# SOEPpapers
## on Multidisciplinary Panel Data Research

# 288

Jörg-Peter Schräpler • Jürgen Schupp • Gert G. Wagner

# Individual and Neighborhood Determinants of Survey Nonresponse

**SOEPpapers on Multidisciplinary Panel Data Research**
at DIW Berlin

This series presents research findings based either directly on data from the German Socio-Economic Panel Study (SOEP) or using SOEP data as part of an internationally comparable data set (e.g. CNEF, ECHP, LIS, LWS, CHER/PACO). SOEP is a truly multidisciplinary household panel study covering a wide range of social and behavioral sciences: economics, sociology, psychology, survey methodology, econometrics and applied statistics, educational science, political science, public health, behavioral genetics, demography, geography, and sport science.

The decision to publish a submission in SOEPpapers is made by a board of editors chosen by the DIW Berlin to represent the wide range of disciplines covered by SOEP. There is no external referee process and papers are either accepted or rejected without revision. Papers appear in this series as works in progress and may also appear elsewhere. They often represent preliminary studies and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be requested from the author directly.

Any opinions expressed in this series are those of the author(s) and not those of DIW Berlin. Research disseminated by DIW Berlin may include views on public policy issues, but the institute itself takes no institutional policy positions.

The SOEPpapers are available at
**http://www.diw.de/soeppapers**

# Individual and Neighborhood Determinants of Survey Nonresponse –

## An Analysis Based on a New Subsample of the German Socio-Economic Panel (SOEP), Microgeographic Characteristics and Survey-Based Interviewer Characteristics[1]

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

[1] Ruhr University Bochum and IT NRW
[2] DIW Berlin (German Institute for Economic Research)
[3] Free University  Berlin
[4] Berlin University of Technology and Max Planck Institute for Human Development, Berlin

March 2010

---

## Abstract

This study examines the phenomenon of nonresponse in the first wave of a refresher sample (subsample H) of the German Socio-Economic Panel Study (SOEP). Our first step is to link additional (commercial) microgeographic data on the immediate neighborhoods of the households visited by interviewers. These additional data (paradata) provide valuable information on respondents and nonrespondents, including milieu or lifestyle, dominant household structure, desire for anonymity, frequency of moves, and other important microgeographic information. This linked information is then used to analyze nonresponse. In a second step, we also use demographic variables for the interviewer from an administrative data set about the interviewers, and, in a third step, we use the results of a special interviewer survey. We use multilevel statistical modeling to examine the influence of neighborhoods and interviewers on non-contacts, inability to participate, and refusals.

In our analysis, we find our additional variables useful for understanding and explaining non-contacts and refusals and the inability of some respondents to participate in surveys. These data provide an important basis for filling the information gap on response and nonresponse in panel surveys (and in cross-sectional surveys). However, the effect sizes of these effects are negligible. Ignoring these effects does not cause significant biases in statistical inferences drawn from the survey under consideration.

## Zusammenfassung

Am Beispiel der Auffrischungsstichprobe "H" des Sozio-oekonomischen Panels (SOEP) wird untersucht, welche Möglichkeiten zur Analyse von Unit-Non-Response in einer Zufallsstich-probe (nach dem Random Route-Verfahren gezogen) georeferenzierte Sozialstatistiken und die Befragung von Interviewern mit Hilfe eines „Interviewer Surveys" bieten (diese Art von Daten, in unserem Fall zusammen mit demographischen Merkmalen der Interviewer, die aus der Buchhaltung des Erhebungs-Instituts kommen, werden neuerdings „Paradaten" genannt) Es zeigt sich, dass alle Arten von Non-Response („kein Kontakt", „nicht in der Lage zu antworten" und „offene Verweigerung") mit georeferenzierten Sozialstrukturdaten und Persönlichkeits-merkmalen von Interviewerinnen und Interviewern zusammenhängen. Gleichzeitig sind die Effektstärken aber nicht nennenswert. D. h. die übliche Nicht-Berücksichtigung von georeferen-zierten Zusatzdaten und Interviewer-Merkmalen bei der Hochrechnung von Stichproben und der Analyse der Erhebungsdaten führt im Falle der SOEP-Stichprobe H faktisch zu keinen nennenswerten Verzerrungen. Wahrscheinlich ist das Ergebnis auch für jede Art von qualitativ hochwertigen Random-Route-Erhebungen mit Adressenvorlauf verallgemeinerbar.

# 1    Introduction

Unit nonresponse within surveys is one of the most important methodological issues in the empirical social sciences. There are many reasons for the differences that emerge between the addresses sampled and the survey actually carried out (see Schnell 1997). This paper examines unit nonresponse in the German Socio-Economic Panel Study (SOEP) in depth, using new types of "paradata" in an attempt to understand the process of nonresponse in the first wave of the most recent SOEP subsample. This analysis is of general interest for survey methodology because the first wave of a panel survey is essentially a new cross-sectional survey.

The SOEP is a longitudinal representative study containing socio-economic information on private households in the Federal Republic of Germany (see Wagner et al. 1993, Wagner et al. 2007) and is thus similar to the US Panel Study of Income Dynamics (PSID), for example. The first two SOEP samples from 1984 covered 12,290 West German respondents, and the study was gradually expanded to include a series of additional subsamples in the years that followed. Our nonresponse study deals with the most recent refreshment sample H which was drawn in spring 2006 (see Rosenbladt et al. 2006).

Two new features of the SOEP study are incorporated into our nonresponse study. First, it is possible to link information about sampling points with external geo-referenced data sets describing the socio-economic status of neighborhoods (see Hintze and Lakes 2009). Secondly, data about address contacts can be matched with almost unique microdata about the characteristics of the interviewers carrying out the fieldwork of SOEP. These data were collected by means of a special "interviewer survey" conducted at the end of 2006 (see Siegel and Stimmel 2007).  Thus, we are able to use new types of "paradata" for our analyses.[2]

---

[2] For the concept of "paradata" see, for example, Kreuter and Casas-Cordero (2010).

# 2    Explaining Unit Nonresponse

Several causes for unit nonresponse have been noted in the literature (Lynn 2008, Rendtel and Harms 2009). It is possible to distinguish between several problem areas and derive hypotheses about a contact's outcome. Unit nonresponse sometimes occurs because of a respondent's lack of ability or motivation to respond or lack of accessibility. However, the reasons for nonresponse also may be on the side of the fieldwork organization and in particular the interviewers and their behavior.

**Accessibility**

At its most basic level, unit nonresponse is not the result of unwillingness or inability but of the impossibility to contact household members. The "reachability" or accessibility of households constitutes a precondition for interviewing and the first step in an analysis of survey participation. Accessibility may be seen as a function of the physical possibility of reaching the household, the circadian rhythm of the household members and – of course – the contact strategy of the interviewer (Groves/Couper 1998). Sometimes it may be that the respondent is at home but does not open the door because he or she does not want to explicitly refuse to participate or because of fear of crime (Daeubler 2002).

Hence, it is possible to distinguish between individuals' levels of accessibility but the above-mentioned causes cannot be measured directly. In most analyses, demographical variables are used instead as a kind of proxy variable. Previous empirical results suggest that variables such as socio-economic status (Goyder 1987), household size, vocational status and the age of the contact person are important factors influencing "reachability" (see Schneekloth/Leven 2003, Koch 1997, Schraepler 2000). It is assumed, for example, that people in single households and middle-aged people are more mobile than others and therefore simply more difficult to reach. In this study, additional microgeographic "neighborhood data" about sample points are used to test these hypotheses.

One important aspect is the characteristics and motivation of the individual interviewer. Some interviewers are more successful than others because they have better contact strategies. We can assume that experienced interviewers and interviewers with a higher

workload are more effective than others.[3] In our study, we use additional data about interviewer characteristics to prove these hypotheses.

**(In)Ability**

At the second level, unit nonresponse depends on the ability of the household members to participate. Inability can occur, for example, when individuals are ill. We can assume that health problems increase with the age of the respondent, and hence, that it is mainly older persons beyond the age of 65 who will be unable to participate (see Schneekloth/Leven 2003, Koch 1997). We can also assume that the statement "unable to participate" is sometimes used as an excuse ("alibi").

We can interpret this behavior as a kind of "soft refusal" (not generally, but a certain percentage) connected to the ability of interviewers to convince people to participate. Therefore, we will examine whether some interviewers obtain more responses of "unable" than others, and whether certain interviewer characteristics can explain this outcome. In most previous empirical studies, "inability" has not been investigated separately from other nonresponse sources. In the present paper, therefore, we examine whether respondent and/or interviewer characteristics can explain the outcome of "unable to participate."

**Motivation**

In most empirical studies dealing with nonresponse, the cooperation of the respondent is explained by rational choice theory ("RC theory"), in particular cost/benefit analysis (Esser 1990). The motivation to participate depends on the respondent's assessment of the interview situation. In RC theory, it is assumed that after assessing the situation by comparing costs and benefits, the respondent evaluates the consequences of possible actions (probability of outcomes) and selects a specific action designed to achieve his or her personal goals. It is assumed that opportunity costs as well as privacy and

---

[3] A higher workload does not mean an easier job, but it may be an indicator per se of more effective contact strategies. This is the case when a higher workload is caused by selection (by the fieldwork organization) and self-selection (by the interviewer). In addition, a higher workload can help to minimize transaction costs of several attempts to contact a household.

confidentiality concerns are important in the decision to participate (see Schraepler 2006).

Opportunity costs: Participation in a survey, especially a longitudinal survey, takes time, meaning that this time is taken away from other activities. If respondents see the survey as not serving a meaningful purpose or not producing other benefits, they are likely to refuse participation. Some empirical studies suggest that due to higher opportunity costs middle-aged working people are more difficult to persuade to participate, and that refusal rates are higher in this group than in other age categories. In our study, microgeographic data are used to explore whether particular types of settlement structures have higher nonresponse rates than others.

Privacy and confidentiality concerns: Sometimes nonrespondents have a general aversion to answering questions that they think are an invasion of their privacy (Singer et al. 1993). In addition, individuals appear to vary in the degree to which they worry about confidentiality, which may affect their willingness to participate (see Sudman and Bradburn 1974; Singer et al. 1993, Pickery et al. 2001). Schneekloth and Leven (2003) hypothesize that in Germany these individuals tend to come from intellectual and more or less post-materialistic neighborhoods that encourage critical distance and possibly mistrust in surveys. In our study, microgeographic data are used to explore whether particular types of milieus have higher nonresponse rates than others.

Fear of crime: Fears of crime and danger, together with high levels of helplessness, may be assumed to cause nonresponse (see Schnell 1997). Empirical studies show that higher fears of danger occur mainly in areas of high population density, including large cities, downtown areas, and anonymous residential zones (see Koch 1997; Goyder 1987; DeMaio 1980). It is hypothesized that older individuals and women in particular have these kinds of insecurities and fears, especially in situations where they have to open the door to strangers or foreigners. Again, we use in our study, microgeographic data to explore whether particular types of neighborhood structures have higher nonresponse rates than others.

Interviewer: Many empirical studies show that the interviewer has a significant influence on respondents' decisions whether or not to participate in surveys (see Koch 1991). Besides interviewer age and gender, their motivations, clothing, and attitudes or other behavioral aspects can be significant for interview outcomes (see Schraepler 2006, 2004, Schraepler/Wagner 2001; Esser 1986, p. 41). In previous studies based on SOEP data, it has been shown that more experienced interviewers have lower refusal rates, and that interviewers with higher workloads are more successful than others (Schraepler 2000). In this study, an almost unique "interviewer survey" is used to test the effects of interviewer characteristics on response behavior of randomly selected households.

# 3    Unit Nonresponse in SOEP Sample H

The first wave of SOEP's sample H was launched in 2006 (Rosenbladt et al. 2006). The survey was conducted by 234 interviewers. Of these, 143 were already members of SOEP's interviewer staff and had experience with SOEP. The other 91 were new to the project and received some special  training by the fieldwork agency TNS Infratest Sozialforschung. All interviews were carried out by Computer Assisted Personal Interviewing (CAPI), and no paper version was available.

**Table 1: Reasons for Nonresponse in Sample H**

| Reasons for Nonresponse in Sample H at household level | N | % |
|---|---|---|
| **Gross Sample** | **3,931** | **100** |
| ./.  *Drop-Out Level I* | | |
| Household not detectable | 169 | 4.30 |
| Not feasible at the moment | 12 | 0.31 |
| **Adjusted Gross Sample** | **3,750** | **100** |
| ./.  *Systematic Drop-Out Level II* | | |
| Not accessible | 485 | 12.93 |
| Refusal | 1,487 | 39.65 |
| Not able to participate (e.g. nursing case) | 172 | 4.59 |
| Whole sample point lost | 15 | 0.40 |
| Individual household not surveyed | 82 | 2.19 |
| **Analyzable Interviews** | **1,509** | **40.24** |

Source: SOEP 2006, Gross Sample H, household level; authors' own calculations

The fieldwork was carried out during a time span of four to five months and separated into three main phases. To begin with, Germany was divided into 53,000 areas (sample points), and 250 sample points were chosen according to the size of the municipalities in each federal state. The second and third phases separated the process of recording addresses from the interviewing process. This separation is important for enforcing quality standards. Hence,

- the interviewer receives fixed addresses from the fieldwork organization,

- the quality of the addresses can be checked before the interviewing process begins,

- the target households can be sent a letter by mail giving notice of the survey, and

the addresses can be used to obtain additional information such as microgeographic data on the target household. A total of 3,931 household addresses were recorded by random walk (Thompson 2006). The first interviewer contact disclosed 181 drop-outs (uninhabited lodging, deceased persons), which can be considered as highly non-systematic drop-outs. Therefore, the "adjusted gross sample" contained 3,750 addresses. Of these target households, 1,487 (39.7%) were unwilling (refused) to participate and 485 (12.9%) households could not be reached. In a small percentage of cases (4.6%), respondents were "unable to participate" (a typical reason was long-term care). In 97 cases, the interviewers did not visit the households, and in one case a whole sample point was lost due to an interviewer's omission. For all reasons we cannot rule out a systematic effect and all nonresponse cases are included in the nonresponse figures. Overall, the response rate was 40.2 % and 1,509 interviews were conducted (see Rosenbladt et al. 2006 for details).

On average, three interviewer contacts per household were necessary to complete all interviews with one household. Fig. 1b shows that up to nine contact attempts were the basis for classifying a household as "unreachable" (non-contact) or as a refusal.

**Fig. 1: Nonresponse by size of township and number of attempts at contact**



a. Size of Township

b. Number of Attempts at Contact
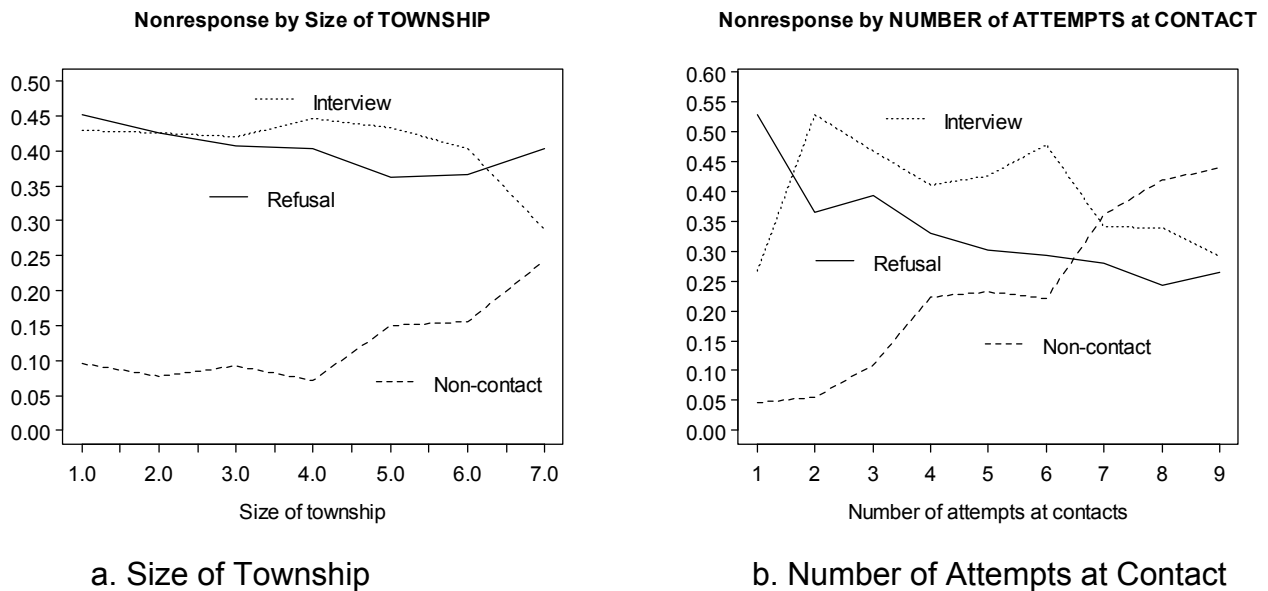
## 3.1   *Microgeographic Data and Systematic Drop-Outs*

To describe the neighborhoods of households contacted, we link additional (commercial) microgeographic data on the immediate vicinity of the households from the MOSAIC data system to the SOEP survey.[4] The MOSAIC data system contains over 75 indicators with neighborhood characteristics. These data are normally used to analyze and describe "customer databases" or "markets." This information is available at the address level and contains approximately 17.8 million buildings in Germany.

One building address covers seven or eight households on average (at least five households). Buildings with less than five households are pooled (for reasons of data protection) with households in the neighborhood that are similar in structure. This means that the linked information is not necessarily in line with the specific reality of a

---

[4] For a general description of possibilities to link geo-referenced "external" data to the microdata of SOEP, see Knies and Spiess (2007) and Hintze and Lakes (2009). For a detailed description of  geo-referenced data that is linkable to SOEP, see Goebel et al. (2009).

particular household in the gross sample but is an approximation for neighborhood characteristics.

The MICROM data set describes households and their environments in the following dimensions:

**MOSAIC "Sinus Milieus"**

The MOSAIC "Sinus milieus" combine the MOSAIC data with what is known as the Sinus milieu model. Sinus is a German company. "Milieus" encompass all the people and objects in the environment around us that are expected to influence "how we live and think." Sinus milieu groups (trademark: Sinus Sociovision) consist of people with similar attitudes towards life and lifestyles. Basic values, attitudes towards work, family, leisure, money and consumption as well as income, profession and education influence members of a particular milieu.[5]

The Sinus milieus turn the focus of attention to the individual and his or her entire living environment and social cosmos. The boundaries between milieus are fluid; there are points of contact and transitions between them. The potential for overlap and the position of the milieus in society, plotted according to social status and basic values, are visualized in a chart, often called a "potato chart." Overall, ten milieus may be distinguished in Germany.

---

[5] The assignment of households to the particular milieus was carried out using cluster analysis and probability models. The probability model was developed on the basis of a "calibration sample" by Sinus Sociovision (see http://www.sociovision.com for details).

**Table 2: A brief description of the Sinus milieus**

| Society's leading milieus | |
|---|---|
| *Well-established* | The self-confident establishment: success ethics, "can-do" mentality and highly exclusive tastes |
| *Post-materialists* | The enlightened post-68 generation: post-materialist values, critique of globalization, and intellectual interests |
| *Modern performers* | The young and unconventional high performers: intensive lifestyle - both at work and at play, multiplicity of options, flexibility and multimedia enthusiasm |
| **Traditional milieus** | |
| *Upper conservatives* | The old German educated classes; conservative critique of modern culture, humanist sense of duty, and cultivated manners |
| *Traditionals* | The security and order-loving wartime generation: rooted in the petty bourgeois world or traditional blue-collar culture |
| *Nostalgics of the former GDR* | The resigned losers of German reunification: clinging to Prussian virtues and old socialist notions of justice and solidarity |
| **Mainstream milieus** | |
| *New middle class* | The status-oriented modern mainstream: attempting to establish themselves professionally and socially, seeking a sheltered and harmonious life |
| *Materialists* | The markedly materialistic lower class: attempting to keep up with the consumer standards of the broad middle classes to compensate for social disadvantages |
| **Hedonistic milieus** | |
| *Experimentalists* | The extremely individualist Bohemian world: unchecked spontaneity, high value placed on their avant-garde lifestyle and self-conception |
| *Escapists / hedonists* | The fun-oriented, modern lower middle class: disregard for conventions and behavioral expectations of achievement-oriented mainstream society |

Source: Sinus Sociovision

Table 2 and Fig. 2 show that in Germany the "well-established," the "post-materialists" and the "modern top performers" make up the milieu of society's leaders. The "upper conservatives," the "traditionalists" and the "nostalgics of the former GDR" represent the various facets of the traditional milieu. The mainstream milieu consists of the "middle class" and "materialists," while the "experimentalists" and "escapists/hedonists" (pleasure seekers) comprise the hedonistic milieu (see Sinus Sociovision 2007).

**Fig. 2: Sinus milieus in Germany 2007**
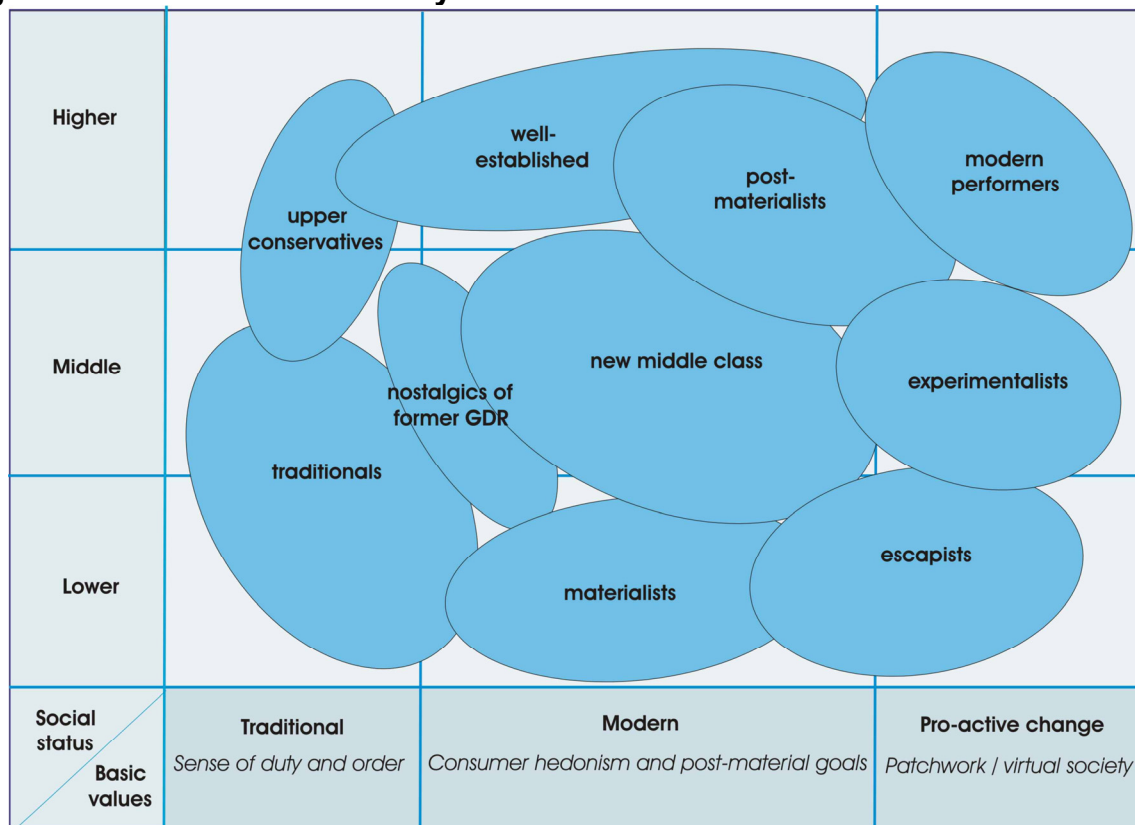


Table 3 presents the distributions of the Sinus milieus in various neighborhoods. The first column shows the distribution for Germany reported by Sinus Sociovision for the year 2007 and the second column the distribution for all SOEP addresses in sample H in 2006. Small differences can be seen across the different columns in the categories "nostalgics of former DDR," "materialists," "hedonists" and "experimentalists." The first category only exists in eastern Germany and is overrepresented, while the shares of the last three milieus are slightly lower for SOEP respondents than for Sinus Sociovision respondents. It must be noted that no-one knows the sample characteristics of the Sinus Sociovision respondents – these respondents and the distribution of "Sinus milieu" certainly do not represent a "benchmark" whose validity cannot be challenged.

Columns three to six show the distribution at a microgeographic level for the responding households, refusals, unreachable households and those unable to participate. On the basis of the distribution of the SOEP's household addresses, we can identify a strong overrepresentation of "well-established" and an underrepresentation of "modern performers," "experimentalists" and "hedonists" amongst the households where an

interview was conducted. The lower percentages might be due to the higher share of unreachable households. Furthermore, we can see that the category "new middle class" has the highest share of refusals and that "unable to participate" was mainly an issue for households living in "traditionalist" mileus.

**Table 3: Distribution of Sinus milieus at a microgeographic level by outcome of the fieldwork for Sample H**

| Sinus milieus | Sinus Sociovision 2007 | SOEP Sample H | | | | |
|---|---|---|---|---|---|---|
| | | Addresses 2006 | Respondents 2006 | Refusals 2006 | Non-contacts 2006 | Unable 2006 |
| Well-established | 10 | 11.6 | **14.4** | 10.5 | 8.5 | 6.6 |
| Post-materialists | 10 | 9.9 | 10.7 | 9.4 | 8.3 | 6.0 |
| Modern performers | 10 | 9.1 | 7.8 | 8.4 | 11.1 | **12.6** |
| Upper conservatives | 5 | 5.7 | 5.9 | 6.1 | 4.3 | **9.6** |
| Traditionalists | 14 | 15.3 | 15.9 | 15.8 | 10.5 | **22.8** |
| Nostalgics of former DDR | 5 | 7.2 | 7.5 | 6.2 | 8.9 | 5.4 |
| New middle class | 15 | 15.4 | 15.3 | **18.0** | 11.1 | 13.8 |
| Materialists | 12 | 10.5 | 10.3 | 11.7 | 11.9 | 7.2 |
| Hedonists/escapists | 11 | 8.8 | 8.1 | 7.8 | 10.6 | **12.0** |
| Experimentalists | 8 | 6.3 | **4.1** | 6.0 | **14.7** | 4.2 |
| | 100 | 100 (3,782) | 100 (1,449) | 100 (1,435) | 100 (470) | 100 (167) |

Source: SOEP 2006, Gross Sample H and MICROM data set; authors' own calculations

**Social status and purchasing power**

Social status classifies household social stratification based on education and income in the neighborhood into nine categories. A value of one identifies the households with the lowest social status, while a value of nine identifies households with the highest social status. Figure 3a shows the share of refusals, interviews conducted and unreachable households (non-contacts) according to the categories of social status. The distributions for the interview outcome do not suggest any significant variations with status.

The category "purchasing power" (or "spending capacity") is defined on the basis of disposable income. It takes into account all sources of income for each member of a household. MICROM developed an index for smaller geographical areas than those included in the official statistical numbers. The average purchasing power per household in 100 euros is used for our analysis. Figure 3b indicates a possible positive effect on the refusal rate only in the case of purchasing power > €50,000 per household. We will test this effect with a multivariate model in the next section.

**Fig. 3: Social status and spending capacity by interview outcome**

Nonresponse by MOSAIC STATUS

Nonresponse by SPENDING CAPACITY



a. Social status

b. Spending capacity per household
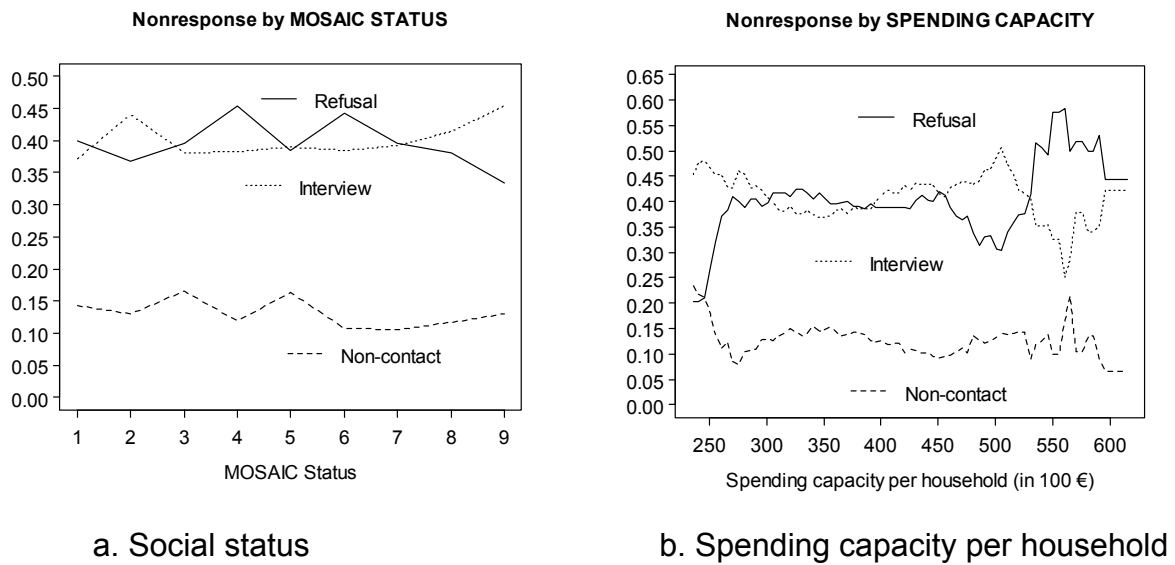
**Desire for Anonymity and Percentage of Foreigners**

The variable "desire for anonymity" is based on published information on individuals about profession, surname and address to create an index of one to nine on the level of neighborhoods and addresses. An index value of one indicates the lowest desire for anonymity, while a value of nine indicates the highest desire for anonymity. The validity of this index is almost unknown, but in our case it contains valid information. Figure 4a indicates an increase in "non-contacts" with an increased desire for anonymity.

The percentage of foreigners is based on analyses of first names to identify the presence of foreigners as heads of household in a certain neighborhood. This variable is also banded into nine groups. Figure 4b suggests there is also an increase in "non-contacts" with the share of foreigners in the neighborhood.

**Fig. 4: Desire for anonymity and share of foreigners by interview outcome**

a. Desire for anonymity



b. Share of foreigners
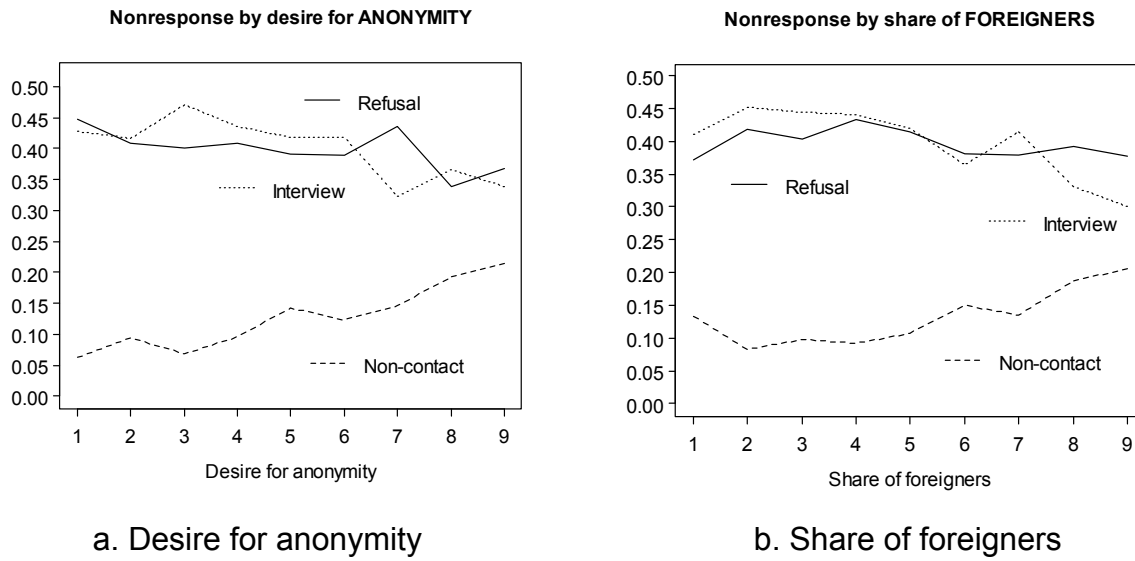
**Removal Intensity**

The variable "removal intensity" means the frequency of moves and expresses the total number of movers within 1,000 households. Again, the variable is banded into nine group values. Fig. 5 shows a strong increase in the percentage of unreachable households with the frequency of moves in a particular neighborhood. The percentage of refusals and respondents are about the same and decrease only slightly with frequency of moves.

**Figure 5: Frequency of moves by interview outcome**

Nonresponse by FREQUENCY of MOVES
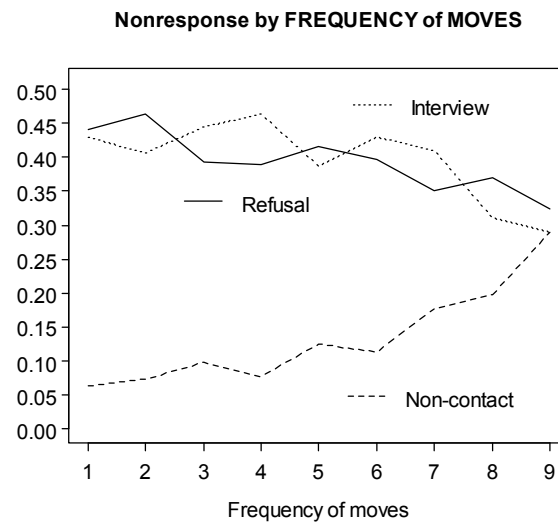


## 3.2 Interviewer Variables from the Interviewer Data Set and the Interviewer Questionnaire

To explore interviewer effects in SOEP, two sources of data on interviewers are utilized:

- a special interviewer data set which is based on bookkeeping data of the fieldwork organization. This data set is available for all waves of SOEP. The data set contains a variety of interviewer characteristics such as age and gender as well as education and some other characteristics (see Schraepler/Wagner 2001 for details).

- a data set based on an almost unique "interviewer survey questionnaire" filled in by the majority of SOEP interviewers containing not only demographic variables but also many personality variables and self-assessments. For subsample H, 187 out of 234 interviewers responded to this special questionnaire (Siegel and Stimmel 2007; Weinhardt et al. 2010).[6]

---

[6] There is some evidence that the nonresponse rate in the interviewer survey is related to the success rate of interviewers for sample H. The success rate of the interviewers who participate in the "interviewer survey" is on average, at 43 %, significantly higher then the success rate of the nonrespondents (35 %).

Fig. 6a and Fig. 6b examine the interview outcomes in relation to interviewer age and gender by local regression. The distributions show that refusals are an issue especially with younger male interviewers, who show a clearly higher share of refusals than older men. Male interviewers above the age of 45 show almost the same refusal and no-contact rates as female interviewers.

**Fig, 6a: Interview outcome by age of interviewer**



Nonresponse by age of interviewer
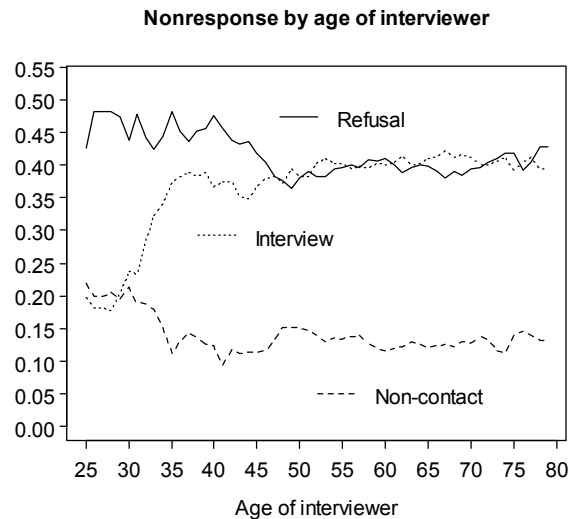
**Fig. 6b: Interview outcome by age and gender of interviewer**



Nonresponse by age of male interviewer

Nonresponse by age of female interviewer
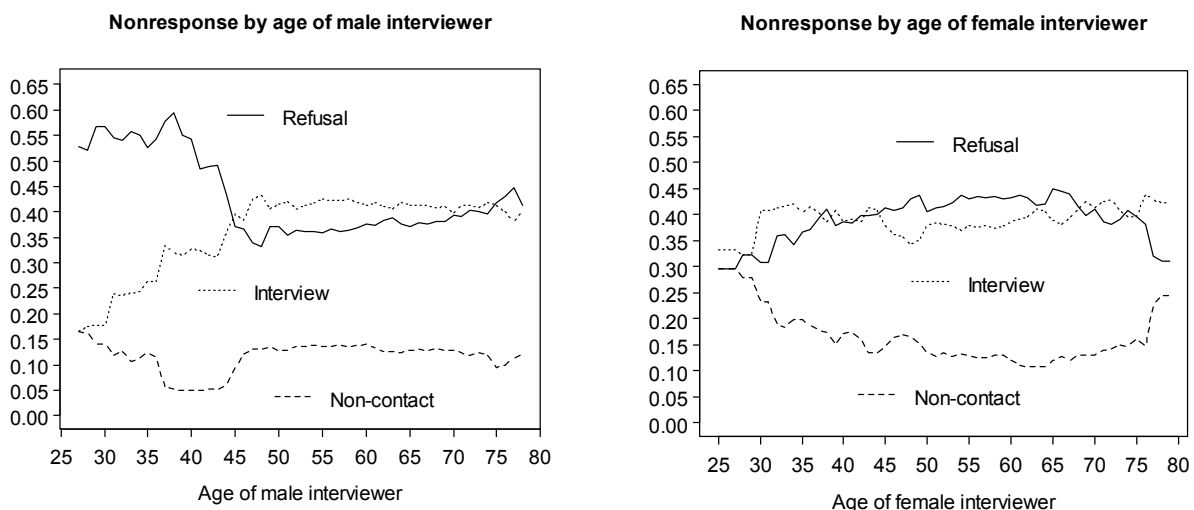
The interviewer questionnaire contains a number of self-assessments as well as an item block that can be used to indicate the interviewer's tendency for social desirability (SD). The tendency for SD is indicated if the interviewer consistently responds in a socially

desirable manner. For the construction of SD seven questions about interviewers' own behavior are used (socially desirable answers in parentheses):

- My first impression of people normally turns out to be right.          (yes)
- I often doubt my own judgment.          (no)
- I always know exactly why I participate in something.          (yes)
- I have received too much change in a shop before without saying anything.          (no)
- I'm always honest with other people.          (yes)
- I occasionally take advantage of others.          (no)

If interviewers consistently respond in a socially desirable manner, the variable SD = 1 and otherwise 0. Concerning sample H on the interviewer level, we found that 67 of the 165 interviewers (40.6 %) who filled out the questionnaire have the indication that SD = 1.[7]

To ascertain the significance of respondent and interviewer attributes, we estimate multilevel regression models in the following section.

## 3.3   Modeling Unit Nonresponse

Unit nonresponse (non-participation) is given when respondents are unable (ill, deceased, or have moved abroad) or unwilling (refusing) to participate in the survey or if they are not reachable. In most of the first two cases, the respondent has chosen between two alternatives: participation or non-participation. In the latter, the interviewer is unable to contact the household members.

The easiest way to determine how various factors influence the attractiveness of the alternatives to different types of individuals or affect the probability of a non-contact is to use a regression framework. Because of the binary response, we use an ordinary logit model with multilevel extension. We estimate four regression models:

---

[7] The SD measurement is problematic in the case of interviewers and has to be interpreted with caution. The relationship between the interviewer and the fieldwork organization conducting the interviewer survey may bias interviewers' response behavior towards higher SD shares.

- Model 1 shows the influence of various factors on the probability of response variable "interview" (participation) vs. nonresponse.
- Model 2 determines the influence of factors on the probability of response variable "refuse to participate" vs. "participate"
- Model 3 reveals the influence of regressors on the probability of response variable "household not reachable" vs. "participate"
- Model 4 presents the influence of regressors on the probability of response variable "household not able to participate" vs. "participate".

Survey data are hierarchical in structure due to multiple nesting; in our case, the respondents (head of the household) are nested in interviewer clusters (see 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:

$$y_{ij} = \begin{cases} 1, \text{if } y_{ij}^* > 0, \text{participation} \\ 0, \qquad \text{otherwise} \end{cases}$$

or

$$y_{ij} = \begin{cases} 1, \text{if } y_{ij}^* > 0, \text{unit-nonresponse (refuse, non-contact, unable)} \\ 0, \qquad\qquad\qquad\qquad \text{otherwise} \end{cases}$$

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

If we specify a two-level random intercept model (model 1), the probability of each response is estimated with the equation:

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

where $x_{h,ij}$ represents values for covariates $x_h$ ($h$ = 1, …,$H$) of respondent $i$ and interviewer $j$. The intercept $\beta_{0j}$ 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 $u_{ij}$.

We estimate three versions of the model with different sets of regressors: in version A, we use only respondent variables estimated on the basis of the household addresses and in version B, we expand our model with microgeographic variables describing the neighborhood and demographic variables from the SOEP interviewer data set.

In the last step, in version C, we explore the effect of the interviewer's opinions about respondents and herself or himself on the interview outcome. Most of the interviewers in sample H filled in the interviewer questionnaire, so we can broaden our model to also include these additional variables.

Regressors can be divided into four groups:

- Demographic and household variables for the **_potential_** respondent (here head of the household at a microgeographic level):
  - "*age*" is the approximate age of the *potential* respondents. The values are based on the basis of first names (eight categories, quantitative)
  - "*status*" indicates the socio-economic status of *potential* respondent's household. Based on estimated income and education (nine categories, quantitative)
  - "*sinus-milieu*" shows the dominant Sinus milieu® in the respondent's area. The Sinus milieu is a typology of Sociovision based on values and lifestyle segmentation (ten categories, qualitative)
  - "*purchasing power per household*" = 1 indicates *potential* high purchasing power (>€53,000 per household), calculated on the basis of data from official statistics, data from GFK Marktforschung as well as MICROM data (dummy variable)
  - "*eastern Germany*" = 1 indicates if household is located in the eastern part of Germany (dummy variable)

- Microgeographic variables for the neighborhood:
  - "*city*" = 1 indicates that the respondent's household is located in a city with over 500,000 residents (dummy variable)
  - "*family*" indicates the dominant household structure in the neighborhood, for example, predominantly single households or higher-than-average share of households with children (nine categories, qualitative)
  - "*hhnumber*" indicates the *potential* number of households in the respondent's house (numeric)
  - "*MOSAIC type*" household classification based on cluster analysis and the following dimensions: "city – rural," "old – young residents," "old – new buildings" and "income of the residents" (38 categories overall, qualitative)
  - "*size of buildings*" (seven categories overall, qualitative)
  - "*anonymous*" desire for anonymity in the area (nine categories, acceding)
  - "*foreigner*" share of foreigners in the area (nine categories, acceding)
  - "*garden*" affinity for garden (nine categories, acceding)
  - "*move*" frequency of moves in respondent's area (nine categories, acceding)

- Demographic variables for the interviewer:
  - "*isex*" = 1 indicates male interviewer (dummy variable)
  - "*iage*" is the age of the interviewer in years (numeric)

- o "*ischool*" highest school degree of the interviewer (three categories, qualitative)
- o "*Workload*" is the number of household addresses of each interviewer.
- o "SOEP experience" = 1 indicates that the interviewer has experience with SOEP interviews

- Interviewer variables from the interviewer questionnaire
  - o "SD" = 1 indicates that interviewer has a tendency for social desirability
  - o "satisfaction with life" (11 categories, quantitative, acceding)
  - o item block "What kind of character do you have?" (16 items, each have seven categories, quantitative)
  - o "risk propensity" (ten categories, quantitative)
  - o "I'm patient" (ten categories, quantitative)
  - o "years of work for SOEP in future" (quantitative)

Tables 4, 5 and 6 show the estimates of the multilevel logit models 1 to 4 in versions A - C for the four response variables. In version A, we use only respondent variables at a microgeographic level as predictors. Model 1 refers to "participation," model 2 to "refuse to participate," model 3 to "non-contact" and model 4 to "inability to participate." The models contain 3,520 to 1,615 observations (address units) from 234 to 222 interviewers.

The estimates of the parameters for the response variables confirm our hypotheses that middle-aged and older people have a lower probability of participating in surveys than younger people. Models 2 and 3 show that the probability of refusing as well as of non-contact is significantly higher in the age category ">45–50 years" for the *potential* (estimated) age of the head of household and for non-contact also in the category ">55–60 years" than in the reference group "<35 years." Our hypothesis that it is more likely that older people will be unable to participate is not supported by our results: potential age is not significant for this response variable.[8] We have to interpret these and the following findings with caution because the microgeographic data used are not accurate respondent characteristics but approximations for neighborhood characteristics ("*potential characteristics*"). Our hypotheses are formulated at an individual level whereas the coefficients of the regressors of our models are estimated at a microgeographic level.

---

[8] This result may depend on which specific categories are used for the upper age range. Unfortunately, the classification of the variable "age" in the MICROM data set has only "65 years and more" as the highest category.

Nonrespondents who refused to participate differ markedly in the "Sinus milieus" from those who were unable to participate or could not be reached. We use the "well-established" milieu as a reference group (see Table 3). The estimates suggest that the probability of refusal vs. participation is higher especially in the mainstream milieus such as the middle class ($\beta$ = 0.405) and the materialists ($\beta$ = 0.331) and also for hedonistic milieus such as the experimentalists ($\beta$ = 0.398). The probability of unreachability vs. participation is highest for the experimentalists ($\beta$ = 1.185) and modern performers ($\beta$ = 0.712). The former are described with characteristics such as unchecked spontaneity, and the latter with intensive living and high flexibility and enthusiasm. Both are fairly young groups with unconventional lifestyles. It seems that households in these milieus are more difficult to reach.

The estimates for households that are unable to participate are difficult to interpret because we obtain significant results for different milieus. The results show significant coefficients for households who relate to upper conservatives and traditionalists as well as for households in modern performer and escapist/hedonist milieus. All milieus have positive coefficients, which means that the reference group of "well-established" has the lowest probability of inability vs. participation. The highest positive coefficients are for households in the traditional milieus: the "upper conservatives" ($\beta$ = 1.659) and "traditionals" ($\beta$ = 1.438).

Other variables such as "purchasing power per household >€53.000" and "social status" are not significant in our models. The dummy variable "eastern Germany" indicates that the response rates are higher in eastern Germany than in western Germany, and model 3 shows that this is caused mainly by a lower share of non-contacts ($\beta$ = - 0.538).

We find significant random effects in all models, indicating that relevant interviewer effects which are not measured directly do exist. The intraclass correlation coefficient (ICC) measures the proportion of variance in outcomes between the level 2 units (interviewers). The estimates show especially high intraclass correlations for the response variables "non-contact" (model 3, ICC = 0.399) and "unable to participate" (model 4, ICC = 0.369).

**Table 4: Multilevel logit models, version A – only microgeographic variables that relate to the head of the household**

| Variable | Model 1 Participation vs. non-participation | | | Model 2 Refusal vs. participation | | | Model 3 Non-contact vs. participation | | | Model 4 Unable vs. participation | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Coeff. | z value | | Coeff. | z value | | Coeff. | z value | | Coeff. | z value | |
| **Fixed effect** | | | | | | | | | | | | |
| (Intercept) | -0.172 | -0.68 | | -0.242 | -0.90 | | -1.762 | -3.74 | *** | -3.721 | -5.47 | *** |
| **Head of the household** | | | | | | | | | | | | |
| Age < = 35 years (ref.) | | | | | | | | | | | | |
| Age > 35 - 40 years | -0.111 | -0.57 | | 0.090 | 0.43 | | 0.439 | 1.21 | | -0.177 | -0.34 | |
| Age > 40 - 45 years | -0.289 | -1.56 | | 0.199 | 1.00 | | 0.420 | 1.21 | | 0.529 | 1.12 | |
| Age > 45 - 50 years | -0.483 | -2.65 | ** | 0.491 | 2.51 | + | 0.640 | 1.87 | + | 0.241 | 0.51 | |
| Age > 50 - 55 years | -0.069 | -0.37 | | 0.028 | 0.14 | | 0.373 | 1.04 | | 0.275 | 0.57 | |
| Age > 55 - 60 years | -0.169 | -0.88 | | 0.072 | 0.35 | | 0.811 | 2.24 | * | -0.263 | -0.50 | |
| Age > 60 - 65 years | -0.327 | -1.68 | + | 0.270 | 1.30 | | 0.579 | 1.55 | | 0.217 | 0.44 | |
| Age > 65 y. | -0.210 | -1.05 | | 0.153 | 0.71 | | 0.458 | 1.15 | | 0.232 | 0.45 | |
| **Sinus milieus** | | | | | | | | | | | | |
| Well-established (ref.) | | | | | | | | | | | | |
| Post-materialists | -0.100 | -0.61 | | 0.177 | 0.99 | | -0.101 | -0.34 | | 0.033 | 0.06 | |
| Modern performers | -0.438 | -2.45 | * | 0.305 | 1.58 | | 0.712 | 2.32 | * | 1.135 | 2.25 | * |
| Upper conservatives | -0.333 | -1.69 | + | 0.328 | 1.55 | | -0.389 | -1.03 | | 1.659 | 3.23 | ** |
| Traditionalists | -0.185 | -1.14 | | 0.213 | 1.23 | | -0.475 | -1.54 | | 1.438 | 3.02 | ** |
| Nostalgics of the former DDR | -0.107 | -0.52 | | 0.064 | 0.28 | | -0.091 | -0.25 | | 0.429 | 0.69 | |
| New middle class | -0.308 | -1.88 | + | 0.405 | 2.33 | * | -0.112 | -0.36 | | 0.883 | 1.79 | + |
| Materialists | -0.280 | -1.56 | | 0.331 | 1.72 | + | 0.173 | 0.54 | | 0.392 | 0.69 | |
| Hedonists/escapists | -0.219 | -1.15 | | 0.116 | 0.56 | | 0.306 | 0.93 | | 1.119 | 2.11 | * |
| Experimentalists | -0.629 | -2.86 | ** | 0.398 | 1.67 | + | 1.185 | 3.33 | *** | 0.941 | 1.43 | |
| Purchasing power per household > 530 € | -0.067 | -0.22 | | 0.247 | 0.78 | | -0.479 | -0.85 | | 0.140 | 0.16 | |
| Status | 0.021 | 0.86 | | -0.012 | -0.47 | | -0.046 | -1.11 | | 0.009 | 0.16 | |
| Eastern Germany | 0.273 | 1.63 | + | -0.260 | -1.49 | | -0.534 | -1.73 | + | -0.043 | -0.12 | |
| **Random effects** | 95% Interval | | | 95% Interval | | | 95% Interval | | | 95% Interval | | |
| $u_{ij}$. | $\pi^2 / 3$ | | | $\pi^2 / 3$ | | | $\pi^2 / 3$ | | | $\pi^2 / 3$ | | |
| $v_{0j}$. (intercept) | 0.673 | (0.53 - 1.02) | | 0.640 | (0.49 - 0.98) | | 2.180 | (1.69 - 3.18) | | 1.92 | (1.09 - 3.13) | |
| ICC | 0.170 | | | 0.163 | | | 0.399 | | | 0.369 | | |
| Interviewer | 234 | | | 231 | | | 226 | | | 222 | | |
| Household addresses | 3,520 | | | 2,883 | | | 1,918 | | | 1,615 | | |
| Log likelihood | -2,247 | | | -1,897 | | | -914 | | | -493 | | |
| Pseudo R² | 0.04 | | | 0.04 | | | 0.06 | | | 0.06 | | |

Note: Significance: + 10%; * 5 %, ** 1 %, *** 0.1 %

Source: SOEP 2006, Sample H and MICROM data set, authors' own calculations

In version B, we extend the models to include additional regressors such as microgeographic variables for the neighborhood from the MICROM data set and interviewer variables from the interviewer data set. The fit of each of the models is measured by the pseudo R² and increases substantially with the added variables. We find the same pattern as in the version A models for respondent variables such as age and Sinus milieu. However, the estimates of version B models also show several

significant interviewer variables. Consequently, the interviewer variance or the intraclass correlation declines in the random part.

One main effect of interviewer gender may be recognized for response variable "unable to participate." Male interviewers have a significantly higher probability of showing this outcome ($\beta$ = 0.599). Therefore, the interaction term for interviewer's age and gender (interviewer age < 40 & male) drives the significant increase in probability of refusals ($\beta$ = -0.508). Furthermore, the estimated coefficients indicate another main effect, which is the interviewer's education. The probability of participation in the survey declines for more highly educated interviewers ($\beta$ = -0.455). The interviewer's workload seems to be relevant for all response variables. While the probability of an interview increases with the interviewer's workload, the probability of refusals, non-contact and "inability to participate" decreases significantly. This finding indicates that interviewers with a high number of addresses work more efficiently than interviewers with lower workloads.[9] The interviewer's past SOEP experience is not significant in this model.

To describe the surroundings in the respondent's neighborhood, we have used a dummy variable "city" and the variable "MOSAIC type," a household classification based on cluster analysis and the following dimensions: "city – rural," "old – young residents," "old – new buildings," and "income of the residents" (38 categories overall). Our reference category is "simple urban row estate," which has the highest frequency of addresses (n=181) in the sample. If we sort the estimated logit coefficients of all categories in descending order, we find the highest probability of participation for "high-income families, new privately owned buildings" ($\beta$ = 1.030), "old families in the outskirts" ($\beta$ = 1.028), "self-employed in new buildings" ($\beta$ = 0.994), "new high-quality detached houses – commuter belt" ($\beta$ = 0.957). In contrast, the lowest probability for participation can be found among "old social housing" ($\beta$ = -0.301, not significant). This finding corresponds with the estimates for the variable "Sinus milieus" of the respondents. High earners and self-employed people and households in high-quality new houses often belong to the "well-established" milieu. It is precisely this milieu that has the highest probability of participation in the survey.

---

[9] This is good news for old SOEP policies of TNS Infratest: rather than defining an artificial upper limit to an individual interviewer's workload (like in some other studies), it shows that the SOEP policy to allow any increases in workload as long as it manageable for interviewers and based on good quality work the individual interviewer also seems to work well for a refreshment sample.

The opposite picture appears for the response variable "refusal". The category "old social housing" has the highest probability of refusals but is not significant ($\beta$ = 0.462). Almost all significant categories have negative coefficients, indicating that the reference category "simple urban row housing" is one of the worst areas or has a higher probability of refusals. The lowest probability of refusals can be found for "old families in the outskirts" ($\beta$ = -1.002), "self-employed in new buildings" and "social climbers, upscale professions, outskirts." As in many other empirical studies, we have found an increasing probability of refusals ($\beta$ = 0.407) and non-contacts ($\beta$ = 0.834) in urban areas such as big cities.

Significantly higher reachability than for the reference category can be found in the case of younger villagers ($\beta$ = -1.856), simple houses in rural districts ($\beta$ = -1.496), self-employed in new buildings ($\beta$ = -1.382), simple vocations in rural areas ($\beta$ = -1.118), villages in the outskirts ($\beta$ = -1.087), and also in old city centers ($\beta$ = -1.398) and astonishingly, in social hotspots ($\beta$ = -1.117).

The coefficients for "unable to participate" indicate lower probabilities for "high-income families and new privately owned homes" ($\beta$ = -3.404), "distinguished detached houses" ($\beta$ = -2.289), "simple vocations in rural areas" ($\beta$ = -2.364), "younger villagers" ($\beta$ = -2.185), "new terraced houses in rural areas" ($\beta$ = -2.138) and "old city centers" ($\beta$ = -2.139).

Besides these area effects, we found virtually no significant effects for different sizes of respondents' homes. We only find significantly lower probabilities of "unable to participate" in accommodations larger than one to two-family homes.

Clear results are found for the variable "dominant family structure in the area." The category "mainly single households" is used as a reference, and the estimates show that the probability of refusal increases and the probability of non-contact decreases significantly with children in households.

The coefficient for the variable "frequency of moves in the area" (nine categories, acceding) indicates an increasing probability of non-contact and "inability to participate" with the frequency of moves.

The estimates for other microgeographic variables from the MICROM data set such as "desire for anonymity in the area" (nine categories, acceding), "share of foreigners in the area" (nine categories, acceding), "affinity for garden" (nine categories, acceding) are more difficult to interpret because the results turn out to be somewhat ambiguous. The increase in non-contacts with an increasing desire for anonymity shown in Fig. 4 is not significant in model 3, version B, and the increasing share of foreigners in the area indicates a lower probability of non-contact. The estimates for "affinity for garden" show inconsistent findings.

**Table 5: Multilevel logit models, version B: microgeographic variables that relate to the head of the household, for the interviewer and microgeographic variables for the neighborhood**

| Variable | Model 1 Participation vs. non-participation | | | Model 2 Refusal vs. participation | | | Model 3 Non-contact vs. participation | | | Model 4 Unable vs. participation | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Coeff. | z value | | Coeff. | z value | | Coeff. | z value | | Coeff. | z value | |
| ***Fixed effect*** | | | | | | | | | | | | |
| (Intercept) | -1.482 | -2.13 | * | 0.360 | 0.47 | | 0.279 | 0.22 | | -1.106 | -0.58 | |
| ***Head of the household*** | | | | | | | | | | | | |
| Age < = 35 years (ref.) | | | | | | | | | | | | |
| Age > 35 - 40 years | -0.160 | -0.78 | | 0.221 | 0.99 | | 0.297 | 0.77 | | -0.238 | -0.38 | |
| Age > 40 - 45 years | -0.292 | -1.52 | | 0.239 | 1.14 | | 0.309 | 0.84 | | 0.550 | 0.97 | |
| Age > 45 - 50 years | -0.454 | -2.39 | * | 0.519 | 2.52 | * | 0.437 | 1.18 | | 0.181 | 0.32 | |
| Age > 50 - 55 years | -0.019 | -0.10 | | 0.050 | 0.24 | | 0.167 | 0.43 | | 0.079 | 0.14 | |
| Age > 55 - 60 years | -0.128 | -0.64 | | 0.102 | 0.47 | | 0.561 | 1.46 | | -0.275 | -0.45 | |
| Age > 60 - 65 years | -0.294 | -1.45 | | 0.313 | 1.43 | | 0.407 | 1.02 | | 0.232 | 0.39 | |
| Age > 65 years | -0.212 | -1.00 | | 0.251 | 1.10 | | 0.306 | 0.71 | | 0.315 | 0.52 | |
| *Sinus milieus* | | | | | | | | | | | | |
| Well-established (ref.) | | | | | | | | | | | | |
| Post-materialists | -0.194 | -1.12 | | 0.282 | 1.51 | | 0.023 | 0.07 | | 0.190 | 0.30 | |
| Modern performers | -0.419 | -2.06 | * | 0.499 | 2.25 | * | 0.247 | 0.68 | | 0.764 | 1.20 | |
| Upper conservatives | -0.249 | -1.20 | | 0.200 | 0.89 | | -0.547 | -1.30 | | 1.986 | 3.27 | *** |
| Traditionalists | -0.108 | -0.62 | | 0.146 | 0.78 | | -0.634 | -1.83 | + | 1.483 | 2.60 | ** |
| Nostalgics of the former DDR | -0.052 | -0.24 | | 0.064 | 0.26 | | -0.255 | -0.63 | | 0.359 | 0.49 | |
| New middle class | -0.288 | -1.64 | + | 0.351 | 1.87 | + | 0.060 | 0.17 | | 0.980 | 1.66 | + |
| Materialists | -0.267 | -1.36 | | 0.336 | 1.58 | | 0.099 | 0.27 | | 0.219 | 0.32 | |
| Hedonists/escapists | -0.222 | -1.02 | | 0.276 | 1.15 | | 0.044 | 0.11 | | 0.769 | 1.13 | |
| Experimentalists | -0.565 | -2.24 | * | 0.508 | 1.84 | + | 0.635 | 1.47 | | 0.695 | 0.80 | |
| | | | | | | | | | | | | |
| Purchasing power per household > 530 € | -0.329 | -0.96 | | 0.452 | 1.22 | | -0.097 | -0.15 | | 0.331 | 0.29 | |
| Status | 0.018 | 0.71 | | -0.017 | -0.63 | | -0.029 | -0.61 | | 0.082 | 1.14 | |
| Eastern Germany | 0.075 | 0.37 | | -0.117 | -0.54 | | -0.199 | -0.50 | | 0.601 | 1.04 | |
| ***Interviewer variables*** | | | | | | | | | | | | |
| Isex (1 - men) | 0.074 | 0.54 | | -0.127 | -0.87 | | -0.271 | -1.02 | | 0.599 | 1.69 | + |
| Age of interviewer | 0.002 | 0.27 | | 0.000 | -0.05 | | -0.002 | -0.15 | | -0.023 | -1.16 | |
| Interviewer age < 40 & male | -0.508 | -1.32 | | 0.785 | 1.95 | + | 0.519 | 0.72 | | -0.981 | -0.86 | |
| Secondary modern school (ref.) | | | | | | | | | | | | |
| Secondary school | -0.068 | -0.40 | | 0.037 | 0.20 | | -0.004 | -0.01 | | -0.081 | -0.18 | |
| High school diploma | -0.455 | -1.67 | + | 0.401 | 1.39 | | 0.618 | 1.22 | | 0.597 | 0.89 | |
| University with and without degree | 0.008 | 0.04 | | 0.049 | 0.23 | | -0.158 | -0.41 | | -0.456 | -0.87 | |
| Workload | 0.018 | 3.26 | ** | -0.019 | -3.17 | ** | -0.018 | -1.63 | + | -0.024 | -1.78 | + |
| SOEP experience | 0.066 | 0.45 | | -0.210 | -1.36 | | 0.323 | 1.16 | | -0.059 | -0.16 | |

Table continued on next page

(continued)

| Variable | Model 1 Participation vs. non-participation | | | Model 2 Refusal vs. participation | | | Model 3 Non-contact vs. participation | | | Model 4 Unable vs. participation | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Coeff. | z value | | Coeff. | z value | | Coeff. | z value | | Coeff. | z value | |
| **Area description** | | | | | | | | | | | | |
| City | -0.442 | -2.58 | *** | 0.407 | 2.22 | * | 0.834 | 2.79 | ** | -0.047 | -0.10 | |
| Simple urban row estate (ref.) | | | | | | | | | | | | |
| Exclusive academic residential area | 0.613 | 1.24 | | -0.476 | -0.88 | | -1.034 | -1.12 | | -1.753 | -1.05 | |
| High-income families, new privately owned home | 1.030 | 2.16 | * | -0.770 | -1.48 | | -1.251 | -1.29 | | -3.404 | -2.31 | * |
| Distinguished detached houses | 0.760 | 2.10 | * | -0.638 | -1.61 | | -0.812 | -1.28 | | -2.289 | -1.92 | + |
| High-quality new detached houses- commuter belt | 0.957 | 1.94 | + | -0.847 | -1.55 | | -1.487 | -1.52 | | -0.488 | -0.39 | |
| High-quality new detached houses | 0.458 | 0.94 | | -0.197 | -0.38 | | -1.426 | -1.32 | | -0.557 | -0.44 | |
| Old city center | 0.834 | 2.33 | * | -0.605 | -1.55 | | -1.398 | -2.12 | * | -2.139 | -1.84 | + |
| New terraced houses in rural areas | 0.537 | 1.44 | | -0.278 | -0.68 | | -0.685 | -0.99 | | -2.138 | -1.86 | + |
| Simple houses in rural districts | 0.752 | 2.15 | * | -0.447 | -1.17 | | -1.496 | -2.25 | * | -1.630 | -1.55 | |
| Old apartment building | 0.515 | 1.44 | | -0.398 | -0.99 | | -0.707 | -1.29 | | -0.939 | -0.98 | |
| Social climber, upscale professions. outskirts | 0.771 | 2.00 | * | -0.697 | -1.63 | + | -1.064 | -1.67 | + | -1.424 | -1.36 | |
| Middle class in rural areas | 0.504 | 1.09 | | -0.321 | -0.63 | | -0.342 | -0.44 | | -1.686 | -1.20 | |
| Social housing, simple apartment buildings | 0.654 | 2.05 | * | -0.588 | -1.64 | + | -0.859 | -1.69 | + | -0.592 | -0.69 | |
| Unrenovated old apartment buildings | 0.443 | 1.30 | | -0.422 | -1.11 | | -0.226 | -0.44 | | -1.040 | -1.05 | |
| Apartment blocks, low standard | 0.644 | 1.66 | + | -0.530 | -1.21 | | -0.947 | -1.64 | | -16.290 | -0.01 | |
| Multi-cultural central zone | 0.387 | 0.91 | | -0.080 | -0.18 | | -0.586 | -0.80 | | -2.283 | -1.40 | |
| Highrise, basic standard | -0.018 | -0.06 | | -0.035 | -0.10 | | -0.050 | -0.10 | | 0.849 | 1.08 | |
| Old social housing | -0.301 | -0.72 | | 0.462 | 1.03 | | -0.180 | -0.30 | | 1.103 | 1.05 | |
| Attractive urban location | 0.022 | 0.06 | | 0.021 | 0.05 | | -0.382 | -0.70 | | 0.919 | 1.05 | |
| Social hotspot | 0.507 | 1.52 | | -0.416 | -1.11 | | -1.117 | -2.09 | * | 0.238 | 0.29 | |
| Young people in older apartments | 0.230 | 0.66 | | -0.129 | -0.33 | | -0.793 | -1.44 | | 0.499 | 0.59 | |
| Middle class in older accommodations | 0.629 | 2.08 | * | -0.478 | -1.42 | | -0.897 | -1.84 | + | -1.426 | -1.61 | |
| Lower class in apartments | 0.297 | 0.91 | | -0.195 | -0.54 | | -0.879 | -1.68 | + | -0.598 | -0.68 | |
| Solitary retired persons | 0.288 | 0.89 | | -0.316 | -0.88 | | -0.508 | -1.00 | | 0.192 | 0.24 | |
| Younger villager | 0.605 | 1.73 | + | -0.234 | -0.61 | | -1.856 | -2.61 | ** | -2.185 | -2.19 | * |
| Simple vocations in rural areas | 0.752 | 2.15 | * | -0.494 | -1.29 | | -1.118 | -1.83 | + | -2.364 | -2.07 | * |
| Low qualified worker | 0.643 | 1.71 | + | -0.428 | -1.05 | | -1.095 | -1.60 | | -2.109 | -1.88 | + |
| Self-employed in new buildings | 0.994 | 2.92 | ** | -0.842 | -2.24 | * | -1.382 | -2.19 | * | -2.015 | -2.01 | + |
| Manufacturer in rural areas | 0.320 | 0.74 | | -0.193 | -0.40 | | -0.632 | -0.80 | | -0.208 | -0.20 | |
| Socially disadvantaged small-towner | 0.609 | 1.10 | | -0.187 | -0.32 | | -14.650 | -0.02 | | -2.824 | -1.64 | + |
| Villages in outskirts | 0.842 | 2.44 | * | -0.732 | -1.92 | + | -1.087 | -1.71 | + | -1.366 | -1.47 | |
| Seniors in surrounding areas | 0.349 | 0.94 | | -0.218 | -0.54 | | -0.308 | -0.49 | | -1.253 | -1.32 | |
| Old families in outskirts | 1.028 | 3.00 | ** | -1.002 | -2.61 | ** | -0.724 | -1.30 | | -1.841 | -1.89 | + |
| Well-off retired persons in semi-detached houses | 0.692 | 1.93 | + | -0.658 | -1.64 | | -0.343 | -0.60 | | -1.865 | -1.81 | + |
| Older people in older houses | 0.006 | 0.01 | | 0.315 | 0.68 | | -0.712 | -0.99 | | -2.121 | -1.37 | |
| Well-off senior citizens in outskirts | 0.633 | 1.76 | + | -0.507 | -1.27 | | -0.919 | -1.51 | | -1.439 | -1.50 | |
| Older rural population | 0.921 | 2.49 | * | -0.642 | -1.59 | | -0.959 | -1.39 | | -16.343 | -0.02 | |
| Rural population | 0.267 | 0.51 | | 0.174 | 0.32 | | -1.784 | -1.40 | | -16.175 | -0.01 | |

(table continued on next page)

| | Model 1 Participation vs. non-participation | | | Model 2 Refusal vs. participation | | | Model 3 Non-contact vs. participation | | | Model 4 Unable vs. participation | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Variable | Coeff. | z value | | Coeff. | z value | | Coeff. | z value | | Coeff. | z value | |
| *Size of houses* | | | | | | | | | | | | |
| 1-2 family houses in homog. street section (ref.) | | | | | | | | | | | | |
| 1-2 family houses in non-homog. street section | -0.167 | -1.12 | | 0.158 | 1.00 | | 0.308 | 0.87 | | -0.033 | -0.07 | |
| 3-5 family houses | -0.011 | -0.06 | | -0.086 | -0.45 | | 0.519 | 1.35 | | -0.494 | -0.88 | |
| 6-9 family houses | 0.133 | 0.62 | | -0.078 | -0.34 | | 0.032 | 0.07 | | -0.976 | -1.52 | |
| Apartment buildings with 10 - 19 households | 0.290 | 1.11 | | -0.234 | -0.82 | | 0.157 | 0.32 | | -2.253 | -2.85 | ** |
| Highrise with 10 and more households | 0.636 | 1.49 | | -0.502 | -1.03 | | -0.127 | -0.19 | | -2.591 | -2.23 | * |
| Mainly commercial use | -0.392 | -0.87 | | 0.196 | 0.40 | | 1.529 | 2.06 | * | -1.725 | -1.03 | |
| *Family structure* | | | | | | | | | | | | |
| Mainly single household (Ref.) | | | | | | | | | | | | |
| Well above average share of single households | -0.005 | -0.02 | | 0.574 | 2.31 | * | -0.597 | -1.90 | + | -0.319 | -0.58 | |
| Above average share of single households | -0.130 | -0.60 | | 0.686 | 2.70 | ** | -0.765 | -2.28 | * | 0.255 | 0.47 | |
| Slightly above average share of single households | 0.104 | 0.47 | | 0.398 | 1.54 | | -0.778 | -2.25 | * | 0.078 | 0.14 | |
| Mixed family structure | -0.087 | -0.39 | | 0.704 | 2.71 | ** | -0.674 | -1.92 | + | -0.052 | -0.09 | |
| Slightly above average share of families with children | 0.169 | 0.73 | | 0.434 | 1.63 | | -1.062 | -2.78 | ** | 0.082 | 0.14 | |
| Above average share of families with children | 0.164 | 0.69 | | 0.460 | 1.69 | + | -1.211 | -3.08 | ** | -0.226 | -0.37 | |
| Well above average share of families with children | 0.135 | 0.55 | | 0.581 | 2.09 | * | -1.605 | -3.70 | *** | -1.052 | -1.56 | |
| Almost exclusively families with children | 0.278 | 1.08 | | 0.451 | 1.55 | | -1.536 | -3.13 | ** | -1.793 | -2.06 | * |
| | | | | | | | | | | | | |
| Number of household members | -0.007 | -0.72 | | 0.008 | 0.72 | | 0.000 | 0.01 | | 0.012 | 0.50 | |
| | | | | | | | | | | | | |
| ANONYM2 | 0.208 | 1.20 | | -0.180 | -0.98 | | 0.040 | 0.11 | | -0.147 | -0.31 | |
| ANONYM3 | 0.383 | 2.19 | * | -0.274 | -1.48 | | -0.456 | -1.12 | | -0.438 | -0.86 | |
| ANONYM4 | 0.245 | 1.38 | | -0.211 | -1.11 | | 0.202 | 0.51 | | -0.639 | -1.25 | |
| ANONYM5 | 0.225 | 1.23 | | -0.154 | -0.78 | | 0.195 | 0.51 | | -1.723 | -2.88 | ** |
| ANONYM6 | 0.391 | 2.06 | * | -0.325 | -1.60 | | -0.032 | -0.08 | | -0.395 | -0.72 | |
| ANONYM7 | -0.012 | -0.06 | | 0.034 | 0.15 | | 0.302 | 0.73 | | -0.111 | -0.19 | |
| ANONYM8 | 0.278 | 1.31 | | -0.306 | -1.32 | | 0.224 | 0.53 | | -0.558 | -0.91 | |
| ANONYM9 | 0.185 | 0.87 | | -0.194 | -0.85 | | 0.138 | 0.32 | | -0.484 | -0.76 | |
| | | | | | | | | | | | | |
| FOREIGN2 | 0.202 | 1.04 | | 0.003 | 0.01 | | -0.630 | -1.54 | | -1.217 | -1.98 | * |
| FOREIGN3 | 0.281 | 1.44 | | -0.118 | -0.56 | | -0.719 | -1.86 | + | -0.283 | -0.54 | |
| FOREIGN4 | 0.208 | 1.04 | | -0.025 | -0.12 | | -0.629 | -1.56 | + | -0.756 | -1.28 | |
| FOREIGN5 | 0.311 | 1.57 | | -0.122 | -0.57 | | -0.529 | -1.38 | | -1.593 | -2.55 | ** |
| FOREIGN6 | -0.028 | -0.13 | | 0.083 | 0.37 | | -0.308 | -0.78 | | 0.404 | 0.75 | |
| FOREIGN7 | 0.239 | 1.21 | | -0.050 | -0.23 | | -0.918 | -2.42 | ** | -0.065 | -0.12 | |
| FOREIGN8 | -0.113 | -0.55 | | 0.225 | 1.02 | | -0.220 | -0.58 | | 0.059 | 0.10 | |
| FOREIGN9 | -0.073 | -0.34 | | 0.142 | 0.60 | | -0.489 | -1.27 | | 0.949 | 1.75 | |
| | | | | | | | | | | | | |
| GARDEN2 | 0.239 | 1.24 | | -0.303 | -1.41 | | -0.035 | -0.11 | | -0.039 | -0.06 | |
| GARDEN3 | 0.289 | 1.39 | | -0.472 | -2.04 | * | -0.117 | -0.33 | | 0.851 | 1.41 | |
| GARDEN4 | 0.035 | 0.17 | | -0.093 | -0.41 | | -0.015 | -0.04 | | -0.047 | -0.07 | |
| GARDEN5 | 0.143 | 0.69 | | -0.137 | -0.60 | | -0.557 | -1.54 | | 0.729 | 1.21 | |
| GARDEN6 | -0.167 | -0.77 | | 0.096 | 0.41 | | 0.220 | 0.58 | | 0.722 | 1.11 | |
| GARDEN7 | -0.226 | -1.00 | | 0.153 | 0.63 | | -0.034 | -0.08 | | 0.921 | 1.36 | |
| GARDEN8 | -0.120 | -0.52 | | 0.049 | 0.20 | | 0.223 | 0.53 | | 1.054 | 1.50 | |
| GARDEN9 | -0.108 | -0.44 | | 0.028 | 0.11 | | 0.097 | 0.21 | | 0.972 | 1.26 | |

(continued)

| Variable | Model 1 Participation vs. Non-participation | | Model 2 Refusal vs. participation | | Model 3 Non-contact vs. participation | | Model 4 Unable vs. participation | |
|---|---|---|---|---|---|---|---|---|
| | Coeff. | z value | Coeff. | z value | Coeff. | z value | Coeff. | z value |
| MOVE | -0.003 | -0.13 | -0.033 | -1.16 | 0.082 | 1.74 + | 0.140 | 1.91 + |
| **Random effects** | | 95% Interval | | 95% Interval | | 95% Interval | | 95% Interval |
| $u_{ij}$. | $\pi^2 / 3$ | | $\pi^2 / 3$ | | $\pi^2 / 3$ | | $\pi^2 / 3$ | |
| $v_{0j}$. (intercept) | 0.400 | (0.38 - 0.83) | 0.395 | (0.39 - 0.94) | 1.300 | (1.25 - 3.0) | 1.370 | (0.52 - 1.6) |
| ICC | 0.108 | | 0.107 | | 0.283 | | 0.294 | |
| Interviewer | 227 | | 224 | | 219 | | 215 | |
| Household addresses | 3,408 | | 2,774 | | 1,825 | | 1,523 | |
| Log likelihood | -2,142 | | -1,805 | | -829 | | -405 | |
| Pseudo R² | 0.07 | | 0.07 | | 0.14 | | 0.22 | |

Note: Significance: + 10%; * 5 %, ** 1 %, *** 0.1 %

Source: SOEP 2006, Sample H and MICROM data set, authors' own calculations

Table 6 shows the estimates for version C of the model. In this version, we have broadened the model to include variables from the interviewer questionnaire study. The interviewer questionnaire was filled in by 187 of the 234 interviewers in sample H. Because of item nonresponse, we can only use the information of 165 interviews. These 165 interviewers worked on 2,592 household addresses. The implementation of additional interviewer variables reduces the interviewer variance and intraclass correlation in the random part by about 50%. The comparability of these results with the results from model version A and B can be problematic, however, because only two thirds of the interviewers in sample H filled out the questionnaire.

Nevertheless, overall we find approximately the same significant respondent and area variables as in version B of the model, although the focus here is on the interviewer variables. We can see that an interviewer's past SOEP experience is now significant, both increasing the probability of participation (β = 0.295) and decreasing the probability of refusals (β = -0.461). Furthermore, interviewers with higher educational degrees (high school or university) are less successful and result in more refusals (β = 0.625 and β = 0.437) and non-contacts (β = 0.975) than interviewers with lower education. As before, male interviewers have more "unable to participate" responses (β = 0.959), and younger male interviewers have higher refusal rates (β = 0.809).

Some findings of general interest (not only for SOEP) result from the personality variables and self-assessments of the interviewers. Self-assessments give insight into the self-perception of the interviewer but do not necessarily provide information about his or her true character. We can see that the probability of a successful interview increases if interviewers describe themselves as being above average amicable ($\beta$ = 0.193), reserved ($\beta$ = 0.147), less patient ($\beta$ = -0.097), if they are above average satisfied with their life ($\beta$ = 0.082) and if they do not get flustered easily ($\beta$ = -0.103). It seems that these interviewers are generally self-confident. They look forward to the future and have a positive coefficient for the desire to work on SOEP in future years (but not significant).

The probability of refusals increases if interviewers describe themselves as above average unsociable, if they are less amicable ($\beta$ = -0.231) but more patient ($\beta$ = 0.082), and if they are less reserved ($\beta$ = -0.104) and less satisfied with their life ($\beta$ = -0.081). The coefficient for the desire to work on the SOEP in future years is negative (but not significant).

The probability of non-contacts increases if interviewers describe themselves less original ($\beta$ = -0.397) and less reserved than average ($\beta$ = -0.336) but more imaginative ($\beta$ = 0.336).

We can find many significant self-assessment variables for the response variable "unable to participate." These interviewers describe themselves as above average sluggish ($\beta$ = 0.276), as patient ($\beta$ = 0.164), as able to forgive others ($\beta$ = 0.318) and easily flustered ($\beta$ = 0.285). They think they have a higher risky propensity ($\beta$ = 0.169). Furthermore, they are below average in terms of being reserved ($\beta$ = -0.235), inquisitive ($\beta$ = -0.341), and communicative ($\beta$ = -0.699). Particularly the last variable, the ability to communicate with others, is one of the most important key qualifications for an interviewer. However, it seems that these interviewers tend to be more passive and lacking in self-confidence. A further finding is the significant social desirability variable ($\beta$ = 0.773), which indicates that interviewers with a strong need for social approval get more "unable to participate" responses than interviewers without this trait. However, we have to interpret this strong effect with caution. The fieldwork organization that conducts the interviewer survey is also the employer of the interviewer. This relationship may bias

the interviewer response behavior. The estimates for the Sinus milieus show that respondents who use the response "unable to participate" come mainly from an upper conservative milieu. If we calculate the predicted probabilities for "unable to participate" for several situations, we see that the probability is highest for potential respondents who live in neighborhoods with "upper conservative households" in attractive urban locations, and doubles in the case of SD = 1 (see Table 9 in the Appendix). A possible explanation could be that households in traditional, higher social milieus use this excuse more often, and that interviewers who value social approval accept this excuse more readily than others.

**Table 6: Multilevel logit models, version C – including variables of the interviewer questionnaire**

| Variable | Model 1 Participation vs. non-participation | | | Model 2 Refusal vs. participation | | | Model 3 Non-contact vs. participation | | | Model 4 Unable vs. participation | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Coeff. | z value | | Coeff. | z value | | Coeff. | z value | | Coeff. | z value | |
| **Fixed effect** | | | | | | | | | | | | |
| (Intercept) | -2.961 | -1.93 | + | 1.275 | 0.73 | | 3.022 | 0.99 | | 1.278 | 0.34 | |
| **Head of the household** | | | | | | | | | | | | |
| Age < = 35 years (ref.) | | | | | | | | | | | | |
| Age > 35 - 40 years | -0.134 | -0.58 | | 0.244 | 0.96 | | 0.296 | 0.66 | | 0.041 | 0.06 | |
| Age > 40 - 45 years | -0.341 | -1.58 | | 0.298 | 1.25 | | 0.516 | 1.21 | | 0.576 | 1.01 | |
| Age > 45 - 50 years | -0.570 | -2.67 | ** | 0.620 | 2.64 | ** | 0.870 | 2.05 | * | 0.338 | 0.58 | |
| Age > 50 - 55 years | -0.071 | -0.32 | | 0.111 | 0.46 | | 0.447 | 1.01 | | -0.056 | -0.09 | |
| Age > 55 - 60 years | -0.165 | -0.73 | | 0.131 | 0.53 | | 0.719 | 1.62 | | -0.121 | -0.19 | |
| Age > 60 - 65 years | -0.347 | -1.51 | | 0.336 | 1.34 | | 0.867 | 1.88 | + | 0.299 | 0.49 | |
| Age > 65 y. | -0.403 | -1.67 | + | 0.484 | 1.85 | + | 0.638 | 1.29 | | 0.746 | 1.16 | |
| **Sinus milieus** | | | | | | | | | | | | |
| Well-established (ref.) | | | | | | | | | | | | |
| Post-materialists | -0.128 | -0.70 | | 0.231 | 1.16 | | -0.104 | -0.31 | | -0.759 | -1.18 | |
| Modern performers | -0.295 | -1.37 | | 0.374 | 1.56 | | 0.096 | 0.25 | | 0.036 | 0.06 | |
| Upper conservatives | -0.163 | -0.68 | | 0.204 | 0.78 | | -0.806 | -1.67 | + | 1.053 | 1.68 | + |
| Traditionalists | -0.111 | -0.57 | | 0.123 | 0.58 | | -0.563 | -1.50 | | 0.860 | 1.54 | |
| Nostalgics of the former DDR | -0.086 | -0.36 | | 0.225 | 0.85 | | -0.634 | -1.38 | | 0.091 | 0.13 | |
| New middle class | -0.286 | -1.48 | | 0.375 | 1.78 | + | -0.141 | -0.38 | | 0.476 | 0.83 | |
| Materialists | -0.103 | -0.47 | | 0.158 | 0.66 | | -0.091 | -0.23 | | -0.617 | -0.85 | |
| Hedonists/escapists | -0.136 | -0.57 | | 0.109 | 0.41 | | 0.078 | 0.19 | | 0.069 | 0.11 | |
| Experimentalists | -0.391 | -1.41 | | 0.440 | 1.42 | | 0.424 | 0.91 | | -0.805 | -0.91 | |
| Purchasing power per household > 530 € | -0.046 | -0.13 | | 0.225 | 0.58 | | -0.640 | -0.96 | | -0.039 | -0.04 | |
| Status | 0.044 | 1.55 | | -0.036 | -1.14 | | -0.046 | -0.89 | | 0.059 | 0.81 | |
| Eastern Germany | 0.155 | 0.90 | | -0.211 | -1.09 | | -0.120 | -0.34 | | 0.130 | 0.29 | |
| **Interviewer** | | | | | | | | | | | | |
| Isex (1 - men) | 0.079 | 0.53 | | -0.241 | -1.45 | | 0.235 | 0.76 | | 0.959 | 2.70 | * |
| Age of interviewer | -0.006 | -0.74 | | 0.010 | 1.10 | | -0.011 | -0.68 | | 0.000 | 0.00 | |
| Interviewer age < 40 & male | -0.266 | -0.60 | | 0.809 | 1.64 | + | -0.717 | -0.78 | | -0.770 | -0.73 | |
| Secondary modern school (ref.) | | | | | | | | | | | | |
| Secondary school | -0.211 | -1.16 | | 0.219 | 1.07 | | 0.269 | 0.71 | | -0.394 | -0.92 | |
| High school diploma | -0.634 | -2.27 | * | 0.625 | 1.99 | * | 0.975 | 1.80 | + | 0.313 | 0.51 | |
| University with and without degree | -0.296 | -1.34 | | 0.437 | 1.75 | + | 0.236 | 0.53 | | -0.583 | -1.13 | |
| Workload | 0.021 | 3.15 | ** | -0.024 | -3.24 | ** | -0.009 | -0.68 | | -0.043 | -2.56 | * |
| SOEP experience | 0.295 | 1.97 | + | -0.461 | -2.71 | ** | 0.014 | 0.05 | | -0.022 | -0.06 | |

(table continued on next page)

(continued)

| Variable | Model 1 Participation vs. non-participation | | | Model 2 Refusal vs. participation | | | Model 3 Non-contact vs. participation | | | Model 4 Unable vs. participation | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Coeff. | z value | | Coeff. | z value | | Coeff. | z value | | Coeff. | z value | |
| ***Interviewer questionnaire*** | | | | | | | | | | | | |
| *What kind of character do you have?* | | | | | | | | | | | | |
| Hard-working (1 - 7) | -0.115 | -1.01 | | 0.204 | 1.60 | | -0.050 | -0.21 | | 0.123 | 0.43 | |
| Communicative (1 - 7) | 0.079 | 0.83 | | -0.030 | -0.28 | | -0.109 | -0.55 | | -0.699 | -3.01 | ** |
| Sometimes too brusque (1 - 7) | 0.031 | 0.48 | | -0.057 | -0.79 | | 0.178 | 1.44 | | -0.340 | -1.86 | + |
| Creative (1 – 7) | 0.052 | 0.75 | | -0.013 | -0.16 | | -0.397 | -2.84 | ** | 0.259 | 1.37 | |
| Often worry about things (1 - 7) | 0.073 | 1.50 | | -0.096 | -1.75 | + | -0.049 | -0.50 | | -0.020 | -0.17 | |
| Forgive others (1 - 7) | -0.054 | -0.92 | | 0.052 | 0.79 | | -0.033 | -0.27 | | 0.318 | 2.01 | * |
| Sluggish (1 - 7) | -0.073 | -1.14 | | 0.079 | 1.09 | | 0.024 | 0.19 | | 0.276 | 1.82 | + |
| Amicable (1- 7) | 0.193 | 2.66 | ** | -0.231 | -2.87 | ** | -0.070 | -0.47 | | -0.073 | -0.40 | |
| Appreciate aesthetic things (1 - 7) | 0.002 | 0.04 | | 0.007 | 0.11 | | -0.002 | -0.02 | | -0.131 | -0.95 | |
| Easily flustered (1 - 7) | -0.103 | -1.76 | + | 0.095 | 1.44 | | 0.071 | 0.60 | | 0.285 | 1.99 | * |
| Efficient (1 - 7) | -0.092 | -0.97 | | 0.083 | 0.79 | | -0.208 | -0.99 | | 0.394 | 1.46 | |
| Reserved  (1 – 7) | 0.147 | 2.86 | ** | -0.104 | -1.80 | + | -0.336 | -3.16 | ** | -0.235 | -1.91 | + |
| Considerate (1 - 7) | 0.078 | 0.83 | | -0.084 | -0.79 | | 0.141 | 0.76 | | -0.334 | -1.41 | |
| Imaginative (1 - 7) | -0.031 | -0.50 | | -0.034 | -0.48 | | 0.336 | 2.62 | ** | 0.017 | 0.11 | |
| Relaxed (1 - 7) | -0.015 | -0.19 | | 0.009 | 0.10 | | 0.224 | 1.35 | | -0.114 | -0.59 | |
| Inquisitive ( 1- 7) | 0.074 | 0.92 | | -0.039 | -0.43 | | 0.030 | 0.19 | | -0.341 | -1.91 | + |
| | | | | | | | | | | | | |
| Own risk propensity ( 1 - 10) | -0.027 | -0.76 | | 0.022 | 0.55 | | -0.052 | -0.74 | | 0.169 | 1.83 | + |
| No trust in other people (1 - 4) | 0.219 | 1.29 | | -0.123 | -0.65 | | -0.475 | -1.31 | | -0.333 | -0.69 | |
| Able to count on someone (1 – 4) | 0.018 | 0.14 | | -0.092 | -0.60 | | -0.214 | -0.83 | | 0.537 | 1.55 | |
| Pay attention to foreigners (1 – 4) | -0.064 | -0.61 | | 0.026 | 0.23 | | 0.312 | 1.40 | | 0.136 | 0.55 | |
| Most people take advantage of others (1/0) | -0.287 | -1.59 | | 0.195 | 0.97 | | 0.277 | 0.76 | | 0.607 | 1.37 | |
| Most people are helpful (1/0) | 0.148 | 1.08 | | -0.067 | -0.43 | | -0.448 | -1.60 | | -0.300 | -0.98 | |
| Years for SOEP in future | 0.215 | 1.54 | | -0.239 | -1.53 | | 0.042 | 0.15 | | -0.297 | -0.92 | |
| Patient (1 – 10) | -0.097 | -2.75 | ** | 0.082 | 2.06 | * | 0.094 | 1.32 | | 0.164 | 1.94 | + |
| Life satisfaction (1 – 10) | 0.082 | 2.02 | + | -0.081 | -1.78 | + | -0.072 | -0.92 | | -0.277 | -2.79 | ** |
| SD (social desirability indicator built on six questions) | -0.163 | -1.12 | | 0.127 | 0.78 | | 0.252 | 0.86 | | 0.773 | 2.28 | * |

(table continued on next page)

34

(continued)

| Variable | Model 1 Participation vs. non-participation | | | Model 2 Refusal vs. participation | | | Model 3 Non-contact vs. participation | | | Model 4 Unable vs. participation | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Coeff. | z value | | Coeff. | z value | | Coeff. | z value | | Coeff. | z value | |
| *Area description* | | | | | | | | | | | | |
| City | -0.197 | -1.02 | | 0.114 | 0.53 | | 0.540 | 1.56 | | -0.852 | -1.73 | + |
| | | | | | | | | | | | | |
| Simple urban row estate (ref.) | | | | | | | | | | | | |
| Exclusive academic residential area | 0.416 | 0.85 | | -0.179 | -0.33 | | -1.416 | -1.59 | | -0.640 | -0.44 | |
| High-income families, new privately owned home | 0.481 | 0.98 | | -0.123 | -0.23 | | -1.302 | -1.39 | | -1.574 | -1.15 | |
| Distinguished detached houses | 0.630 | 1.74 | + | -0.408 | -1.01 | | -1.004 | -1.63 | | -1.579 | -1.21 | |
| High-quality new detached houses – commuter belt | 0.444 | 0.81 | | -0.349 | -0.57 | | -1.612 | -1.29 | | 0.729 | 0.61 | |
| High-quality new detached houses | 0.007 | 0.01 | | 0.386 | 0.67 | | -1.652 | -1.03 | | 0.372 | 0.26 | |
| Old city center | 0.462 | 1.27 | | -0.182 | -0.45 | | -1.589 | -2.44 | * | -0.695 | -0.64 | |
| New terraced houses in rural areas | 0.285 | 0.75 | | 0.058 | 0.14 | | -0.772 | -1.12 | | -0.758 | -0.68 | |
| Simple houses in rural districts | 0.941 | 2.52 | * | -0.512 | -1.25 | | -2.363 | -3.01 | ** | -1.142 | -1.09 | |
| Old apartment building | 0.419 | 1.12 | | -0.202 | -0.47 | | -0.949 | -1.67 | + | -0.039 | -0.04 | |
| Social climber, upscale professions, outskirts | 0.921 | 2.12 | * | -0.777 | -1.59 | | -0.987 | -1.42 | | -16.223 | -0.01 | |
| Middle class in rural areas | 0.309 | 0.62 | | 0.047 | 0.09 | | -0.982 | -1.07 | | -0.938 | -0.69 | |
| Social housing, simple apartment buildings | 0.675 | 1.97 | * | -0.486 | -1.25 | | -1.143 | -2.13 | * | -0.410 | -0.45 | |
| Unrenovated old apartment buildings | 0.182 | 0.47 | | -0.113 | -0.26 | | -0.227 | -0.39 | | -0.164 | -0.17 | |
| Block buildings low standard | 0.852 | 2.00 | * | -0.798 | -1.59 | | -1.419 | -2.28 | * | -16.392 | -0.01 | |
| Multi-cultural central zone | 0.153 | 0.34 | | 0.239 | 0.49 | | -0.755 | -0.98 | | -0.778 | -0.58 | |
| Highrise, basic standard | 0.299 | 0.87 | | -0.092 | -0.23 | | -1.039 | -1.88 | + | 0.697 | 0.92 | |
| Old social housing | 0.036 | 0.08 | | 0.257 | 0.51 | | -0.969 | -1.49 | | 1.087 | 1.08 | |
| Attractive urban location | 0.052 | 0.15 | | 0.172 | 0.42 | | -0.872 | -1.62 | | 0.338 | 0.38 | |
| Social hotspot | 0.438 | 1.23 | | -0.164 | -0.41 | | -1.456 | -2.52 | * | 0.697 | 0.89 | |
| Young people in older apartments | 0.154 | 0.42 | | 0.076 | 0.18 | | -1.043 | -1.87 | + | 0.747 | 0.91 | |
| Middle class in older accommodations | 0.733 | 2.34 | * | -0.550 | -1.56 | | -1.085 | -2.22 | * | -0.994 | -1.17 | |
| Lower class in apartments | 0.252 | 0.72 | | -0.158 | -0.40 | | -0.725 | -1.36 | | 0.271 | 0.32 | |
| Solitary retired persons | 0.327 | 0.94 | | -0.253 | -0.64 | | -0.713 | -1.30 | | 0.254 | 0.34 | |
| Younger villager | 0.329 | 0.92 | | 0.106 | 0.27 | | -2.056 | -2.72 | ** | -0.277 | -0.31 | |
| Simple vocations in rural areas | 0.577 | 1.53 | | -0.207 | -0.50 | | -1.556 | -2.34 | * | -1.272 | -1.17 | |
| Low qualified worker | 0.623 | 1.58 | | -0.327 | -0.75 | | -1.226 | -1.81 | + | -1.399 | -1.06 | |
| Self-employed in new buildings | 0.755 | 2.16 | * | -0.473 | -1.22 | | -1.687 | -2.60 | ** | -0.599 | -0.65 | |
| Manufacturer in rural areas | 0.291 | 0.65 | | 0.035 | 0.07 | | -0.907 | -1.17 | | 0.250 | 0.19 | |
| Socially disadvantaged small-towner | 0.885 | 1.27 | | -0.554 | -0.74 | | -13.731 | -0.03 | | 0.782 | 0.48 | |
| Villages in outskirts | 0.532 | 1.47 | | -0.359 | -0.89 | | -1.056 | -1.72 | + | 0.176 | 0.21 | |
| Senior citizens in surrounding areas | -0.047 | -0.12 | | 0.251 | 0.58 | | -0.822 | -1.28 | | 0.814 | 0.94 | |
| Old families in outskirts | 0.633 | 1.79 | + | -0.498 | -1.24 | | -0.999 | -1.77 | + | -0.345 | -0.38 | |
| Well-off retired persons in semi-detached houses | 0.232 | 0.60 | | -0.031 | -0.07 | | -0.460 | -0.74 | | -1.065 | -1.00 | |
| Older people in older houses | -0.072 | -0.15 | | 0.437 | 0.84 | | -1.065 | -1.26 | | -0.798 | -0.58 | |
| Well-off senior citizens in outskirts | 0.457 | 1.27 | | -0.179 | -0.44 | | -1.405 | -2.29 | * | -0.152 | -0.18 | |
| Older rural population | 0.579 | 1.45 | | -0.200 | -0.46 | | -1.444 | -1.93 | + | -16.095 | -0.01 | |
| Rural population | -0.218 | -0.37 | | 0.752 | 1.21 | | -1.457 | -1.12 | | -16.281 | 0.00 | |

(table continued on next page)

(continued)

| Variable | Model 1 Participation vs. non-participation | | Model 2 Refusal vs. participation | | | Model 3 Non-contact vs. participation | | | Model 4 Unable vs. participation | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Coeff. | z value | Coeff. | z value | | Coeff. | z value | | Coeff. | z value | |
| **Family structure** | | | | | | | | | | | |
| Mainly single household (ref.) | | | | | | | | | | | |
| Well above average share of single households | -0.007 | -0.03 | 0.771 | 2.70 | ** | -0.901 | -2.54 | * | -0.476 | -0.81 | |
| Above average share of single households | -0.027 | -0.11 | 0.713 | 2.42 | * | -1.059 | -2.87 | ** | 0.056 | 0.10 | |
| Slightly above average share of single households | -0.054 | -0.22 | 0.678 | 2.27 | * | -0.929 | -2.44 | * | 0.368 | 0.66 | |
| Mixed family structure | -0.072 | -0.29 | 0.777 | 2.60 | ** | -0.836 | -2.16 | * | 0.116 | 0.20 | |
| Slightly above average share of families with children | 0.131 | 0.52 | 0.601 | 1.97 | * | -1.068 | -2.59 | ** | 0.017 | 0.03 | |
| Above average share of families with children | 0.200 | 0.78 | 0.521 | 1.68 | + | -1.226 | -2.92 | ** | -0.437 | -0.68 | |
| Well above average share of families with children | 0.105 | 0.39 | 0.766 | 2.39 | * | -1.839 | -3.85 | *** | -0.770 | -1.09 | |
| Almost exclusively families with children | 0.066 | 0.23 | 0.811 | 2.39 | * | -1.831 | -3.24 | ** | -1.823 | -1.80 | + |
| Number of household members | 0.006 | 0.79 | 0.000 | 0.03 | | -0.020 | -1.77 | + | 0.011 | 0.53 | |
| MOVE | -0.009 | -0.33 | -0.029 | -0.94 | | 0.109 | 2.09 | * | 0.060 | 0.84 | |
| **Random effects** | | | | | | | | | | | |
| $u_{ij}$. | $\pi^2 / 3$ | | $\pi^2 / 3$ | | | $\pi^2 / 3$ | | | $\pi^2 / 3$ | | |
| $v_{0j}$. (intercept) | 0.180 | | 0.243 | | | 0.740 | | | 0.242 | | |
| ICC | 0.052 | | 0.069 | | | 0.184 | | | 0.069 | | |
| Interviewer | 165 | | 165 | | | 163 | | | 162 | | |
| Household addresses | 2,592 | | 2,111 | | | 1,409 | | | 1,184 | | |
| Log likelihood | -1641 | | -1368 | | | -636 | | | -310 | | |
| Pseudo $R^2$ | 0.07 | | 0.07 | | | 0.14 | | | 0.22 | | |

Note: Significance: + 10%; * 5 %, ** 1 %, *** 0.1 %

Source: SOEP 2006, Sample H and MICROM data set, authors' own calculation

## Table 7: Summary of findings in the multilevel logit models

| | Easy to persuade | Refusal more likely | Difficult to contact | Use "not able to participate" |
|---|---|---|---|---|
| **Head of household** | | | | |
| Sinus milieus | well-established | new middle class<br>experimentalists<br>modern performers | experimentalists<br>modern performers | upper conservatives<br>traditionalists<br>new middle class<br>modern performer |
| Age | age <= 35 years | age > 45 - 50 years | age > 45 - 50 years<br>age > 55 - 60 years | |
| Family structure | | families with children | single households | |
| **Interviewer** | | | | |
| Gender | | | | male interviewer |
| Age | | age < 40 years & male int. | | |
| Education | | higher level of education | higher level of education | |
| Workload | high | Low | | low |
| SOEP experience | increase | decrease | | |
| Character: | amicable<br>satisfied with own life<br>reserved<br>not easily flustered | not amicable<br>dissatisfied with own life<br>not reserved<br>Patient | not creative<br>not reserved<br>fancifulness | need for social approval<br>not communicative<br>dissatisfied with own life<br>sluggish<br>not inquisitive<br>easily flustered<br>not reserved<br>patient |
| **Area** | smaller than cities | Cities | cities | smaller than cities |
| MOSAIC type | high-income families<br>new private owned house<br>old families in outskirts<br>self employed in new buildings<br>high-quality new detached houses<br>social climber, upscale professions | simple urban estate<br>less for social climber, upscale prof.<br>less for old families in outskirts<br>less for villages in outskirts<br>less for self employed in new buildings | less for younger villagers<br>less for self-employed in new houses<br>less for old city centers<br>less for simple houses in rural districts<br>less for simple vocations in rural areas | less for social disadvantaged small-towner<br>less for high-income families, new private home<br>less for simple vocations in rural areas |
| Frequency of moves | | | increase | increase |

# 4    Illustration of the Nonresponse Bias at Household Level

In order to give a better idea of the bias due to nonresponse, we use the "gap" between the empirical distribution of the microgeographic categories in the gross sample and the net sample (realized sample). And we use this gap to construct "design weights" (or "correction factors") for the net sample.[10]

Often, this gap is called "initial nonresponse". A variable $c_i$ is defined for all enumerated household addresses:

$$c_i = \begin{cases} 1, \text{if household } i \text{ is in the sample} \\ 0, \qquad\qquad \text{otherwise} \end{cases}$$

We can consider two cases. If it is a random nonresponse "0", treatment is not necessary and respondents can be considered as non-selected. However, the analyses above show that this is not the case here: the non-response is not random but connected to a few variables.

Assume we are interested in the total of a variable $y_i$, for example, the household income of household i.

$$t_y = \sum_{i=1}^{N} y_i$$

Then we could use the common Horvitz-Thompson (HT) estimator to estimate the gross sample total of $y_i$

$$\hat{t}_y = \sum_{i=1}^{N} w_i y_i \quad \text{with} \quad E(w_i) = 1$$

$$w_i = \frac{c_i}{\pi_i}$$

---

[10] We use a common cause model and assume that the chosen microgeographic variables are causal factors for the survey participation decision. These variables are measured for respondents and nonrespondents. Hence, there is hope that we can remove the nonresponse bias by weighting class adjustment using these variables (see Groves 2006).

with $\pi_i = E(c_i) = P(c_i = 1)$. This means that each household in the sample is weighted with the inverse of its selection probability in order to apportion the nonresponse among respondents.

In the previous section, we showed that the microgeographic variables "Sinus milieu," "age," "family structure" and "size of township" are relevant for the nonresponse process. We use these variables as well as the variable "eastern Germany" to calculate the predicted values of the selection probability (fixed by a logistic regression). When interpreting the results, we have to be aware that the nonresponse process occurs at household level, but the correction can only be made at a microgeographic level. Overall, Table 8 shows that the unweighted and weighted shares and average values from several important household variables are all rather close. In most cases we cannot find any serious bias. Nevertheless, we can see that the share of apartment owners is overestimated and the share of main tenants is underestimated in the sample due to nonresponse. After the correction has been made, the share of main tenants is higher than that of apartment owners.

**Table 8: Comparison of unweighted and weighted measurements**

| Variable | Share of "yes" | | | Average amount | | |
|---|---|---|---|---|---|---|
| | Unweighted | Weighted | Est. bias (%) | Respondent | Weighted | Est. bias (%) |
| Child allowance 2005 | 34.0 | 32.8 | 3.7 | €254.7 | €253.0 | 0.7 |
| Child allowance 2006 | 33.1 | 32.0 | 3.4 | €256.4 | €254.4 | 0.8 |
| Unemployment benefit II (ALG II 2005) | 8.5 | 8.9 | -4.5 | €547.0 | €540.6 | 1.2 |
| Unemployment benefit II (ALG II 2006) | 7.8 | 8.4 | -7.1 | €545.6 | €535.1 | 2.0 |
| Housing assistance 2005 | 3.6 | 3.8 | -5.3 | €192.5 | €192.1 | 0.2 |
| Housing assistance 2006 | 3.5 | 3.6 | -2.8 | €159.1 | €166.0 | -4.1 |
| Support for care of sick family 2005 | 1.9 | 1.8 | 5.6 | €360.8 | €376.7 | -4.2 |
| Support for care of sick family 2006 | 2.1 | 1.9 | 10.5 | €383.0 | €380.0 | 0.8 |
| | | | | | | |
| Main tenant | **47.6** | **52.7** | -9.7 | | | |
| Subtentant | 1.1 | 1.0 | 10.0 | | | |
| Apartment owner | **51.1** | **46.2** | 10.6 | | | |
| Total living space in apt. | | | | 102.8 qm | 98.2 qm | 4.6 |
| Number of rooms in flat | | | | 4.0 | 3.8 | 4.2 |
| Monthly rent | | | | €470.0 | €469.9 | 0.0 |
| | | | | | | |
| Household income | | | | €2,300.4 | €2,244.5 | 2.5 |
| Observations | 1,509 | 3,794 | | 1,509 | 3,794 | |

Source: SOEP 2006, Sample H; authors' own calculations

# 5    Conclusion

In this study, we have investigated why households (addresses) that were randomly selected for a survey interview did not participate. Our basis is the refreshment sample H of the SOEP. We used information from 3,931 German household addresses, detailed information about interviewers from an "interviewer survey" and an interviewer data set based on bookkeeping data from the fieldwork agency, and microgeographic characteristics of the neighborhood of the selected household addresses.

Nonresponses in the SOEP study are categorized according to three reasons: refusals, non-contact, and "unable to participate." We found that refusals and non-contacts as well as "inability to participate" relate to different respondent characteristics at the microgeographic level. The description in Section 3 and the estimates in the multivariate analysis in Section 6 presented evidence that younger respondents and respondents living in the "well-established milieu" are comparatively easy to persuade. Respondents middle class areas, respondents in areas with children and middle-aged people, however, are more likely to refuse. Milieus of "experimentalists" and "modern performers," and households with potentially middle-aged and older heads of households as well as households where singles most likely live are often difficult to contact.

One surprising result of our study was the empirical explanation for the interview outcome "unable to participate." We found that this does not indicate illness of respondents as expected, but there is evidence that it is an alibi used by households in higher traditionalistic milieus such as "upper conservatives" to avoid participation by an excuse which is diplomatic and easy to communicate. We can interpret this behavior as a "soft refusal." Moreover, our estimates show that this outcome is more likely with interviewers who describe themselves as above average "uncommunicative," "sluggish," "not inquisitive" and "dissatisfied with life" and highest for those who have a strong need for social approval. The latter interviewer characteristic doubles the probability that the reporting "household was unable to participate".

The self-assessments of interviewers with higher shares of refusals pointed in a similar direction. They more often describe themselves as dissatisfied with life and unsociable.

They have lower workloads and are often new to the SOEP staff. Furthermore, the probability of refusals increases with younger male and better educated interviewers. In big cities and urban areas, the refusal rates are higher than in the outskirts.

The interview outcome "non-contact" was more likely among interviewers who describe themselves as above average "not creative" and "not reserved". They often have a higher level of education. We also found more non-contacts in big cities and fewer in rural districts. The frequency of moves in the area significantly increases this outcome.

The best results—that is, the highest probabilities of securing and completing an interview—were achieved by experienced interviewers who described themselves as more self-confident, amicable, and satisfied with life. They tend to have a higher workload than other interviewers. Causal interpretation must be handled with caution: It could be the case that success in the interview process make interviewers more confident and happier with life. Furthermore, higher response rates are achieved in "good areas" where high-income families live, in areas with new privately owned houses, in areas where self-employed people live in new houses, and among old families in the outskirts. These results point towards potential weighting schemes using proxy measures and other correlates of survey outcomes to adjust for nonresponse (Kreuter et al. 2010).

In the last section, we used the gap between the frequency distributions of the microgeographic characteristics in the gross sample (enumerated addresses) and the net sample (realized sample) to construct design weights (or "correction factors") for the sample realized at household level. We found that the estimated nonresponse bias for the most selected variables is more or less negligible, despite statistically significant gaps. Nevertheless, one exception is the share of apartment owners, which turns out to be overestimated and the share of main tenants, which appears to be underestimated in subsample H. Based on this result  re-weighting of the survey households is easy to do.

# References

Daeubler, Th. 2002. Nonresponseanalysen der Stichprobe F des SOEP. Materialien 15. DIW Berlin.

Esser, H. 1986. Über die Teilnahme an Befragungen. *ZUMA-Nachrichten* 18: 38-47.

Esser, H.(1990). "Habits", "Frames" und "Rational Choice". Die Reichweite von Theorien der rationalen Wahl (am Beispiel der Erklärung des Befragtenverhaltens). *Zeitschrift für Soziologie* 19(4): 213-247.

Goebel, J., C. K. Spiess, N. R. J. Witte and S. Gerstenberg. 2007. Die Verknüpfung des SOEP mit MICROM-Indikatoren: Der MICROM-SOEP Datensatz. DIW Data Documentation No. 26. Berlin.

Goyder, J. 1987. The Silent Minority - Nonrespondents on Sample Surveys. Cambridge: Polity Press.

Groves, R.M. and Couper, M. 1998. Nonresponse in household interview surveys. New York: Wiley.

Groves, R.M. 2006. Nonresponse Rates and Nonresponse Bias in Household Surveys. *Public Opinion Quarterly,* 70(5), 646-675.

Hintze, P. and T. Lakes. 2009. Geographically Referenced Data in Social Science - A Service Paper for SOEP Data Users. DIW Data Documentation No. 46. Berlin.

Hox, J.J. 1994. Hierachical regression models for interviewer and respondent effects. *Sociological Methods and Research,* 22(3), 300-318.

Knies, G. and C. K. Spiess. 2007. Regional Data in the German Socio-Economic Panel Study (SOEP) . DIW Data Documentation No. 17. Berlin.

Koch, A. 1991. Zum Zusammenhang von Interviewermerkmalen und Ausschöpfungsquote. *ZUMA-Nachrichten,* 28, 41-53.

Koch, A. 1997. Teilnahmeverhalten beim ALLBUS 1994. *Kölner Zeitschrift für Soziologie und Sozialpsychology,* 49, 66-90.

Krason A., M. Mayr and M. Schipperges. 2003. The modern upper middle classes in Western and Central Eastern Europe. Presentation on ESOMAR-Congress "Excellence in Consumer Insights", April, 1st. 2003, Madrid.

Kreuter, F. et al. 2010. Using proxy measures and other correlates of survey outcomes to adjust for non-response: examples from multiple surveys. *Journal of the Royal Statistical Society, Series* A, 173(3).

Kreuter, F. and C. Casas-Cordero. 2010. Paradata. RatSWD Working Paper. Berlin.

Lynn, P. 2008. The problem of nonresponse in: DeLeeuw, E.D.; J.J. Hox, and D.A. Dillman. (Eds.), *International Handbook of Survey Methodology*, New York - London: Lawrence Erlbaum, 35-55.

Pickery, J., Loosveldt, G. and A. Carton. 2001. The effects of interviewer and respospondent characteristics on response behaviour in panel surveys. *Sociological Methods and Research,* 29(4), 509-523.

Plöger, W., M. Schipperges, and M. Mayr. 2005. Die Sinus-Milieus in Deutschland, Österreich und der Schweiz: Hintergrund, Ergebnisse, Anwendung. microMarketing Forum 2005. Sinus Sociovision, Heidelberg.

Rendtel, U., and T. Harms. 2009. Weighting and Calibration for Household Surveys. In: P. Lynn (ed.), *Methodology of Longitudinal Surveys*, Chichester: John Wiley & Sons, 265-286.

Rosenbladt, B. von, S. Stimmel and F. Stutz. 2006. SOEP 2006 – Erstbefragung der Ergänzungsstichprobe H: Methodenbericht. Munich: http://www.diw.de/documents/dokumentenarchiv/17/diw_01.c.83169.de/meth_2006_h.pdf

Schneekloth U. and I. Leven. 2003. Woran bemisst sich eine "gute" allgemeine Bevölkerungsumfrage? Analysen zu Ausmaß, Bedeutung und zu den Hintergründen von Non-response in zufallsbasierten Stichprobenerhebungen am Beispiel des ALLBUS. *ZUMA-Nachrichten* 53, 16-57.

Schnell, R. 1997. Nonresponse in Bevölkerungsumfragen. Opladen: Leske und Budrich.

Schraepler, J.-P. 2000. Was kann man am Beispiel des SOEP bezüglich Nonresponse lernen? *ZUMA-Nachrichten* 46, 177-149.

Schraepler, 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(1), 118-156.

Schraepler, J.-P. 2006. Explaining Income-Nonresponse - A Case Study by Means of the British Household Panel Study (BHPS). *Quality & Quantity, 40(6) , 1013 - 1036.*

Schraepler, J.-P. and G.G. Wagner. 2001. Das Verhalten von Interviewern - Darstellung und ausgewählte Analysen am Beispiel des Interviewer-Panels des Sozio-oekonomischen Panels. *Allgemeines Statistisches Archiv*, 85(1), 45-66.

Siegel, N. and S. Stimmel. 2007. SOEP-Interviewerbefragung 2006: Methodenbericht. Munich: http://www.diw.de/documents/dokumentenarchiv/17/diw_02.c.239773.de/meth_2006_int erviewerbefragung.pdf**.**

Singer, E., Mathiowetz, N. and M. Couper. 1993. The impact of privacy and confidentiality concerns on survey participation: The case of the 1990 U.S. Census. *Public Opinion Quarterly,* 57, 465-482

Sudman, S. and N. Bradburn. 1974. Response Effects in Surveys. Chicago: Aldine Publishing Company.

Thompson, S. K. (2006). Targeted Random Walk Designs. *Survey Methodology* 32(1): 11-24.

Wagner, G. G., R. V. Burkhauser, and F. Behringer. 1993. "The English Language. Public Use File of the German Socio-Economic Panel." *Journal of Human Resources* 28(2), 413–415.

Wagner, G. G., J. R. Frick and J. Schupp. 2007. The German Socio-Economic Panel Study (SOEP) – Scope, Evolution and Enhancements. In: *Schmollers Jahrbuch*, 127(1), 139-169.

Weinhardt, M, F. Kreuter, and J. Schupp. (2010). Interviewer Personality and Survey Error: Interviewer's Conscientiousness as Predictor for Item- and Unit Nonresponse in a Panel Household Study. SOEPpaper *(forthcoming)*.

# Appendix

**Predicted probabilities for response variable "unable to participate" for selected situations**

| Sinus milieu | Area | SD | P(Y = 1) Unable to participate |
|---|---|---|---|
| Upper conservative | Attractive urban location | 1 | 0.350 |
| | | 0 | 0.199 |
| | Self-employed in new buildings | 1 | 0.175 |
| | | 0 | 0.089 |
| | Exclusive area, academic area | 1 | 0.168 |
| | | 0 | 0.085 |
| | Middle class in rural areas | 1 | 0.130 |
| | | 0 | 0.065 |
| | Middle class in older apartments | 1 | 0.124 |
| | | 0 | 0.062 |

Note: respondent's age = 40-45 years. Other variables in the equation are evaluated at the sample mean.

Source: SOEP. sample H. interviewer questionnaire; authors' own estimations.