addresses	28		24	
Household addresses	994	100.0	998	100.0
Not reached	40	4.0	52	5.2
Finally refused	399	40.1	411	41.2
Not evaluable	5	0.5	6	0.6
Faked household interviews	12	1.2	11	1.1
Number of realized household interviews	538	54.1	518	51.9

[a] Apartments were uninhabited or resident had died.

Table 2. Data collection methods in the Subsamples E1 and E2, Wave 1 (1998)

	E1 (PAPI)		E2 (CAPI)		Total	
Method	*N*	%	*N*	%	*N*	%
Households						
PAPI	432	80.3	90	17.4	522	49.4
CAPI	16	3.0	398	76.8	414	39.2
Self completed	41	7.6	6	1.2	47	4.5
Mixed	25	4.6	6	1.2	31	2.9
Mail	22	4.1	16	3.1	38	3.6
Not known	2	0.4	2	0.4	4	0.4
Total	538	100.0	518	100.0	1,056	100.0
Persons						
PAPI	650	65.4	158	17.2	808	42.3
CAPI	24	2.4	679	74.1	703	36.8
Proxy	2	0.2	–	–	2	0.1
Self completed	146	14.7	43	4.7	189	9.9
Mixed	127	12.7	15	1.7	142	7.4
Mail	21	2.1	19	2.1	40	2.1
Not known	24	2.4	2	0.2	26	1.4
Total	994	100.0	916	100.0	1,910	100.0

Source: SOEP 1998, Sample E.

Figures 1 and 2 show the development of the methods used in subsamples E1 and E2 on the individual level. We can recognize that it was used to maintain the method split between CAPI and PAPI in the first two waves. Figure 3 shows that after the second wave, the PAPI mode is replaced systematically by CAPI.

Table 3 shows the distribution of the data collection methods in the first five waves of the whole sample E on the household and individual levels. In the first wave, 49% of all

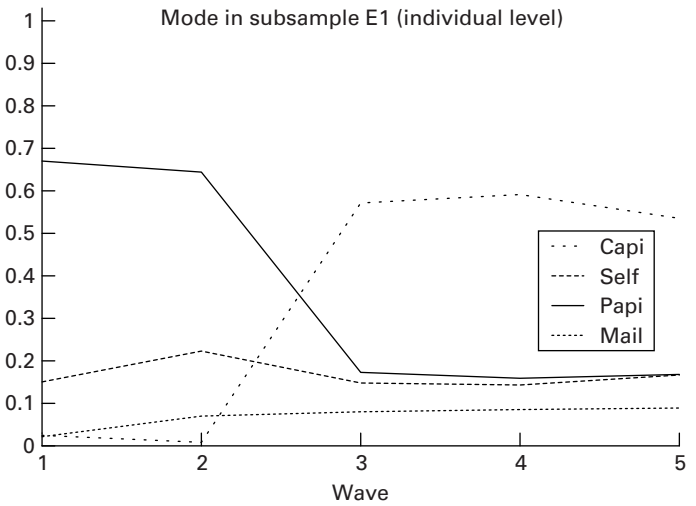

Fig. 1. SOEP, Subsample E1: Share of data collection methods

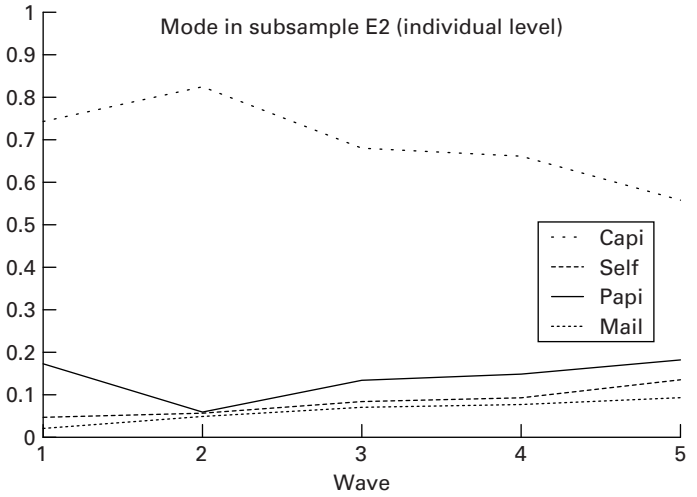

Fig. 2. SOEP, Subsample E2: Share of data collection methods

households are interviewed face-to-face by means of PAPI and 39% by means of CAPI. In the second wave the proportion of PAPI-based face-to-face interviews declines to 40% and CAPI is used in 41% of all household interviews. We can observe that the proportion of the PAPI-based face-to-face interviews declines from 49% in the first wave to 19% in Wave 5 on the household level and from 42% to 17.5% on the individual level. At the same time, the proportion of computer-assisted personal interviewing increases from 39% in Wave 1 to 57% in Wave 5 on the household level and from 36.8% to 54.6% on the individual level.

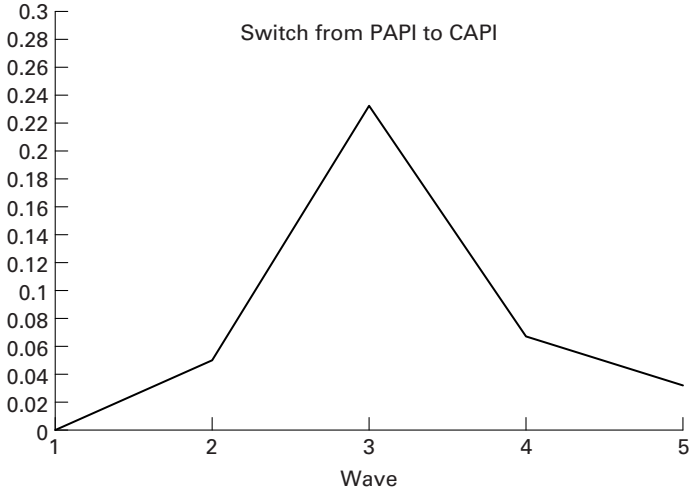

Fig. 3. SOEP, Sample E: Share of mode change from PAPI to CAPI

Table 3. Development of the data collection methods in Sample E, 1998–2002

	1998		1999		2000		2001		2002	
	n	%	*n*	%	*n*	%	*n*	%	*n*	%
Households										
PAPI-based methods										
Face-to-face	522	49.4	358	40.4	130	15.4	135	16.6	149	19.3
Self-completed	47	4.5	84	9.5	84	10.0	81	10.0	90	1.6
Mixed	31	2.9	21	2.4	10	1.2	9	1.1	18	2.3
Mail	38	3.6	57	6.4	68	8.1	73	9.0	78	10.1
CAPI	414	39.2	363	41.0	547	65.0	510	62.9	438	56.7
d.k.	4	0.4	3	0.3	3	0.4	3	0.4	–	–
Total	1,056	100	886	100	842	100	811	100	773	100
Persons										
PAPI-based methods										
Face-to-face	808	42.3	589	36.2	237	15.3	225	15.4	240	17.5
Self-completed	189	9.9	232	14.2	180	11.6	174	11.9	208	15.1
Mixed	142	7.4	51	3.2	44	2.8	29	2.0	50	3.6
Proxy	2	0.1	1	0.1	1	0.1	1	0.1	–	–
Mail	40	2.1	97	6.0	116	7.5	119	8.1	125	9.1
CAPI	703	36.8	647	39.7	958	61.8	913	62.4	750	54.6
d.k.	26	1.4	10	0.6	13	0.8	3	0.2	–	–
Total	1,910	100	1,629	100	1,549	100	1,464	100	1,373	100

Source: SOEP, Sample E, 1998–2002.

3.2. Field Experiences with CAPI in Sample E

Although the response rate in the CAPI subsample E2 is, at 51.9%, slightly lower than that in the PAPI sample E1 (54.1%) in Wave 1 (see Table 1), we cannot conclude that respondents exhibit a more reserved attitude toward interviews conducted using laptops. The results show that the decision to participate or not to participate is made before the interviewer has unpacked the laptop (Infratest 1998). Furthermore no problems about respondents' acceptance of CAPI are reported by the interviewers. And the difference between the two response rates is not significant.

One advantage of computer-assisted interviewing is that some errors like routing mistakes are not possible. Some data and consistency checks that are normally done after the data collection can be done automatically in CAPI during the interview process. Therefore the editing group of the fieldwork organization has had less work with the CAPI data set in the SOEP.

Another important finding is that in large households with many respondents, the interview process as a whole takes more time with CAPI than with a flexible mix of traditional PAPI methods, where self-completion of questionnaires is allowed either in the presence or in the absence of the interviewer. Table 4 shows the percentages of CAPI interviews in the CAPI split sample E2 by the number of respondents in the household. We can see that with increasing numbers of respondents, interviewers increasingly used PAPI-based methods. It can be assumed that the time required for CAPI interviewing of all household members is the reason behind this tendency.

4. Mode Effects on Data Quality

As mentioned above, one reason to move from PAPI to CAPI is the expectation of data quality improvements based on several different calculations. However, even in the best case this could create survey artefacts due to mode effects in the SOEP and could create a break in time series within the longitudinal study. In this section we use some key indicators to examine data collection mode effects in sample E. These indicators are unit nonresponse, missing values, and gross income nonresponse.

4.1. Hypotheses

Based on recent results in the literature and the first fieldwork experiences of the SOEP group with the move from PAPI to CAPI, we derive three hypotheses:

Table 4. Data collection method in the CAPI split sample E2 by number of respondents in the household

	Number of respondents in household			
	1	2	3	4
CAPI	75.2	75.4	63.3	55.6
Self-completed	3.2	4.8	13.3	16.7
Other methods	21.6	19.8	23.4	27.7
Total	100	100	100	100

Source: SOEP Sample E, 1998.

Respondent acceptance: Baker (1992) describes broad respondent acceptance in the case of CAPI and Groves et al. (2004) report that there is no evidence that the technology used affects response rates. Moreover, no problems regarding respondent acceptance with CAPI are reported by the SOEP interviewers. We can assume that respondents who were asked to respond to the survey using CAPI, but were unhappy with this method, will refuse to participate in the following wave. On this basis, we derive our first hypothesis: we can assume that there are no significant differences between PAPI and CAPI for the probability of nonparticipation in the following wave. However, the effect may be small.

Implausible and missing values: Some studies report that CAPI may reduce routing errors due to the use of consistency checks during the interview process. Our second hypothesis is that CAPI interviews have a lower number of implausible values than PAPI. However, the SOEP is edited carefully, so a nonsignificant effect is also possible.

Willingness to disclose sensitive information: Baker et al. (1995) and de Leeuw et al. (1995) report a greater willingness of respondents to disclose sensitive information when using CAPI. They assume that respondents are not concerned about having their information stored on the computer. Monthly income is one of these sensitive items. Therefore we can assume that we will not find significant differences between income nonresponse rates for CAPI and PAPI. However, because there has been much public discussion about privacy issues in computer databases in Germany, there could be a significant effect here as well.

4.2. Respondent's Acceptance – Unit Nonresponse

In this subsection we examine the probability of participation in the next wave after a CAPI interview took place. Unit nonresponse (nonparticipation) is given when respondents are unsuccessful (ill), deceased (dead), or unwilling (refusing) to participate in the survey.[7] A few households could not be found during the fieldwork. Table 5 shows the frequencies of these categories in Waves 2 up to 5 in Sample E.[8] Interviewers classify over 80 percent of this attrition as unwilling respondents and refusals. Again, note that we restrict our nonresponse analysis to respondents who participate in at least one wave. We concentrate our further analysis on unwilling respondents and exclude the independent categories of ill, household not found, moved abroad and deceased, which could dilute real differences.

Figure 4 shows the proportion of unwilling respondents in subsamples E1 and E2 whose questionnaires were administered by an interviewer (not mail) and who dropped out of the survey. We see that there are only small differences between the two subsamples. The share of unwilling lost respondents declines over the first few waves, and is slightly lower in subsample E1 than in subsample E2.

To show the influence of earnings-related institutions and occupations on unit nonresponse we use the classification in Table 6 and classify occupations into three groups. These groups are defined by the type of position (wage, salary, or civil service) and

[7] In the case of "unwilling" the respondent refused only temporarily in this wave and we can contact him/her again in the next wave, whereas in the case of "final refusal" there will be no additional attempts to make contact.
[8] Further details about panel attrition and sample sizes in the SOEP can be found in Kroh/Spieß (2006).

Table 5. Attrition in Sample E

	Wave									
	2		3		4		5		Total	
	N	%	*N*	%	*N*	%	*N*	%	*N*	%
Unsuccessful at the time (e.g., sick)	13	3.4	9	3.6	13	5.9	5	2.1	40	3.6
Unwilling	92	24.2	72	28.5	74	33.8	96	39.7	334	30.5
Final refusal	271	71.1	146	57.7	110	50.2	116	47.9	643	58.7
Dead			15	5.9	13	5.9	20	8.3	48	4.4
HH not found	5	1.3	11	4.4	9	4.1	5	2.1	30	2.7
Total	381	100	253	100	219	100	242	100	1,095	100

Source: SOEP, Sample E, 1998–2002.

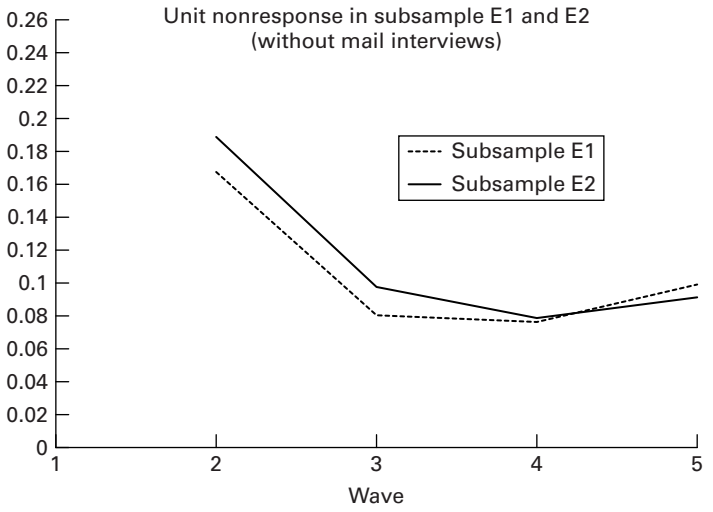

Fig. 4. SOEP, Subsamples E1 and E2: Share of lost respondents (without mail interviews)

occupational skills. Figures 5 and 6 show the unit nonresponse rate by the mode of data collection used in the previous wave and by the respondents' vocational position. A fairly inconsistent pattern can be seen in Figure 6. After a strong decrease in drop-outs for respondents in low earning positions from 20% in Wave 2 to approximately 5% in Waves 3 and 4, the rate increases again to 20% in Wave 5. The rates for medium and high earners decrease moderately from 18% to 11% in Wave 5. Schräpler (2004) has shown that in sample A of the SOEP the mail mode is a strong indicator for cooperation problems and that respondents who answer by mail often drop out of the survey in the next wave. The graph in Figure 5 shows a similar pattern: mail interviews have higher unit nonresponse rates in the following wave than the other modes.[9] Furthermore, it seems that CAPI performs slightly better than PAPI, because the rate of lost respondents declines from wave to wave and is only 8% in Wave 5 whereas for PAPI the unit nonresponse rate increases to 16% in the last wave. In Wave 5, these increasing rates could indicate serious cooperation problems for respondents in low occupational states, and the effect of the PAPI mode is seen in Figures 5 and 6. For the explanation of unit nonresponse and the impact of the interview mode we estimate multilevel logit regression models.

4.2.1. Modeling Unit Nonresponse

After participation in the survey, the respondent chose between two alternatives, participation or nonparticipation in the following wave. We use a regression framework to determine how various factors influence the attractiveness of the alternatives to different

[9] But we have to interpret this finding as a compound effect: the higher drop-out rates in the case of mail interviews are probably not due to the mode alone, i.e., respondents do a mail interview because they are less cooperative, and it is this that determines their subsequent response.

Table 6. Classification of the vocational position

	Vocational position	Occupation
LOW	Hourly-paid worker	Unskilled worker, semiskilled worker
MEDIUM	Hourly-paid worker	Skilled worker, foreman, master,
	Salaried employee	Industry and works foreman, employee with simple activity, skilled activity
	Civil servant[a]	Minor and lower-grade civil service
HIGH	Salaried employee	Highly skilled activity, executive function
	Civil servant[a]	High and senior service

[a] Civil servant includes also government officials.

types of individuals (cf. Dubin and Rivers 1989:373). The difference between the utility of two alternatives is

$$y_i^* = \tilde{U}_{i1} - \tilde{U}_{i2}$$

If $y_i^* > 0$, the first alternative (refuse to participate) yields higher utility and is preferred; otherwise, the second one is preferred.

Because of the binary response character, we can use a probit or an ordinary logit model. The hierarchical structure of the survey data suggests the use of multilevel models as the appropriate method of analysis (cf. Hill 1991; Hox 1994).

Level 1 consists of i respondents and Level 2 represents the aggregate level, which is formed by j interviewers. Hence for respondent i and interviewer j one dichotomous variable y_{ij} is observed:

$$y_{ij} = \begin{cases} 1, & \text{if } y_{ij}^* > 0, \text{ refuse to participate (next wave)} \\ 0, & \text{otherwise} \end{cases}$$

$$y_{ij} = \pi_{ij} + u_{ij}$$

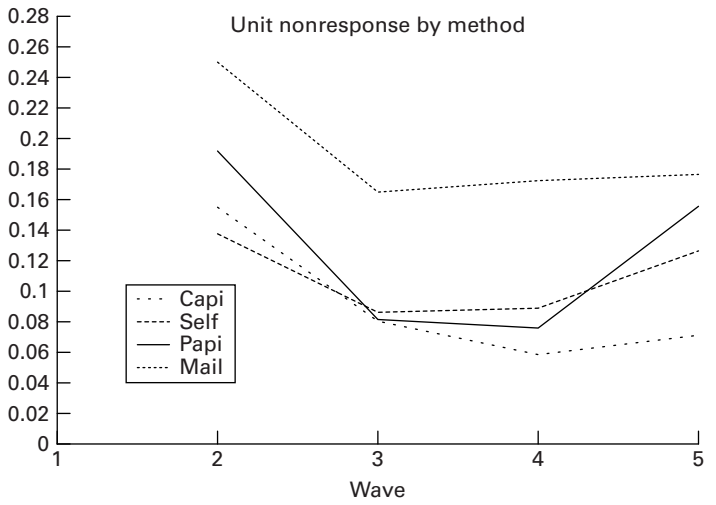

Fig. 5. SOEP, Sample E: Share of unit nonresponse by data collection method

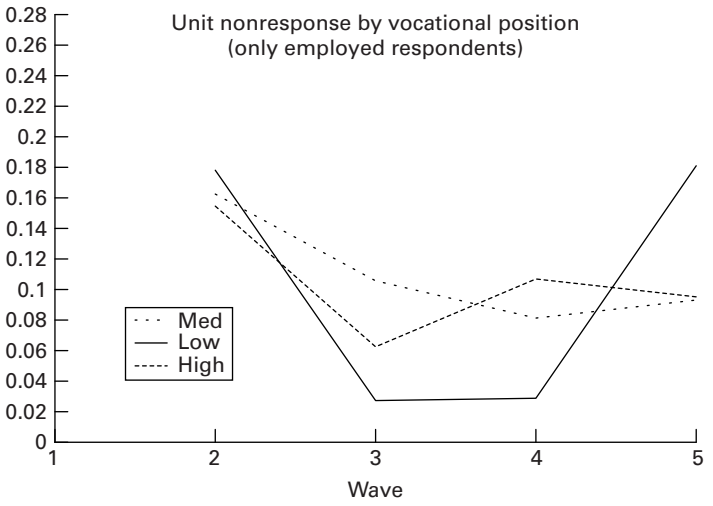

Fig. 6. SOEP, Sample E: Share of unit nonresponse by vocational position

If we specify a two-level random intercept model (Model 1) the probability π_{ij} for each response is estimated from:

$$\pi_{ij} = \left[1 + \exp\left(-\left(\beta_{0j} + \sum_{h=1}^{H} \beta_{h,ij} x_{h,ij} + v_{0j}\right)\right)\right]^{-1} \tag{1}$$

where $x_{h,ij}$ represents values for covariates $x_h (h = 1, \ldots, H)$ of respondent i and interviewer j. The intercept β_{0j} is specified as random on Level 2 (interviewer level) and the variance is estimated as v_{0j}. The random variation among the respondents on Level 1 is estimated as the variance u_{ij}.

Up to this point, we have assumed that the effects of the explanatory variables are the same for each interviewer. We will now modify this assumption in Model 2 by allowing the effect of the CAPI mode to vary across interviewers. Therefore, we have to introduce a random coefficient for CAPI.

$$\beta_{capi,j} = \beta_{capi} + v_{capi,j} \tag{2}$$

Hence we specify a two-level random coefficient model (Model 2) with the probability π_{ij}:

$$\pi_{ij} = \left[1 + \exp\left(-\left(\beta_{0j} + \sum_{h=1}^{H} \beta_{h,ij} x_{h,ij} + x_{capi} v_{capi,j} + v_{0j}\right)\right)\right]^{-1} \tag{3}$$

where $v_{capi,j}$ is a normally distributed random effect with mean zero and variance $\sigma^2_{v,capi}$

$$\begin{bmatrix} v_{0j} \\ v_{capi,j} \end{bmatrix} \sim N(0, \Sigma_R) : \Sigma_R = \begin{bmatrix} \sigma^2_{v,0} & \\ \sigma_{v,capi0} & \sigma^2_{v,capi} \end{bmatrix}$$

Allowing the coefficient of CAPI to vary across interviewers has also introduced the parameter $\sigma_{v,capi0}$ which is the covariance between v_{0j} and $v_{capi,j}$.

4.2.2. Regressors of the Unit Nonresponse Model

Regressors can be considered in three groups:

1. *Demographic and household variables for the respondent*: "age" is the age of the respondent in years, "sex" = 1 indicates male respondent (dummy), "low occup.", "med. occup." and "high occup." as well as "not empl./trainees", "self-empl." and "milit./civil serv." are dummy variables that indicate the corresponding occupational status (ref. category is "medium occup."), "size of HH." indicates the number of persons living in the respondent's household. "Move" = 1 indicates that the respondent has changed residence in the past 12 months (dummy).
2. *Demographic variables for the interviewer*: "isex" = 1 indicates male interviewer (dummy).[10]
3. *Variables that describe the interview situation:* "CAPI" = 1 indicates a CAPI interview (dummy), "self-completed" = 1 indicates a self-completion mode of response in the presence of the interviewer (dummy), "mixed" = 1 indicates a mixed mode (dummy), "change of interviewer" = 1 indicates a change in interviewer (dummy), "switch to CAPI" = 1 indicates the switch from PAPI to CAPI (dummy, Wave 2 upwards) and "Sample E1" = 1 indicates the subsample E1 (dummy, Wave 2 upwards).[11]

4.2.3. Estimates

Table 7 shows estimates of two univariate logit models for Waves 1 to 4.[12] Model 1 is a random intercept model where only the intercept is allowed to vary across the interviewers. Model 2 is a random coefficient model, where we also allow this variation for the slope for CAPI. The sample contains a total of 1,583 respondents who participated in Wave 1, with 110 interviewers. In the following the samples in Waves 2 to 4 decline due to attrition.

The estimates of Models 1 and 2 in the first wave show no significant effects of respondent characteristics on the probability of refusals. But we find a strong positive significant effect on unit nonresponse for moving respondents (move) and also for the change of the interviewer (change of int.).

We were interested mainly in mode effects. But we have to interpret the results with caution. We cannot measure pure mode effects because our experimental design in Wave 1 is not implemented completely and in the following waves we have compound effects with other response-related characteristics and interviewer behavior.

Our first hypothesis states that we will not find significant differences between the coefficients for PAPI and CAPI. Although the coefficient for CAPI is negative in all waves, we find significantly better performance than the reference category PAPI only in the first and the fourth wave. In Wave 4, the PAPI mode has the highest attrition rate of all

[10] For further details of available data about interviewer characteristics in the SOEP see Schräpler and Wagner 2001.
[11] We did not introduce the dummy variable "Sample E1" in Wave 1 because of possible multicollinearity problems with the variable CAPI.
[12] The analysis is done with MLwiN 2.02 (Rasbash et al. 1999). We used the iterative generalized least squares (IGLS) algorithm.

Table 7. Multilevel Logit-model for refusal in the following wave, Model 1 – random intercept model; Model 2 – random coefficient model

	Wave 1				Wave 2			
	Model 1		Model 2		Model 1		Model 2	
	$\hat{\beta}$	S.E.	$\hat{\beta}$	S.E.	$\hat{\beta}$	S.E.	$\hat{\beta}$	S.E.
Fixed								
Intercept	−2.536***	0.46	−2.546***	0.48	−1.286	1.09	−1.011	1.13
Respondent								
Sex (1 – men)	−0.058	0.15	−0.081	0.16	−0.011	0.21	−0.004	0.22
Age (year)	0.007	0.01	0.005	0.01	−0.016*	0.01	−0.014	0.01
Med. (ref)								
Low occup.	−0.112	0.34	−0.078	0.35	−1.208*	0.64	−1.168*	0.64
High occup.	0.255	0.28	0.349	0.29	−0.832*	0.46	−0.734	0.47
Not empl./trainees	0.424	0.37	0.378	0.38	0.660	0.43	0.629	0.45
Self-empl.	0.138	0.35	0.280	0.37	0.066	0.41	0.165	0.43
Military/civil serv.	0.014	0.21	0.052	0.21	−0.315	0.27	−0.322	0.28
Size of HH	−0.063	0.07	−0.077	0.07	−0.296**	0.11	−0.265**	0.11
Move	0.790*	0.46	0.942**	0.47	1.437***	0.50	1.549***	0.52
Interviewer								
Isex (1 – men)	0.876**	0.28	0.957***	0.29	−0.143	0.30	−0.151	0.29
Situation								
Change of interviewer	0.496*	0.29	0.452	0.30	2.175***	0.33	2.071***	0.34
Sum of participation					0.304	0.48	0.059	0.49
Papi (ref)								
Capi	−0.298*	0.17	−0.316	0.26	−0.417	0.44	−1.011	1.13
Self completed	0.044	0.32	−0.017	0.35	−0.030	0.36	−0.181	0.40
Mixed	−0.199	0.34	−0.366	0.37	0.235	0.58	0.119	0.61
Sample E1					−0.365	0.40	−0.467	0.43
Switch to CAPI					0.136	0.51	0.157	0.56

Table 7. Continued

	Wave 1				Wave 2			
	Model 1		Model 2		Model 1		Model 2	
	$\hat{\beta}$	S.E.	$\hat{\beta}$	S.E.	$\hat{\beta}$	S.E.	$\hat{\beta}$	S.E.
Random								
Respondent level								
σ_u^2	$\pi^2/3$		$\pi^2/3$		$\pi^2/3$		$\pi^2/3$	
Interviewer level								
σ_v^2	1.018***	0.24	1.520***	0.39	0.983***	0.30	1.692***	0.55
$\sigma_{v,capi}^2$			2.525***	0.84			3.298**	1.25
$\sigma_{v,capi0}$			−1.352**	0.49			−1.913***	0.74
Interviewer cluster	110		110		115		115	
Persons	1,583		1,583		1,477		1,477	

	Wave 3				Wave 4			
	Model 1		Model 2		Model 1		Model 2	
	$\hat{\beta}$	S.E.	$\hat{\beta}$	S.E.	$\hat{\beta}$	S.E.	$\hat{\beta}$	S.E.
Fixed								
Intercept	−3.158***	1.13	−3.280	1.23	−3.127***	0.96	−3.196***	0.97
Respondent								
Sex (1 – men)	0.238	0.26	0.176	0.26	0.036	0.26	0.066	0.26
Age (year)	−0.020*	0.01	−0.018*	0.01	−0.013	0.01	−0.013	0.01
Med. (ref)								
Low occup.	−0.228	0.72	−0.343	0.73	1.425**	0.49	1.424***	0.49
High occup.	0.531	0.50	0.609	0.52	0.638	0.49	0.650	0.49
Not empl./trainees	−0.454	0.75	−0.460	0.76	0.479	0.60	0.443	0.61

Table 7. Continued

	Wave 3				Wave 4			
	Model 1		Model 2		Model 1		Model 2	
	$\hat{\beta}$	S.E.	$\hat{\beta}$	S.E.	$\hat{\beta}$	S.E.	$\hat{\beta}$	S.E.
Self-empl.	1.044*	0.52	0.899*	0.53	0.803	0.51	0.782	0.51
Military/civil serv.	0.874**	0.34	0.794**	0.34	0.189	0.35	0.193	0.36
Size of HH	−0.353**	0.14	−0.371**	0.14	0.358**	0.12	0.343**	0.12
Move	2.253***	0.49	2.284***	0.51	0.871	0.70	0.867	0.72
Interviewer								
Isex (1 – men)	0.551	0.42	0.402	0.40	0.216	0.42	0.226	0.43
Situation								
Change of interviewer	3.388***	0.35	3.583***	0.35	4.299***	0.43	4.285***	0.45
Sum of participation	0.077	0.31	0.118	0.34	−0.203	0.18	−0.205	0.18
Papi (ref)								
Capi	−0.082	0.44	0.066	0.51	−1.053**	0.44	−0.974**	0.48
Self completed	0.394	0.53	0.125	0.62	−0.219	0.51	−0.020	0.52
Mixed	1.593**	0.61	1.587**	0.69	−2.587	1.70	−2.413	1.68
Sample E1	0.381	0.32	0.405	0.33	0.059	0.27	0.064	0.28
Switch to CAPI	−1.044**	0.46	−0.995**	0.45	0.143	0.67	0.384	0.68
Random								
Respondent level								
σ_u^2	$\pi^2/3$		$\pi^2/3$		$\pi^2/3$		$\pi^2/3$	
Interviewer level								
σ_v^2	2.045***	0.55	4.486***	1.41	2.023***	0.57	1.998*	0.86
$\sigma_{v,capi}^2$			1.456	1.58			2.775	2.26
$\sigma_{v,capi0}$			2.510*	1.30			−1.141	0.13
Interviewer cluster	129		129		134		134	
Persons	1,420		1,420		1,340		1,340	

Source: SOEP Sample E, individual questionnaire, 1998–2002; significance: *10%; **5%; ***1%.

interview modes. In addition to this main effect we find that the CAPI coefficient varies significantly between the interviewers in the random coefficient Model 2 $\left(\sigma^2_{v,capi}\right)$ in Waves 1 and 2. This means that the impact of the CAPI mode at time t on the participation in the following Wave $t + 1$ depends on the interviewer's performance especially in the first two waves. We can assume that this finding is caused by interviewer's skill in managing the new data collection method. Interviewers who are confident with the new technique might act in a more trustworthy manner than interviewers who lack expertise in the use of CAPI.

The main shift from PAPI to CAPI occurs in Wave 3. An interesting result is that in Wave 3 the coefficient for the switch shows a significant decrease in the probability of nonparticipation in the next wave. It seems that the change in mode had no negative effect on the participation rate. The coefficient of the dummy variable for subsample E1 indicates no significant effect in any of the waves.

In addition we find a gender interviewer effect: male interviewers lost significantly more respondents after the first wave than female interviewers. Besides this identifiable systematic effect, we find significant interviewer/area variances σ^2_v in all waves and significant covariances $\sigma_{v,capi0}$ between the interviewer and the CAPI variance $\sigma^2_{v,capi}$ in the first two waves.

4.3. Item Nonresponse

In the literature, some studies can be found (Sebestik et al. 1988; Olsen 1992) that report lower levels of missing data in the case of CAPI. They assume that CAPI avoids routing errors and implausible values. Routing errors are not caused by cooperation or cognitive problems such as refusals and "don't knows." The SOEP distinguishes between implausible values and other types of missing values in Sample E. Therefore we can explore whether there are differences in these rates by varying data collection modes.

4.3.1. Missing Values and Implausible Values

Table 8 shows the average number of missing values in the individual questionnaires by the data collection method. Because employed persons have to answer more questions than unemployed persons, we calculate the average number for employed respondents separately. The values in the table show a rather consistent result: the average number of missing values is highest in the case of employed respondents and mail and self-completed questionnaires, and lowest in the case of face-to-face interviews. CAPI interviews lie somewhere in between these groups.

Next we looked at the average number of implausible values in the questionnaires. Implausible values may be a result of coding errors caused either by untrained interviewers who enter wrong values in the questionnaires or by confused respondents who do not understand the question and answer in an invalid way. Well-trained interviewers should be able to detect these implausible values and call attention to them. Furthermore a well-programmed CAPI system should be able to detect values that are out of range automatically and should indicate this on the screen of the laptop immediately.

Therefore we can assume that CAPI interviews will have lower rates of implausible values than face-to-face interviews. Table 9 shows the average number of implausible

Table 8. Average number of missing values in the individual questionnaires in Sample E by method

	Wave 1		Wave 2		Wave 3		Wave 4		Wave 5	
Missing values	All	Empl.	All	Empl.	All	Empl.	All	Empl.	All	Empl.
PAPI	5.28	5.88	1.97	2.81	1.66	2.14	2.60	3.14	2.52	2.97
Self	6.76	7.44	2.89	2.94	2.53	3.13	3.44	3.71	3.41	3.69
Mail	6.48	6.10	3.66	3.91	5.72	6.17	4.91	5.38	5.22	5.44
CAPI	5.46	5.87	2.84	3.37	1.99	2.38	2.57	3.22	4.15	4.87
Total	5.68	6.25	2.61	3.22	2.38	2.97	2.97	3.58	3.82	4.37

Source: SOEP, Sample E, individual questionnaire, 1998–2002 (own calc.).

Table 9. Average number of implausible values in the individual questionnaires in Sample E by method

	Wave 1		Wave 2		Wave 3		Wave 4		Wave 5	
	All	Empl.	All	Empl.	All	Empl.	All	Empl.	All	Empl.
PAPI	0.07	0.09	0.12	0.12	0.11	0.11	0.04	0.04	0.04	0.05
Self	0.06	0.05	0.13	0.11	0.11	0.12	0.15	0.17	0.08	0.09
Mail	0.03	0.03	0.09	0.07	0.06	0.06	0.13	0.04	0.09	0.10
CAPI	0.11	0.12	0.06	0.06	0.04	0.05	0.03	0.04	0.03	0.05
Total	0.09	0.10	0.10	0.09	0.07	0.08	0.06	0.06	0.05	0.07

Source: SOEP, Sample E, individual questionnaire, 1998–2002 (own calc.).

values in the individual questionnaires by the interview mode used. We see that the total maximum rate of implausible values is only 0.1% in Waves 1 and 2.

In addition we see that – with the exception of Wave 1 – CAPI reduces the number of implausible values in the data set. In Waves 2 and 3 the average number for CAPI is half of the average number for face-to-face. In Waves 4 and 5, the two have nearly the same low rates. We can assume that CAPI has a higher rate in the first wave because of transposition problems. The software used has to be adjusted. Overall it seems that CAPI is the best mode for avoiding implausible values.

4.3.2. Willingness to Disclose Sensitive Information – Income Nonresponse

In this section, we explore whether the CAPI mode has a significant effect on respondents' decisions to reveal their earnings. A detailed conceptual and empirical explanation of the reasons for income nonresponse is given in Schräpler (2004, 2006). Our comparative study reveals that the same patterns for "refusals" and "don't knows" occur in the SOEP as in the BHPS, and we have shown that it is important to distinguish between the two types of missing values. We do not want to repeat our conceptual framework and empirical results, but do have to repeat some statistical procedures.

Table 10 shows the income nonresponse rate for the gross income question of employed persons in Sample E. We exclude self-employed persons and trainees in our analysis.

Table 10. Item Nonresponse rates for the gross income question among employed persons in the SOEP, Sample E (in percent)

	Including self-employed and trainees			Excluding self-employed and trainees		
Wave	Employed respondents	Missing	%	Selected respondents	Missing	%
1	1,032	272	26.4	870	206	23.7
2	886	167	18.8	736	113	15.4
3	858	151	17.6	716	106	14.8
4	805	153	19.0	658	95	14.4
5	746	131	17.6	613	89	14.5
Total	4,327	874	20.2	3,593	609	16.9

Source: SOEP, Sample K, 1998–2002 (own calc.).

Table 11. Gross Income Nonresponse rate by data collection method in Sample E, employed persons

	Wave					
Method	1	2	3	4	5	*N*
PAPI	21.4	13.0	12.9	4.1	8.3	864
Mixed	14.6	12.5	14.8	28.6	23.0	173
Self-completed	22.8	12.9	8.9	12.1	1.6	546
Mail	22.2	15.0	23.3	16.2	22.1	305
CAPI	27.2	18.6	15.5	16.9	16.0	1,676
N	870	736	716	658	613	3,593

Source: SOEP, Sample E, 1998–2002 (own calc.).

The nonresponse rate is, at 23.7%, highest in the first wave, declines to 15.4% in the second wave and then remains relatively constant between 14% and 15%.

Table 11 shows the income nonresponse rate by data collection mode used. We see that CAPI interviews have the highest rates of all the modes in the first two waves. This finding suggests that respondents have some reservations regarding computer-based interviewing on their first encounter with it. Moreover, CAPI interviews always have higher rates than PAPI interviews.

Because refusals are not distinguished from "don't knows" in the SOEP, we have to use the same approach as described in Schräpler 2004 in the following. Table 12 shows the cross-tabulation of missing gross and net income, pooled over five waves. We have already established that it is reasonable to assume that respondents who do not state their gross income but do state their net income have cognitive problems in the majority of cases, and that we can classify this behavior as a "don't know" answer. In cases where respondents state neither their gross nor their net income, it is reasonable to assume that they are more or less uncooperative and that we can classify this as a refusal (see Schräpler 2006).

Table 12 shows that – under these presumptions – the refusals are, at 10%, slightly higher than the "don't knows," at 7%. Figures 7 and 8 display the income nonresponse rates by data collection methods. We can see that CAPI has, in four of five waves, the highest and PAPI (face-to-face) in all cases the lowest refusal rate. Furthermore PAPI shows a higher variation of the "don't know" rate than CAPI.

Figures 9 and 10 explore mode-induced differences in the refusal rate separated by respondents' gender. We see that the refusal rates are highest for male respondents in the

Table 12. Missing gross and net income in Sample E, Wave 1–5

	Net income					
	Valid		Missing			
Gross income	*N*	%	*N*	%	Total	%
Valid	2,831	78.8	149	4.2	2,980	83.0
Missing	249	7.0	360	10.0	609	17.0
Total	3,080	85.8	509	14.2	3,589	100.0

Source: SOEP, Sample E, 1998–2002 (own calc.).

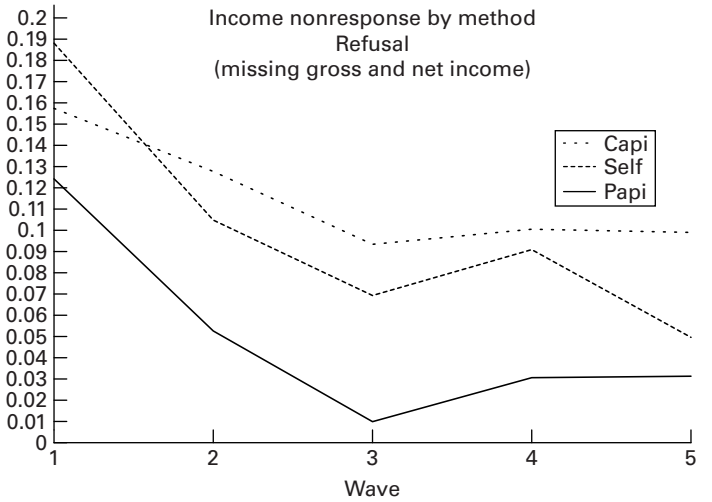

Fig. 7. SOEP, Sample E: Share of income nonresponse (refusals) by data collection method

self-completion mode and for female respondents in the CAPI mode. Nevertheless, we do not find gender differences for the nonresponse rates in CAPI. It seems that the gender effect lies in the fact that male respondents use the self-completion mode more often for refusing than do female respondents. We have shown before that the interviewer has less control over the interview process if the respondents fill out their questionnaires by themselves (in front of the interviewer). In this situation it is much easier to skip an unpleasant statement.

Figures 11 and 12 show the influence of occupational positions on income nonresponse. Respondents in high earning positions tend to refuse to reveal their income whereas respondents in low occupational states have higher rates of "don't knows."

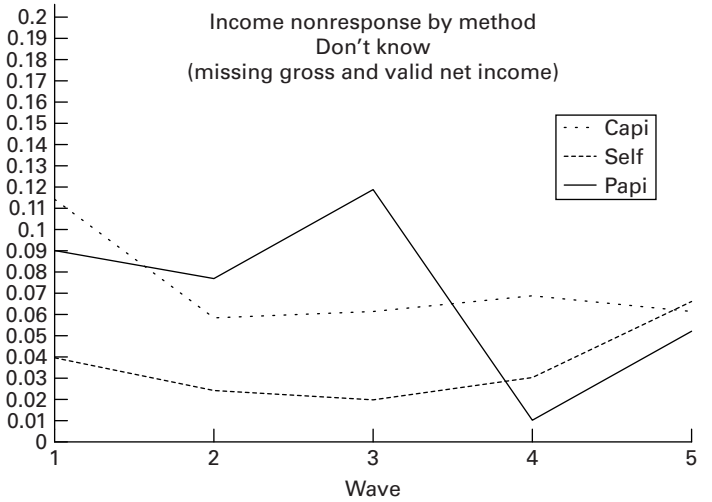

Fig. 8. SOEP, Sample E: Share of income nonresponse (don't know) by data collection method

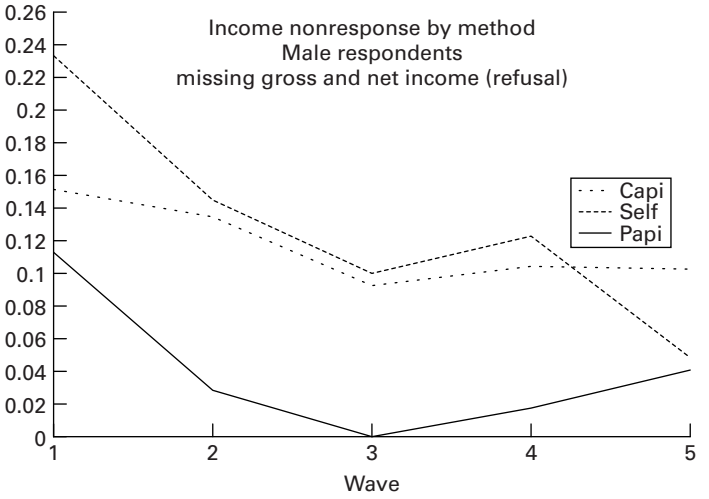

Fig. 9. SOEP, Sample E: Share of refusals by data collection method and male respondent

Modeling Income Nonresponse

We estimate two logit models separately for Waves 1, 2 and 3.[13] First a univariate logit model for the indicator "income nonresponse" and second, a multivariate logit model with three response variables "refuse," "don't know" and "unit response in the following wave."[14] The difference in utility for each response alternative is described with y_i^*. If $y_i^* > 0$, the first alternative yields higher utility and is preferred; otherwise, the second one is preferred. Again we account for the hierarchical structure of the survey data and use a multilevel model. Level 1 represents the different response variables in the multivariate model, Level 2 represents j respondents and Level 3 consists of k interviewers. Hence we estimate a multivariate logit model with three levels:

For respondent j and interviewer k one dichotomous variable y_{ijk} is observed:

$$y_{ijk} = \pi_{ijk} + u_{ijk}$$

$$y_{1jk} = \begin{cases} 1, & \text{if } y_{1jk}^* > 0, \text{ refuse} \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

$$y_{2jk} = \begin{cases} 1, & \text{if } y_{2jk}^* > 0, \text{ don't know} \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

$$y_{3jk} = \begin{cases} 1, & \text{if } y_{3jk}^* > 0, \text{ unit-response (next wave)} \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

[13] The analysis is done with MLwiN 2.02 (Rasbash et al. 1999). We used the iterative generalized least squares (IGLS) algorithm.

[14] A similar model for income nonresponse with a probit specification can be found in Schräpler 2004.

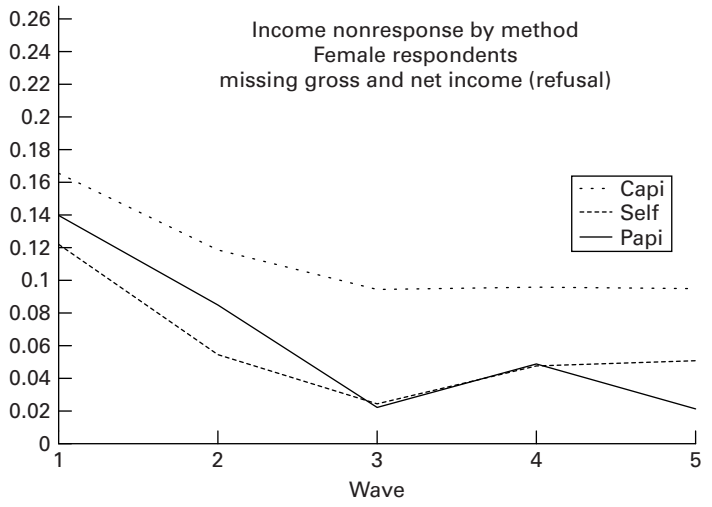

Fig. 10. SOEP, Sample E: Share of refusals by data collection method and female respondent

The probability π_{ijk} for each response variable i is estimated from:

$$\pi_{ijk} = \left[1 + \exp\left(-\left(\beta_{0k} + \sum_{h=1}^{H} \beta_{h,ijk} x_{h,ijk} + v_{0k}\right)\right)\right]^{-1} \tag{7}$$

where $x_{h,ijk}$ represents values for covariates $x_h\,(h = 1, \ldots, H)$ of respondent j and interviewer k. The intercept is specified as random on level 3 (interviewer level) and the variance is estimated as v_{0k}. The random variation among the respondents on level 2 is estimated as the variance/covariance u_{ijk}. If they are dependent binomial variables, we have to estimate the residual variances $\sigma^2_{u_{ii}}$ and covariances $\sigma_{u_{ii'}}$.[15]

The regressors we used to explain income nonresponse are the same as in the section before. The only difference is that we restrict our sample to employed respondents and exclude self-employed persons and trainees.

Estimates

Tables 13–15 show estimates of the univariate and the multivariate logit models for Waves 1 to 3. The sample in Wave 1 contains a total of 702 employed respondents from 106 interviewers. In Wave 2 the sample size declines to 652 and in Wave 3 to 637 employed respondents caused by the attrition process. The number of interviewers increases to 110 in Wave 2 and 119 in Wave 3.

The first column (0) in the tables refers to the univariate logit model (Model 1) with gross income nonresponse as response variable. In this model we can see a consistent significant positive effect in Waves 1 and 2 for CAPI in the fixed part of the model. This means that the CAPI mode produces more missing values for gross income than a PAPI based face-to-face mode. Furthermore, in Model 2 these missing values are separated into

[15] The correlation between the residual variance of "refuse" and "don't know" on Level 2 has to be restricted to zero because the respondents can choose only one of the two alternatives (see Schräpler 2004).

Fig. 11. SOEP, Sample E: Share of refusals by occupation

"refuse" (missing for gross and net income) and "don't know" (missing for gross and valid answer for net income). The estimates show that in all three waves, the CAPI mode has a strong positive effect on the category "refuse" but no significant effect on "don't know." Interviewers who use computer-assisted personal interviewing in Sample E of the SOEP have a higher probability of having respondents refuse to state their gross and net income than interviewers using the traditional PAPI mode. This finding rejects our third hypothesis and is also not in line with the previous findings of Baker et al. (1995) and de Leeuw et al. (1995). It seems that at least in their very first contacts, CAPI respondents in Sample E have more of a problem disclosing their income statement than those using PAPI. But it also cannot be ruled out that these results are caused by the interviewer.

Fig. 12. SOEP, Sample E: Share of don't knows by occupation

Table 13. Multiv. multilevel logit model for income nonresponse, sample E, Wave 1

	Model 1 – w1		Model 2 – w1					
	(0)		(1)		(2)		(3)	
	Item nonresponse		Refuse		Don't Know		Unit Response $(t+1)$	
	$\hat{\beta}$	S.E.	$\hat{\beta}_1$	S.E.	$\hat{\beta}_2$	S.E.	$\hat{\beta}_3$	S.E.
Fixed								
Intercept	−2.190***	0.548	−2.944***	0.362	−3.467***	0.635	2.448***	0.541
Respondent								
Sex (1 – men)	−0.235	0.194	−0.077	0.203	−0.494***	0.204	−0.270	0.195
Age (year)	0.009	0.009	0.005	0.010	0.013	0.010	−0.005	0.009
Med occup. (ref)								
Low occup.	0.408	0.279	−0.140	0.321	0.868***	0.272	0.097	0.289
High occup.	0.164	0.246	0.505***	0.241	−0.603*	0.307	−0.210	0.237
Size of HH	0.007	0.085	0.068	0.088	−0.047	0.094	0.182***	0.083
Move	−0.349	0.852	0.087	0.743	0.000	0.000	−0.998*	0.603
Interviewer								
Isex (1 – men)	−0.039	0.329	−0.062	0.347	0.088	0.482	−1.018***	0.335
Situation								
Change of interviewer	0.190	0.405	0.620	0.409	−0.479	0.447	−0.291	0.355
Papi (ref)								
Capi	0.510**	0.218	0.482***	0.230	0.357	0.234	0.044	0.211
Self-completed	0.344	0.379	0.785***	0.362	−1.053*	0.577	0.020	0.374
Random								
Respondent level								
	u1		u1		u2		u3	
u1	0.861	0.049	0.666	0.038				
u2			0.000°	0.000	0.404	0.023		
u3			−0.077***	0.028	−0.023	0.021	0.703	0.040

Table 13. Continued

	Model 1 – w1		Model 2 – w1					
	(0)		(1)		(2)		(3)	
	Item nonresponse		Refuse		Don't Know		Unit Response $(t+1)$	
	$\hat{\beta}$	S.E.	$\hat{\beta}_1$	S.E.	$\hat{\beta}_2$	S.E.	$\hat{\beta}_3$	S.E.
Interviewer level								
	v1		v1		v2		v3	
v1	1.259	0.343	1.439	0.382				
v2			−0.522	0.395	2.981	0.718		
v3			0.008	0.258	0.176	0.360	1.210	0.325
Interviewer cluster	106				106			
Persons	702				702			
−2 * LogLikelih	−516.4				−921.5			

Note: ° Constrained to zero; Significance; *10%; **5%; ***1%.
Source: SOEP, Sample E, 1998, employed respondents without self-employed and trainees, without mail interviews (own calc.).

Table 14. Multiv. multilevel logit model for income nonresponse, sample E, Wave 2

	Model 1 – w2		Model 2 – w2					
	(0)		(1)		(2)		(3)	
	Item nonresponse		Refuse		Don't Know		Unit Response ($t+1$)	
	$\hat{\beta}$	S.E.	$\hat{\beta}_1$	S.E.	$\hat{\beta}_2$	S.E.	$\hat{\beta}_3$	S.E.
Fixed								
Intercept	−4.022***	0.761	−4.705***	0.843	−6.607***	0.965	1.481	0.834
Respondent								
Sex (1 – men)	0.019	0.202	−0.012	0.214	−0.246	0.241	−0.063	0.235
Age (year)	0.018*	0.010	0.000	0.010	0.043***	0.012	0.023**	0.011
Med occup. (ref)								
Low occup.	−0.193	0.325	−0.044	0.362	−0.746**	0.346	0.985*	0.471
High occup.	−0.131	0.262	0.182	0.275	−797**	0.351	0.727*	0.341
Size of HH	−0.038	0.095	−0.036	0.105	−0.035	0.111	0.155	0.115
Move	1.445**	0.640	1.680**	0.601	0.000	0.000	−1.288**	0.589
Interviewer								
Isex (1 – men)	0.139	0.444	−0.101	0.497	0.363	0.760	0.030	0.413
Situation								
Change of interviewer	−0.426	0.556	−0.486	0.542	−1.130	1.137	−1.776***	0.457
Papi (ref)								
Capi	1.399**	0.547	2.297***	0.624	0.561	0.564	0.488	0.598
Self-completed	0.640*	0.357	1.397***	0.386	−0.903*	0.458	−0.514	0.375
Switch to CAPI	1.016*	0.516	1.367**	0.506	0.000	0.000	−0.571	0.629
Sample El	0.590	0.501	1.150**	0.570	0.107	0.502	0.120	0.554
Random								
Respondent level								
	u1		u1		u2		u3	
u1	0.537	0.032	0.413	0.025				
u2			0.000°	0.000	0.237	0.014		

Table 14. Continued

	Model 1 – w2		Model 2 – w2					
	(0)		(1)		(2)		(3)	
	Item nonresponse		Refuse		Don't Know		Unit Response ($t+1$)	
	$\hat{\beta}$	S.E.	$\hat{\beta}_1$	S.E.	$\hat{\beta}_2$	S.E.	$\hat{\beta}_3$	S.E.
u3			−0.049	0.019	0.007	0.015	0.493	0.029
Interviewer level								
	v1		v1		v2		v3	
v1	3.165	0.665	3.932	0.831				
v2			0.643	0.897	6.205	1.712		
v3			0.201	0.510	2.217	0.827	2.051	0.549
Interviewer cluster	110				110			
Persons	652				652			
−2 * LogLikelih	−284.7				−3,972.9			

Note: ° Constrained to zero; Significance; *10%; **5%; ***1%.

Source: SOEP, Sample E, 1999, employed respondents without self-employed and trainees, without mail interviews (own calc.).

Table 15. Multiv. multilevel logit model for income nonresponse, sample E, Wave 3

	Model 1 – w3		Model 2 – w3					
	(0)		(1)		(2)		(3)	
	Item nonresponse		Refuse		Don't Know		Unit Response ($t+1$)	
	$\hat{\beta}$	S.E.	$\hat{\beta}_1$	S.E.	$\hat{\beta}_2$	S.E.	$\hat{\beta}_3$	S.E.
Fixed								
Intercept	−3.102***	0.700	−4.494***	0.808	−4.999***	0.885	4.247***	1.410
Respondent								
Sex (1 – men)	0.010	0.215	0.287	0.238	−0.540**	0.238	−0.240	0.298
Age (year)	−0.002	0.011	0.001	0.012	−0.008	0.012	0.067***	0.017
Med occup. (ref)								
Low occup.	0.132	0.340	−0.315	0.419	0.952**	0.350	0.269	0.499
High occup.	−0.084	0.279	0.259	0.285	−1.116**	0.420	−1.143**	0.426
Size of HH	0.049	0.098	0.068	0.108	0.015	0.109	1.055***	0.182
Move	−0.166	0.575	−1.000	0.743	1.042	0.669	−2.443***	0.513
Interviewer								
Isex (1 – men)	−0.042	0.412	−0.328	0.480	0.339	0.700	0.562	1.544
Situation								
Change of interviewer	0.364	0.467	1.169*	0.476	−0.191	0.590	−11.988***	1.052
Papi (ref)								
Capi	0.840**	0.396	1.227*	0.473	0.816	0.502	−0.742	0.767
Self-completed	−0.167	0.440	0.809	0.523	−0.702	0.516	−1.574*	0.739
Switch to CAPI	−1.149**	0.384	−0.805*	0.409	−1.297**	0.459	5.163***	0.801
Sample E1	0.563	0.305	0.141	0.324	1.121***	0.377	−2.199***	0.445
Random								
Respondent level								
	u1		u1		u2		u3	
u1	0.575	0.035	0.425	0.026				
u2			0.000°	0.000	0.252	0.015		

Table 15. Continued

	Model 1 – w3		Model 2 – w3					
	(0)		(1)		(2)		(3)	
	Item nonresponse		Refuse		Don't Know		Unit Response ($t+1$)	
	$\hat{\beta}$	S.E.	$\hat{\beta}_1$	S.E.	$\hat{\beta}_2$	S.E.	$\hat{\beta}_3$	S.E.
u3			0.000	0.013	0.006	0.010	0.197	0.012
Interviewer level								
	v1		v1		v2		v3	
v1	2.429	0.581	3.163	0.776				
v2			0.217	0.827	5.845	1.562		
v3			−3.856	1.848	−1.944	2.730	31.938	6.975
Interviewer cluster	119				119			
Persons	637				637			
−2 * LogLikelih	−205.9				−6,172.3			

Note: ° Constrained to zero; Significance: *10%; **5%; ***1%.

Source: SOEP, Sample E, 2000, employed respondents without self-employed and trainees, without mail interviews (own calc.).

The additional variable "switch," explores the effect of mode change from PAPI to CAPI. We can see that the coefficient for the probability of refusal is significant positive in Wave 2 but significant negative in Wave 3. For a small portion of the respondents, the mode changed early, in Wave 2. This caused more income refusals than for respondents with unchanged mode. But the majority of the respondents changed from PAPI to CAPI in Wave 3. In this case, the coefficient indicates a decrease in probability of income nonresponse for refusals as well as for "don't knows." This inconsistent finding is hard to interpret. It could be a result of respondents' confidence. Respondents who changed their mode later might have more trust than those who made an earlier change. Or it could be that the interviewers were more confident with the new technology in Wave 3 than in Wave 2. The positive effect of mode switch in Wave 3 on participation in the next wave was already shown in the unit nonresponse analysis in Table 7.

Besides these definite CAPI effects, we find another mode effect: respondents who used a self-completion mode and filled out their questionnaires by themselves in front of the interviewer more often refused than in situations where the interviewers asked them orally. The self-completion mode partly reduces the interviewer's control over the interview situation and makes it easier for the respondent to skip embarrassing questions. Furthermore respondents in low earning positions have significantly more "don't knows" and respondents in high earning positions more refusals than respondents in medium positions. These findings are in line with previous results of analysis on other samples of the SOEP as well as of the BHPS (Schräpler 2004, 2006).

The interviewer variances in the random part of the model are more than three times their standard error and indicate interviewer or area influences on all three response categories. Nevertheless, we could find an identifiable influence of an interviewer gender effect only in a single case. In Wave 1, male interviewers lost more respondents than female interviewers. Hence we can conclude that the interviewer variance is caused mainly by unmeasured interviewer characteristics such as overall performance and skill.

5. Summary and Conclusion

This article assesses the effect of a change from the Paper-and-Pencil Interviewing (PAPI) method to Computer-Assisted Personal Interviewing (CAPI) in Sample E of the German Socio-Economic Panel (SOEP). Sample E contains over 1,000 German households and is split into two subsamples with the same structure, E1 and E2, using twin sample points. The 16 addresses in each sample point were surveyed in the first wave alternating between PAPI and CAPI. With this method split, we try to analyze CAPI effects and interviewer effects separately in Wave 1. After the second wave, the PAPI mode was systematically replaced by CAPI. Therefore we are able to measure pure mode effects at best only in Wave 1, and in the following waves we have to interpret our results as compounded effects of mode and other response-related characteristics such as respondent, interviewer, and area effects.

One important reason to change from PAPI to CAPI is to improve data quality. We have examined data collection mode effects using quality indicators like unit nonresponse, missing values, implausible values and gross income nonresponse.

The interviewers did not report problems with respondents' acceptance of CAPI during the fieldwork. Hence our first hypothesis is that we will not find a CAPI mode effect on respondents' decision not to participate in the following wave. We use random coefficient multilevel logit models to explore mode effects. The estimates in Wave 1 and Wave 4 show a direct negative effect of CAPI on refusing to participate in the following wave. Furthermore in the first two waves we find a significant interviewer variation of the CAPI coefficient. This finding suggests that the impact of the CAPI mode at time t on the participation in the following wave $t + 1$ depends on the interviewers' skill in managing the data collection method, especially in the first waves. A further result is that the coefficient for the change in data collection mode in Wave 3 indicates a significant decrease in the probability of nonparticipation in the next wave. This is important because in Wave 3, most of the PAPI respondents changed to CAPI. It seems that the switch in mode did not produce significant negative effects on the participation rate.

The second hypothesis is that CAPI reduces the number of implausible and missing values. Our descriptive analysis can support this assumption only partly: the rates of implausible values in Waves 2 to 5 are lowest in the CAPI mode. In Wave 1, the CAPI mode has the highest rate. It can be assumed that in the first wave the CAPI software system had to be adjusted to the special requirements of the SOEP. It may be that some transposition problems occurred in the first wave that were fixed later. In the case of the average number of missing values CAPI fails to meet the expectations. The results show that the CAPI technique does not show any improvements in comparison to PAPI.

The third hypothesis is that CAPI respondents do not have greater reservations about providing sensitive information such as gross income than respondents in the traditional PAPI mode. To explore this assumption we classify the missing values into two components: refusals and "don't knows." The estimates of the multivariate multilevel logit models show that in the first three waves CAPI interviews have a significantly higher probability of refusals (missing gross and net income) than PAPI interviews. One possible explanation is that the use of laptops increases privacy or confidentiality concerns. This interpretation of this finding is not in line with the assumption that respondents trust the confidentiality of computer-based data collection more than that of the traditional mode (de Leeuw et al. 1995). But of course, we can interpret this finding in another way as well, namely that the interviewers are not confident with CAPI and convey their fears of the technology to the respondents. But in any case, this result is important because the computer-assisted personal interviewing methods have increasingly replaced the traditional paper-and-pencil methods. In our study, we investigated only the gross income statement, but further research is needed to reinforce this finding. However, one general conclusion of our analysis is that this problem is crucial to address, and that work still remains to be done to decrease respondents' mistrust of the new data collection technology.

6. References

Baker, R.P. (1992). New Technology in Survey Research: Computer-assisted Personal Interviewing (CAPI). Social Science Computer Review, 10, 145–157.

Baker, R.P. and Bradburn, N.M. (1992). CAPI: Impacts on Data Quality and Survey Costs. Information Technology in Survey Research, Discussion Paper 10.

Baker, R.P., Bradburn, N.M., and Johnson, R.A. (1995). Computer-Assisted Personal Interviewing: An Experimental Evaluation of Data Quality and Cost. Journal of Official Statistics, 11, 415–434.

Bemelmans-Spork, M. and Sikkel, D. (1985). Data Collection with Hand-held Computers. Proceedings of the International Statistical Institute, 3.

Biemer, P.P. and Lyberg, L.E. (2003). Introduction to Survey Quality. Wiley: New Jersey.

Couper, M.P. (2005). Technology Trends in Survey Data Collection. Social Science Computer Review, 23, 486–501.

Couper, M.P. and Nicholls, W.L. II (1998). The History and Development of Computer Assisted Survey Information Collection. In Computer Assisted Survey Information Collection. New York: Wiley.

Danielsson, L. and Maarstad, P. (1982). Statistical Data Collection with Hand-Held Computers: A Consumer Price Index. Örebro, Sweden: Statistics Sweden.

De Leeuw, E.D. (2005). To Mix or Not to Mix Data Collection Modes in Surveys. Journal of Offical Statistics, 21, 233–255.

De Leeuw, E.D., Hox, J.J., and Snijkers, G. (1995). The Effect of Computer-Assisted Interviewing on Data Quality. International Journal of Market Research, 37, 325–344.

De Leeuw, E.D. and Nicholls, W.L. II (1996). Technological Innovations in Data Collection: Acceptance, Data Quality and Costs. Sociological Research Online, 1(4).

Dubin, J.A. and Rivers, D. (1989). Selection Bias in Linear Regression, Logit and Probit Models. Sociological Methods & Research, 18, 360–390.

Edwards, B., Bittner, D., Edwards, W.S., and Sperry, S. (1993). CAPI Effects on Interviewers: A Report from Two Major Surveys. Paper presented at the U.S. Bureau of the Census Annual Research Conference, Washington D.C.

Groves, R.M., Fowler, F.J., Couper, M.P., Lepkowski, J.M., Singer, E., and Tourangeau, R. (2004). Survey Methodology. New York: John Wiley & Sons.

Haisken-DeNew, J.P. and Frick, J.R. (eds) (2005). Desktop Companion to the German Socio-Economic Panel (SOEP). Berlin: DIW Berlin.

Hill, D.H. (1991). Interviewer, Respondent, and Regional Office Effects on Response Variance: A Statistical Decomposition. In Measurement Errors in Surveys, P. Biemer, R. Groves, L. Lyberg, N. Mathiowetz, and S. Sudman (eds). New York: Wiley, 463–483.

Hox, J.J. (1994). Hierarchical Regression Models for Interviewer and Respondents Effects. Sociological Methods & Research, 22, 300–318.

INFRATEST (1998). SOEP 98. Erstbefragung der Stichprobe – Methodenbericht. München. [In German]

INFRATEST (2000). SOEP 99. Methodenbericht zum Befragungsjahr 1999 (Welle 16) des Sozio-ökonomische Panels. München. [In German]

Kroh, M. and Spieß, M. (2006). Documentation of Sample Sizes and Panel Attrition in the German Socio-Economic Panel (1984–2005). DIW Data Documentation 15.

Laurie, H. (2003). From PAPI to CAPI: Consequences for Data Quality on the British Household Panel Study. Working Paper of the Institute for Social and Economic Research, Paper 2003–14, Colchester, University of Essex.

Lynn, P., Jäckle, A., Jenkins, S.P., and Sala, E. (2006). The Effect of Dependent Interviewing on Responses to Questions on Income Sources. Journal of Official Statistics, 22, 357–384.

Martin, J. and Manners, T. (1995). Computer Assisted Personal Interviewing in Survey Research. In Information Technology for the Social Scientist, R. Lee (ed.). London: UCL Press.

Martin, J., O'Muircheartaigh, C., and Curtice, J. (1993). The Use of CAPI for Attitude Surveys: An Experimental Comparison with Traditional Methods. Journal of Official Statistics, 9, 641–661.

Nicholls, W.L. II, Baker, R.P., and Martin, J. (1997). The Effect of New Data Collection Technologies on Survey Data Quality. In Survey Measurement and Process Quality, L. Lyberg, P. Biemer, M. Collins, E. de Leeuw, C. Dippo, N. Schwarz, and D. Trewin (eds). New York: Wiley.

Nicholls, W.L. II and de Leeuw, E.D. (1996). Factors in Acceptance of Computer-Assisted Interviewing Methods: A Conceptual and Historical Review. Proccedings of the Section of Survey Research Methods, American Statistical Association, 758–763.

Olsen, R. (1992). The Effects of Computer Assisted Interviewing on Data Quality. Paper presented at the fourth Social Science Methodology Conference, Trento, Italy.

Rasbash, J., Browne, W., Goldstein, H., Yang, M., Draper, D., Plewis, I., Healy, M., and Woodhouse, G. (1999). MLwiN User's Guide. London: Institute of Education.

Riede, T. and Dorn, V. (1991). Zur Einsetzbarkeit von Laptops in Haushaltsbefragungen in der Bundesrepublik Deutschland. Wiesbaden: Statistisches Bundesamt. Heft 20 der Schriftenreihe Ausgewählte Arbeitsunterlagen zur Bundesstatistik. [In German]

Rothschild, B. and Wilson, L. (1988). Nationwide Food Consumption Survey 1987: A Landmark Person Interview Survey Using Laptop Computers. Proceedings of the Fourth Annual Research Conference, 341–356.

Schräpler, J.-P. (2004). Respondent Behavior in Panel Studies – A Case Study for Income-Nonresponse by Means of the German Socio-Economic Panel (SOEP). Sociological Methods & Research, 33, 118–156.

Schräpler, J.-P. (2006). Explaining Income-Nonresponse – A Case Study by Means of the British Household Panel Study (BHPS). Quality & Quantity, 40, 1013–1036.

Schräpler, J.-P. and Wagner, G.G. (2001). Das "Interviewer-Panel" des Sozio-oekonomischen Panels – Darstellung und ausgewählte Analysen. Allgemeines Statistisches Archiv, 85(1). [In German]

Schupp, J. and Wagner, G.G. (2002). Maintenance of and Innovation in Long-term Panel Studies: The Case of the German Socio-Economic Panel Study (GSOEP). Allgemeines Statistisches Archiv, 86, 163–175.

Sebestik, J., Zelon, H., De Witt, D., O'Reilly, J., and McGowan, K. (1988). Initial Experiences with CAPI. Paper presented at the U.S. Bureau of the Census Annual Research Conference, Washington, DC.

Tourangeau, R., Rasinski, K., Jobe, J., Smith, T.W., and Pratt, W. (1997). Sources of Error in a Survey of Sexual Behavior. Journal of Official Statistics, 13, 341–365.

van Bastelaer, A., Kerssemakers, F., and Sikkel, D. (1988). Data Collection with Hand-held Computers: Contribution to Questionnaire Design. Journal of Official Statistics, 4, 141–154.

von Rosenbladt, B. and Stutz, F. (1998). SOEP '98 – Erstbefragung der Stichprobe E. Methodenbericht. SOEP-Gruppe in der Infratest Burke Sozial-forschung. München, November. [In German]

Voogt, R.J. and Saris, W.E. (2005). Mixed Mode Designs: Finding the Balance Between Nonresponse Bias and Mode Effects. Journal of Official Statistics, 21, 367–387.

Wagner, G.G., Frick, J.R., and Schupp, J. (2007). The German Socio-Economic Panel Study (SOEP) – Scope, Evolution and Enhancements. Schmollers Jahrbuch. Journal of Applied Social Science Studies, 127, 161–191.

Weeks, M. (1992). Computer-Assisted Survey Information Collection: A Review of CASIC Methods and their Implications for Survey Operations. Journal of Official Statistics, 8, 445–465.

Received August 2006
Revised June 2008