SOEP Das Sozio-oekonomische Panel

626
2013

SOEPpapers
on Multidisciplinary Panel Data Research

Conversion of Non-Respondents in an Ongoing Panel Survey: The Case of the German Socio-Economic Panel (SOEP)

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

DIW SOEP

Conversion of Non-Respondents in an Ongoing Panel Survey:
The Case of the German Socio-Economic Panel (SOEP)

Jörg-Peter Schräpler[1,2,4], Jürgen Schupp[2,4], and Gert G. Wagner[2,5]

1) Ruhr University Bochum, Germany
2) German Institute for Economic Research (DIW Berlin), Germany
3) Centre for interdisciplinary regional research (ZEFIR), Germany
4) Free University Berlin (FUB), Germany
5) Max Planck Institute for Human Development, Berlin, and Berlin University of
Technology (TUB), Germany

December 2013

Abstract

The results of a resurvey of non-respondents to the SOEP study carried out in 2006 show that this special effort of reinterviewing was relatively ineffective in two respects. First, the rate of successful conversions of passive to active respondents was low (less than 20 percent). Second, the composition of the longitudinal file did not improve. The same groups that showed high dropout rates in the SOEP (particularly low-income respondents and respondents in major cities) showed lower conversion rates in the resurvey. The resulting file of active respondents after the resurvey is therefore even more selective than the original file with regard to population means. Furthermore, due to the low rate of resurveying, the resurvey did not significantly improve the overall statistical power of the SOEP sample. In sum, the resurvey was an interesting experiment, but it shows little potential as a means for improving the standard fieldwork of SOEP.

1. Introduction

This study examines the potential benefits of efforts to convert inactive into active respondents in an ongoing longitudinal survey: the German Socio-Economic Panel Study (SOEP). In recent years, fieldwork organizations have been faced with declining response rates for most surveys of individuals and households. This phenomenon is pronounced in a number of European countries. In a cross-national study, De Leeuw and de Heer (2002) conducted a meta-analysis of non-response trends in 16 European surveys over a 20-year period and found that the noncontact rate had increased on average by 0.2% per year and the refusal rate by 0.3% per year. These increasing non-response rates can affect the quality of survey estimates (Singer 2006, Groves 2006). If the likelihood of responding is strongly related to a variable of interest in the survey, the non-response bias for the respondent mean will be large (Groves et al. 2009, p. 189).

Considerable resources have therefore been devoted to maximizing response rates and to reducing non-random non-response (selectivity) in surveys. Several response optimization techniques have been explored, including incentives for respondents (Singer 2002) and extended contact strategies for interviewers. One common fieldwork strategy is the conversion of non-respondents (referred to in the context of cross-sectional surveys as "refusers") into respondents (DeMaio 1980, Stoop 2004, Burton et al. 2006). In this case, it is important to understand how converted cases affect survey results in terms of response rates and especially data quality (Retzer et al. 2004, p. 4,984). Several studies have found no significant differences in results between respondents who never refused and converted refusers (Keeter 2000 and Curtin et al. 2000), indicating that the efforts and costs of refusal conversion do not pay off because converted refusers have the same characteristics as regular respondents but are significantly more expensive to survey. Furthermore, Miller and Wedeking (2003) showed that the quality of the data from refusal conversion is inferior to data on respondents who never refused to participate. However, all these studies dealt with cross-sectional surveys and thus with individuals who had never participated in the respective survey: the problem was therefore to convert an unwilling person into a first-time respondent. The present study, in contrast, focuses on the benefits of conversion in the case of an ongoing longitudinal survey, the German Socio-Economic Panel Study (SOEP), and thus with individuals who have taken part at least once.

In ongoing longitudinal studies, fieldwork organizations often do not treat spontaneous unwillingness to participate as a final outcome but rather make further attempts to convince the individual to participate in the study again. Burton et al (2006) studied the extent to which BHPS sample members were successfully converted after a spontaneous non-response in one wave into respondents in subsequent waves, examining the effect of the conversion on

the sample composition and survey estimates. In our paper, we do not look at respondents who were successfully "converted" during the standard SOEP fieldwork but instead focus on former SOEP respondents who expressed unwillingness to take part in a subsequent wave and after several additional attempts by the fieldwork organization TNS Infratest Sozialforschung were treated as dropouts.

On the basis of a special project (financed by the Federal Department of Education and Research),[1] TNS Infratest Sozialforschung undertook an extensive conversion study in 2006 and contacted all SOEP non-respondents from 2001 to 2004 to explore the opportunities for possible non-respondent conversion (TNS Infratest 2011). In this study, we analyze the main reasons for non-response and examine which members of this group were willing to take part in the SOEP again after a "break." We explore the effect of including converted non-respondents in the main SOEP again, also using selected substantive estimates to determine whether the effort to convert non-respondents pays off in terms of better research results.

Another aspect often disregarded in previous research is that the non-response process may vary by geographic location. There have been a few studies showing that the tendency to cooperate also depends on microgeographic characteristics of the sampled regions (Schräpler et al. 2010). In our study, we do not use microgeographic characteristics but investigate instead whether non-respondents in the SOEP were geographically selective at the county level. We also take local geographic effects into account in our statistical models.

2. A Resurvey of Non-Respondents

The SOEP survey is a longitudinal representative survey collecting socio-economic information on private households in the Federal Republic of Germany (Wagner et al. 2007). The SOEP infrastructure unit at DIW Berlin (German Institute for Economic Research) manages the SOEP study.

In the SOEP, all household members aged 16 and older are surveyed. The first wave of data, collected in 1984 in the Federal Republic of Germany—that is, prior to German reunification—contains 5,921 households (samples A and B). The original sample was supplemented by a sample of East German residents in 1990 (2,179 households, sample C) and a sample of immigrants in 1994-1995 (522 households, sample D). Additional refreshment samples were added in 1998 (1,056 households, sample E), 2000 (6,052

households. sample F) and 2006 (1,506 households, sample H). In 2002, an oversample of high-income earners was launched (1,224 households, sample G).

The resurvey of non-respondents was carried out beginning of 2007 by mail and focused on the 2001 to 2004 cohorts that make up SOEP subsample F (which began in 2000).[2]

The main purposes of the resurvey were:

- to explore how many former respondents to the SOEP would again be willing to take part in at least a short form of the survey
- to explore the reasons for their non-response.

In the SOEP, non-response is declared a "temporary refusal" when several attempts were made and failed to persuade a respondent to participate in a particular wave. A temporary refusal is classified as a "dropout" if further attempts and contacts appear useless (Haisken-DeNew/Frick 2005, p. 164). Table 1 shows the distribution of the respondents in subsamples A–F who participated in at least one of the years 2000, 2001, 2002, or 2003 and dropped out in the years 2001–2004. The results show that 60.7% of the former respondents were classified as "dropouts" one year after their last individual interview.

Table 1: Building the sample for the resurvey of non-respondents in 2001-2003

Last successful individual interview	Non-Respondents classified as "dropouts" of the survey in …				
	2001	2002	2003	2004	Total
2000	1,521	731	35	7	2,294
2001		733	636	27	1,396
2002			423	527	950
2003			8	344	352
Total	1,521	1,464	1,102	905	4,992

Sources: SOEP Resurvey 2006; TNS Infratest 2011

In 2006, the fieldwork organization TNS Infratest Sozialforschung sent out a short two-sided questionnaire by mail to all 4,992 non-respondents during the years under consideration. In this letter, TNS underscored that if respondents were unwilling to participate in the survey this would be respected, but also said that the fieldwork phase was not yet completed and that TNS would like to know whether it had been conducted correctly. In the case that a respondent had been misclassified as unwilling to participate, TNS underscored that they would be happy to reactivate that individual as a survey participant.

Table 2 shows that in 923 cases, a mail interview was carried out, resulting in a response rate of 18.5%. In addition, in 55 other cases, information about outcomes such as final

---

[2] The resurvey was part of the "SOEP 2006 Innovation Projects" (TNS Infratest 2011) and started in 2006.

refusal (0.4%), death of respondent (0.4%), or undeliverable mail (0.7%) was recorded, but for over 80% of the cases there was no response (no return of the questionnaire) from the respondent. It remains unclear whether this percentage of outcomes can be attributed to undeliverable mail, refusals, or other causes.

Table 2: Result of the fieldwork

	Number	%
Interview conducted	923	18.5
Refusal	18	0.4
Deceased	21	0.4
Undeliverable	34	0.7
No response (no return)	3,996	80.1
Total	4,992	100

Sources: SOEP Resurvey 2006; TNS Infratest Sozialforschung 2011.

Table 3 shows the distribution of the mail interviews conducted with "non-respondent cohorts." The majority of the participants in the resurvey had their last successful individual interview in the year 2000 and were classified as dropouts one year later. These respondents originated mainly (235 out of 281) from subsample F and participated only in one wave. The non-respondents in 2002 came in large part (61.5%) from subsample F as well.

Table 3: Successful mail interview by non-respondent cohorts

Last successful Individual interview	Non-respondents classified as "dropout" of the survey in …				
	2001	2002	2003	2004	Total
2000	281	130	6	1	418
2001		122	105	6	233
2002			93	94	287
2003				85	85
Total	281	252	204	186	923

Sources: SOEP Resurvey 2006; TNS Infratest Sozialforschung 2011; authors' calculations.

In Table 4, we see the interview outcomes of the resurvey for the various SOEP subsamples. For sample A, the response rate was 23.8% (159/669), for sample F 19.1% (604/3,168) and sample E 16.9% (44/261). In the case of samples B (foreigners), C (East Germans), and D (immigrant sample), the response rate was lower than 10%.

Table 4: Interview outcome by SOEP subsample

Subsample	Interview outcome					
	Successful mail interview	Refused	Deceased	Undeliverable	No response	Total
A	159	6	2	2	500	669
B	21	0	0	2	273	296
C	82	1	3	3	376	465
D	13	0	0	3	117	133
E	44	0	1	3	213	261
F	604	11	15	21	2,517	3,168
Total	923	18	21	34	3,996	4,992

Sources: SOEP Resurvey 2006; TNS Infratest Sozialforschung 2011; authors' calculations.

Next we will examine the interview outcome by respondent characteristics such as age, gender, and household variables and present some bivariate statistics. We compare the conditional distributions of survey participants with those of no return, refusals, undeliverable mail and deceased sample members.

Figure 1 shows that participation in the resurvey is strongly related to respondents' age. We see that a higher percentage of respondents aged 60 to 70 took part in the resurvey. Furthermore, it comes as no surprise that the share of respondents in the category "deceased" increased with respondents' age.

Figure 1: Interview outcome by age of the respondent
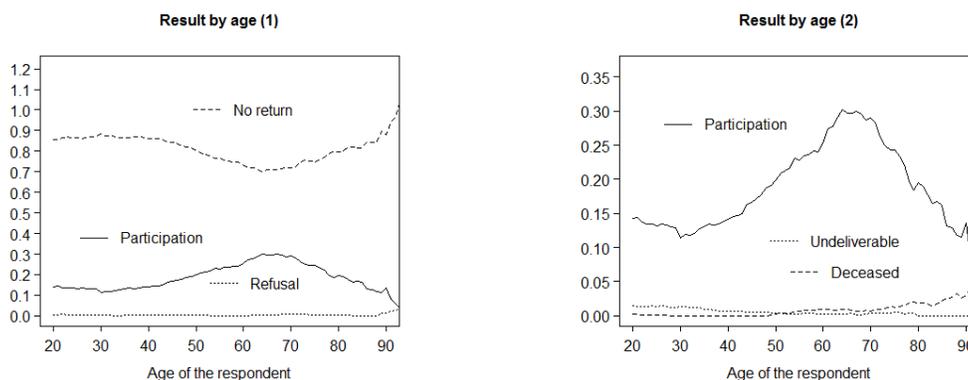
Sources: SOEP Resurvey 2006; authors' calculations.

Figure 2 shows the interview outcome by the reported size of respondents' households (HHGR) in the last completed interview. Note that this value may differ from household size in the year of the resurvey. Participation is rather low in the case of single households and highest in the case of two- and three-person households. A further increase in the size of the household, however, leads to a lower participation rate.

Figure 2: Interview outcome by number of persons in respondent's household.
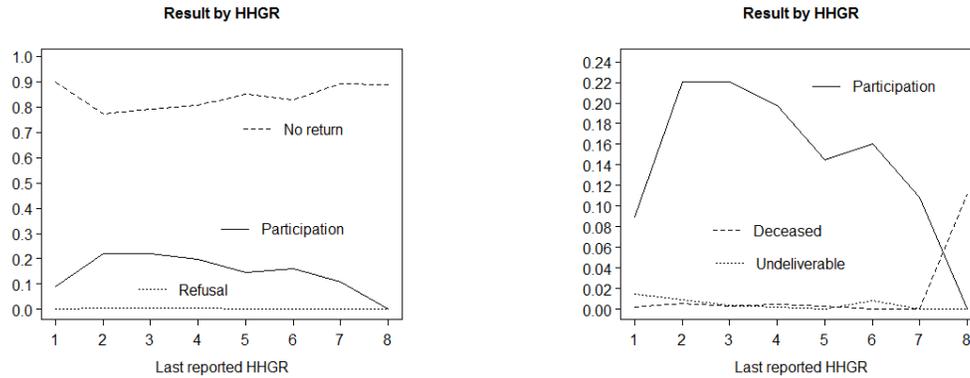
Sources: SOEP Resurvey 2006; authors' calculations.

In Figure 3, we examine interview outcomes in relation to the last reported equivalent household income of the respondent. We use quintiles to ensure the comparability over the different years of non-response. Figure 3 shows an almost linear increase of participation in the resurvey by equivalent income. High earners (5th quintile) have a response rate more than twice as high as low earners (1st quintile). This finding is important because it shows that the below-average participation rate of low-income households in the regular SOEP cannot be overcome by regular resurveying and efforts at non-respondent conversion.

Figure 3: Interview outcome by equivalent HH- income of the respondent (quintiles)
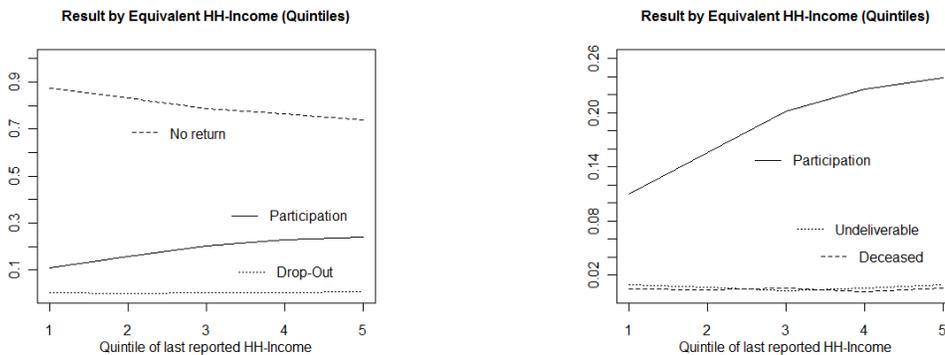
Sources: SOEP Resurvey 2006; authors' calculations.

Table 5 shows the frequency of the main reasons for classifying respondents as dropouts. Unfortunately, only 291 of the 923 non-respondents interviewed reported this information. As the table shows, 43% of them mentioned that the SOEP questionnaires were too complex and too long and 17% were bored or no longer took time to participate. Hence, approximately 60% of non-response was due to motivational issues. Reasons such as illness, advanced

age, care for relatives, and death in the family represent 19%, fieldwork-related reasons such as non-acceptance of interviewer, loss of contact, and concerns with privacy protection represent 16% of all dropouts by refusal.

Table 5: Main reasons for dropping out

	N	%
Questionnaire too complex/long	125	43.0
Bored, not willing to take time	49	16.8
Illness, advanced age	35	12.0
Care for relatives or death in family	22	7.6
Temporary loss of contact	19	6.5
Interviewer non-acceptance	16	5.5
Privacy concerns	10	3.4
Other reasons	15	5.2
Total	291	100

Sources: SOEP Resurvey 2006; authors' calculations.

Since only one-third of those who were reinterviewed reported their main reasons for dropping out, we abstain from further analysis of this variable. The main goal of the study was to measure the percentage of conversions. Table 6 shows that of all former respondents classified as dropouts, approximately 14% could be convinced take part in the SOEP again. This group can be referred to as "converted non-respondents" and will be investigated in more detail below.

Table 6: Share of non-respondents who were willing to take part again

	N	%
Respondent willing to take part again	715	14.3
Respondent unwilling to take part	198	4.0
Missing value	10	0.2
Result of fieldwork of the resurvey: refusal, deceased, undeliverable, no response	4,069	81.5
Total	4,992	100

Sources: SOEP Resurvey 2006; authors' calculations.

3. Modelling Strategy

Given our main goal of examining if there is a benefit of non-respondent conversion in a longitudinal study like the SOEP, we include county-specific spatially correlated random

effects to account for spatial heterogeneity that is not explained by other covariates (Fahrmeier and Lang 2001). For our empirical analysis, we use geoadditive regression models which extend generalized linear and additive regression in a semi-parametric fashion to simultaneously incorporate linear and nonlinear nonparametric effects of usual covariates and spatial effects into a geoadditive predictor (Fahrmeir, Sagerer, and Sussmann 2007). The geographic information is given in spatially aggregated form, which means that we only know the living area or *Kreis*, which we refer to in the following as "district," of the SOEP respondent.

As an appropriate modelling tool, Fahrmeier and Lang (2001) as well as Fahrmeier, Kneib, and Lang (2004) suggest geoadditive models with spatial components called Gaussian Markov Random Fields (GMRFs). This type of geoadditive model is implemented in the public domain software program BayesX (Brezger, Kneib, and Lang 2005) and is used for our analysis. We follow a three-step procedure:

1. In the first step, we explore the characteristics of the reinterviewed non-respondents and analyze which ones decided to participate again after several years, at least in a shortened version of the SOEP questionnaire.

2. In the second step, we examine the attrition process due to non-response (classified as dropouts) from wave to wave in sample F and estimate several cross-section models as well as longitudinal models for the attrition process.

3. In the last step, we explore the effect of conversions on sample composition and estimate a regression model on life satisfaction with and without the dropout conversions.

3.1 Semiparametric geoadditive models

For our analysis, we use geoadditive models. A detailed description of these regression models can be found in Fahrmeier et al. (2007) or Fahrmeier et al. (2009). First, we assume regression data of the following form

$$(y_i, x_i, z_i, s_i) \quad i = 1, \ldots, n,$$

where i is a general observation index of the respondent, y is a binary response variable[3] (e.g. 1: final dropout, 0: re-interview), x is a vector of binary (e.g. sex) or continuous covariates (e.g. life satisfaction) with linear effects, z is a vector of continuous covariates with nonlinear effects (e.g. birth or size of household), and s is a discrete location variable for the geographic district from which the observation derives, and $s_i \in \{1, \ldots, S\}$ denotes the district (*Kreis*) where the respondent i lives.

---

[3] In the last step of our analysis, the response variable y is a continuous variable, and we use a linear regression model.

In a generalized linear model (GLMs), the conditional expectation $\mu_i = E(y_i \mid x_i)$ of the response is linked to a linear combination $\eta_i = x_i' \beta$ of the covariates x_i via a response function h, i.e.

$$\mu_i = h(\eta_i)$$

Geoadditive regression models extend GLMs by generalizing the linear predictor to a geoadditive predictor.

$$\eta_i = x_i' \beta + f_1(z_{i1}) + \ldots + f_p(z_{ip}) + f_{spat}(s_i)$$

The functions $f_1(z_1), \ldots, f_p(z_p)$ are nonlinear effects of continuous covariates $z_1, \ldots, z_p$. The functions will be approximated through penalized regression splines (P-splines), which were introduced by Eilers and Marx (1996).[4] The function $f_{spat}(s)$ represents the geographic effect of the location variable s. As shown in Fahrmeier et al. (2007), the function can be split up into the sum

$$f_{spat}(s) = f_{str}(s) + f_{unstr}(s)$$

of structured (spatially correlated) and unstructured (uncorrelated) effects. The geographic effect can be seen as a substitute of several underlying unobserved influential factors, which may have structured (correlated) local (district_spatial) or unstructured (uncorrelated) global (district_random) spatial effects.

In most of our applications, we have a binary response variable y_i

$$y_i = \begin{cases} 1, \text{ if occasion (e.g. participation or drop-out) is present} \\ 0, \text{ otherwise} \end{cases}$$

where

$$y_i = \pi_i + \varepsilon_i$$

π_i is the probability of the presence of an occasion of interest and ε_i an individual error term. In the logistic case we will get

$$\pi_i = P(y_i = 1) = \frac{\exp(\eta_i)}{1 + \exp(\eta_i)}$$

where:

$$\eta_i = \beta_0 + \beta_1 x_{1i} + \ldots + \beta_p x_{pi} + f_1(z_{1i}) + \ldots + f_k(z_{ki}) + f_{str}(s) + f_{unstr}(s)$$

---

[4] An application of P-splines in the context of nonresponse analyzes can be found in Seiler (2010).

We use two different estimations procedures for our analysis. In the first step, we intend to estimate the participation in the resurvey of non-respondents and their willingness to participate in the SOEP again. In the second step, we estimate cross-sectional logit models for the attrition process due to non-response in sample F. In both cases, we use a generalized linear mixed-model (GLMM) approach and a restricted maximum likelihood (REML) to estimate the regression parameters (see Brezger et al. 2005).

In addition, we expand our attrition analysis to more complicated panel models and estimate models for the first four waves of sample F. The geoadditive predictor of these models takes the following form:

$$\eta_{it} = \beta_0 + \beta_1 x_{1i} + \ldots + \beta_p x_{pi} + f_1(z_{1i}) + \ldots + f_k(z_{ki}) + f_{str}(s) + f_{unstr}(s) + v_i$$

where v_i is an individual-specific random effect. In the last step, we estimate a linear regression model on life satisfaction with and without the conversions and explore the effect of the inclusion of former non-respondents in the sample. For the panel model as well as the linear regression model, we use Markov Chain Monte Carlo (MCMC) simulation techniques (see Brezger et al. 2005) to estimate the regression parameters.

3.2 Modeling the probability of participation in the conversion study and the willingness to participate again in the SOEP

We start with the data from the conversion study and estimate two logit models, one to explore participation in the resurvey and the other to find an empirical explanation for the respondents' willingness to participate again in the SOEP.

For the analyses of participation in the resurvey, we use the following covariates:

- Variables for the respondents:
  - "*year of birth*" indicates the age of the respondents (in years).
  - "sex" is the gender of the respondents (1 – men).
  - "*Migration*" indicates respondents' migration background (1 – migration background).
  - "*life satisfaction*" is measured on an 11-point scale and indicates respondents' current satisfaction with life.
  - "*bad health*" indicates the health status of the respondents (1 – bad health).
  - "*family status*" is a categorical variable which is split into the categories: married (living together), married (living separately), single, separated, and widowed. These categories are dummy variables.
- Variables for the households:
  - "*income*" equivalent household income of the respondents, measured with the quintile in the last year of participation in the survey
  - "*number of persons in HH*" indicates the size of the household in the last year of participation.
- Variables for regions:
  - "*city*" indicates if the respondent lives in a big city (more than 500,000 inhabitants). This information is stored in a dummy variable.

      ○ "*district*" indicates the district (Kreis) in Germany. This location variable is used to explore the geographic effect, which is split into a structured and an unstructured effect.

The analysis of the willingness to participate again in the SOEP is based only on the participants in the resurvey (N = 907)[5]. From these persons, overall 715 (approximately 79%) agreed to participate again, at least in a shortened form of the SOEP. In the resurvey, the respondents were asked about their life satisfaction (measured on a 0-11 scale), health status (dummy variable) and family status. Table 7 shows the estimates of the parametric terms of the two logit models, the nonparametric terms are shown in Figures 4 and 5.

Table 7: Estimates of the parametric terms of the semiparametric
logit models for "participation" and "conversion"

	Participation in the resurvey vs. nonparticipation		Conversion: Willingness to participate in the SOEP again vs. final dropout	
	β	p-value	β	p-value
Fixed Effects				
Const.	-1.758***	0.00	0.834**	0.04
Sex (1 – male)	0.080	0.30	-0.039	0.82
Migration	-0.652***	0.00	-0.666***	0.01
Life satisfaction	-	-	0.085**	0.05
Bad health	-	-	0.049	0.84
Married (liv. togeth. ref.)			-	-
Married (liv. sep.)			-0.174	0.78
Single			-0.684**	0.02
Separated			0.899*	0.09
Widowed			-0.554*	0.09
City	0.269	0.25	0.355	0.23
Random				
Sigma²_e	$\pi^2/3$		$\pi^2/3$	
District_spatial	0.013		0.000	
District_random	0.464		0.001	
-2LL	4,052.2		896.2	
Number of districts	411		411	
N	4,978		907	

Sources: SOEP Resurvey 2006; authors' calculations.

First we discuss the estimates for participation in the conversion study. The parametric estimates correspond with the results of the descriptive analysis above. They suggest that

---

[5] Due to missing values the sample size was reduced from N=927 to N=907.

migration background has a negative effect on the probability to participate in the resurvey. The effects of the last stored equivalent household income, the year of birth, and the size of the household are modelled by cubic P-splines. The estimates of the functions f(HH-Income), f(size of household) and f(year of birth), together with 80% and 95% pointwise confidence intervals, can be found in Figure 4.

We can recognize a significant positive relationship between participation and the last measurement of equivalent household income. Respondents of retirement age (born before 1940) have the highest rate of participation in the resurvey (which took place in 2006); older and younger respondents have a significantly lower rate. Furthermore, respondents who live in a three-person household have the highest participation rate, whereas the rates for smaller and larger households are significantly lower.

The parametric estimates for the willingness of non-respondents to participate again in the SOEP are shown in the last two columns of Table 7. The probability depends on the migration background, life satisfaction, family status, as well as the equivalent household income and the age of the respondent. Willingness shows a positive nonlinear relationship with income and positive linear relationship with year of birth. It seems that high earners and younger respondents are more willing to participate than others. Having a migration background and being single also significantly reduce the willingness to participate.

Figure 4: Nonparametric functions of the metric covariates of the semiparametric logit models for "participation" and "conversion"

	Participation in the resurvey vs. nonparticipation	Conversion: Willingness to participate in the SOEP again vs. final dropout
Effect of the equivalent HH-income		
Effect of the year of birth		
Effect of the size of household		
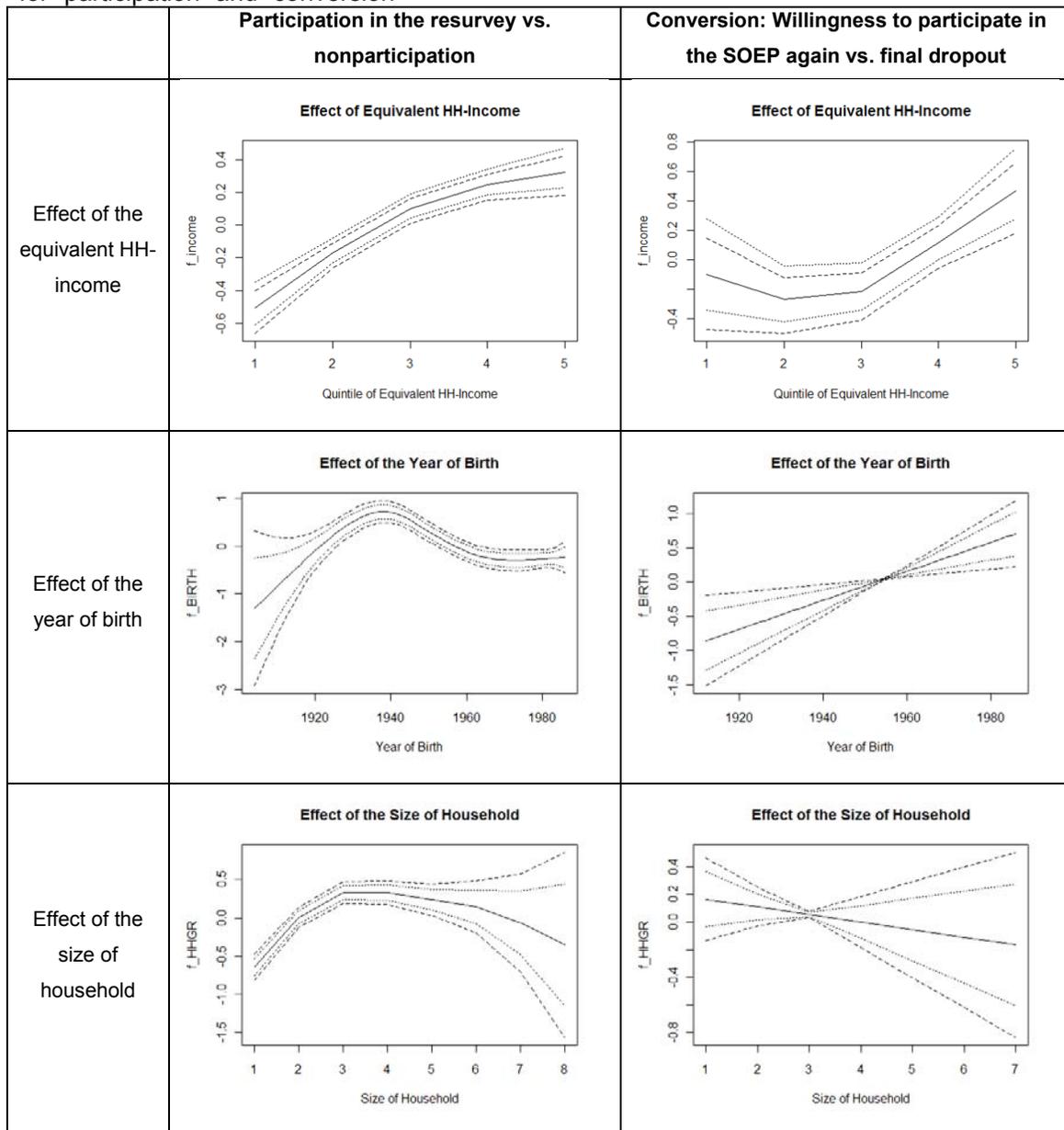
Confidence bands: dashed line: 95% and dotted line 80%

Sources: SOEP Resurvey 2006; authors' calculations.

In Figure 5 the estimated smooth geographic effect is shown. Areas in dark grey indicate lower and areas in light grey higher participation rates than expected. These results give evidence of reverse area effects for respondents' participation in the resurvey and their willingness to participate again in the SOEP. Participation appears higher in the south and lower in the north, whereas the future participation is higher in the north and lower in the south. Nevertheless, these effects are not significant and, particularly in the case of the conversion, very small.

Figure 5: Structured spatial effect of the semiparametric logit models for "participation" and "conversion"

	Participation in the resurvey vs. nonparticipation	Conversion: Willingness to participate in the SOEP again vs. final dropout
Structured spatial effect		
Significant Areas (80% confidence interv.)		
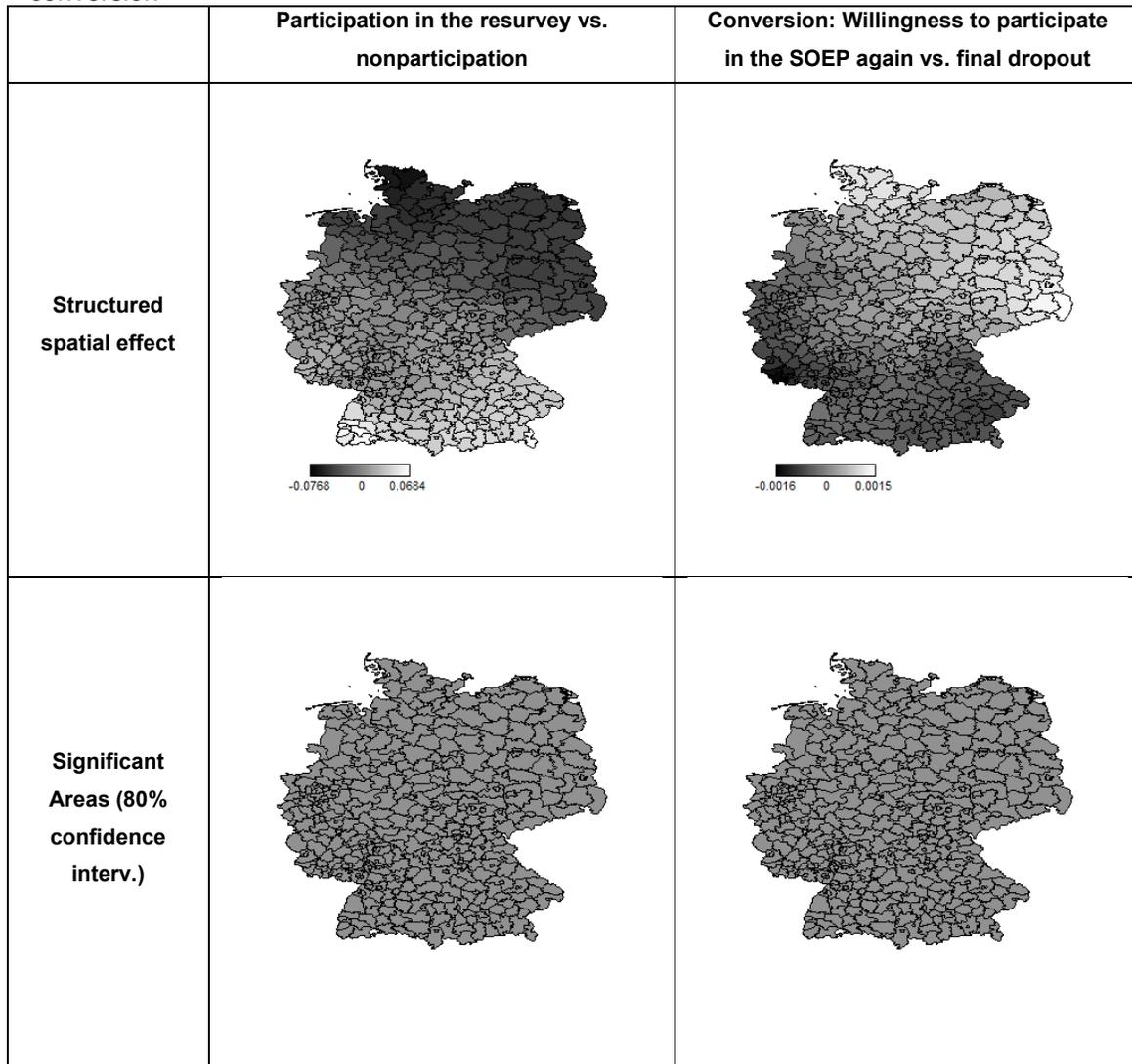
Sources: SOEP Resurvey 2006; authors' calculations.

Now, in the second step, we examine the attrition process due to non-response from wave to wave in sample F and estimate several cross-section models as well as longitudinal models for the attrition process.

4. Panel Attrition Due to Non-response in Sample F

The development of sample sizes and the attrition process due to non-response is well documented for all SOEP samples by the DIW Berlin (Kroh and Spieß 2008). For our analysis, we concentrate on the panel attrition due to non-response in sample F. Table 8 shows the development for the first five waves. We select only households that participated in wave 1 of sample F.

Table 8: Panel attrition process due to fieldwork related factors (without moving abroad and death) on household level

Wave	Realized	Not possible, not able	Temporary "dropouts"	"Dropouts"	Temporarily unreachable	Unreachable	Total
1	5,976	0	0	0	0	0	5,976
2	4,822	56	335	790	39	18	6,060
3	4,488	13	206	611	10	13	5,341
4	4,282	15	214	299	6	4	4,820
5	4,090	18	167	280	4	7	4,566

Sources: SOEP (v26), Sample F, 2000-2004; household level, authors' calculations.

We see that the majority of the attrition process is caused by non-response, which is classified as "dropping out". The relationship between non-response and realized interviews started in wave 2 at 16.4% and declined to 6.8% in wave 5.

4.1 Modeling the probability of non-response vs. re-interviewing in Sample F

Our goal is to give an empirical explanation for non-response. We use the first four waves and define a response variable with 1 for "non-response" and 0 for a successfully conducted interview. In our analysis, we focus solely on non-respondents, hence temporary dropouts such as one-time refusals and respondents stating an interview was "not possible at the moment" as well as those who could not be reached for an interview are left out. Based on the household and interview characteristics measured at time t, we explore the probability of non-response vs. participation re-contacted at time t + 1. Kroh and Spieß (2008) investigated non-response at the household level across all samples. As regressors, they used individual characteristics such as gender and age of the household head as well as household size and interview characteristics such as a change of interviewer or mode. Similar variables are used here. Again, the regressors can be divided into three groups:

- Variables for the respondents (household head):
  - "*year of birth*" indicates the age of the respondents
  - "sex" is the gender of the respondents (1 – men)
  - "*Migration*" indicates respondents' migration background (1 – migration background)

- o "*Life satisfaction*" is measured on an 11-point scale and treated as a continuous variable
- Variables for the households:
  - o "*income*" equivalent household income of the respondents, measured with the quintile in the last year of participation in the survey
  - o "*number of persons in HH*" indicates the size of the household in the last year of participation.
- Variables for the interview situations:
  - o "*change in interviewer*" indicates if the interviewer changed (0 if mode is mail or no change in interviewer)
  - o "*mail*" indicates if the interview mode is mail
- Variables for the geographic locations:
  - o "district" indicates geographic district (statistical regions called "Kreise" ("counties"))

4.2   Estimates for the cross-sectional attrition model

Table 9 shows the variables that are parametrically estimated including all dummy variables such as sex, migration status, mail interview, and change of interviewer as well as the continuous variable "life satisfaction." From this, we see that a mail interview (wave 3 and 4) and a change of interviewer (wave 3) significantly increase the probability of being classified as "dropout" of SOEP sample F[6]. Furthermore, an increase in the reported life satisfaction of the respondent will lower the probability of non-participation (wave 2). These findings correspond with household-level findings on "dropouts" reported in Kroh and Spieß (2008).

The effects of the remaining continuous covariates are modelled by cubic P-splines. The estimates of the functions f(HH-Income), f(Size of Household) and f(year of birth), together with 80% and 95% pointwise confidence intervals, can be found in Figure 6.

---

[6] The effect of "change in interviewer" refers to the sum of the categories "mail" or "no change in interviewer".

Table 9: Estimates of the parametric terms of the logit models for the dropout classification in Sample F

	F Dropout Classification vs. Reinterviewing					
	Wave 2		Wave 3		Wave 4	
	Coeff.	p - value	Coeff.	p – value	Coeff.	p – value
Fixed Effects						
Const.	-1.595***	0.000	-2.421***	0.000	-2.809***	0.000
Sex (1 – male)	0.044	0.615	-0.195	0.170	-0.138	0.764
Migration	0.075	0.414	0.161	0.226	0.196	0.314
Life satisfaction	-0.042*	0.055	-0.043	0.233	-0.073	0.169
Mail			1.112***	0.000	1.237***	0.000
Change of interviewer			0.393**	0.029	0.336	0.360
Random						
Sigma²_ε	$\pi^2/3$		$\pi^2/3$		$\pi^2/3$	
district_spatial	0.095		0.321		0.167	
district_random	0.136		0.091		0.357	
-2LL	4,232.1		3,469.8		1,013.2	
N	5,480		4,532		4,215	

Sources: SOEP (v26), Sample F, waves 1 to 4, household level; authors' calculations.

The variable equivalent household income has a strong effect on the probability of being classified as dropout between wave 1 and wave 2. In the first quintile, we estimate an effect *f(HH-Income = 1)* = 0.12. The odds of dropping out therefore increase by the factor exp(0.12) = 1.127. In the last quintile, we estimate only an effect of *f(HH-Income = 5)* = -0.12. Hence, the odds decrease by the factor exp(-0.12) = 0.887 if all other variables in the model remain constant. With increasing quintiles of the equivalent household income, the chance of being classified as a dropout decreases; the narrow confidence intervals show the high accuracy of the estimates. However, this finding is not confirmed in subsequent waves. From wave 2 to 3, we find the lowest chance of being classified as a dropout in quintile 3 and higher chances in quintiles 1 and 5. From wave 3 to 4, the chance increases slightly with quintiles. Nevertheless, we get broad confidence intervals in both cases, which suggests that these findings are not significant. Hence, we find only consistent and significant results from wave 1 to 2 for the equivalent household income.

Similar results can be found for the effect of household size. From wave 1 to 2, we can see that singles have the highest chance of being classified as dropouts and that the chance

declines with the size of the household. From wave 2 to 3, no significant effect can be found. From wave 3 to 4, the chance is slightly lower for singles and smaller households.

Furthermore, we see a strong, nonlinear age effect from wave 1 to 2 and from wave 2 to 3. In both cases, the chance of being classified as a dropout increases from age 60 upwards. As there are only few observations beyond the age of 80 (year of birth < 1920), confidence intervals become wider here. Nevertheless, the effects remain significant. In wave 3 to 4, no significant age effect can be observed.

The smooth geographic effect estimated in the models stands out in Figure 7, with estimates between -0.341 and 0.416 from wave 1 to 2, -0.527 and 1.195 from wave 2 to 3 and -0.853 and 2.054 from wave 3 to 4. Areas in light grey indicate lower, and areas in dark grey higher dropout classification rates than expected. Underneath these maps, the areas that are significant on the 80% confidence interval level are shown. We see here that from wave 1 to 2, the dropout classification rates are significantly lower in the area around Berlin and higher than expected in southern Germany.

Figure 6:   Nonparametric functions of the metric covariates of logit models for the dropout classifications in sample F

Wave 1 to Wave 2	Wave 2 to Wave 3	Wave 3 to Wave 4
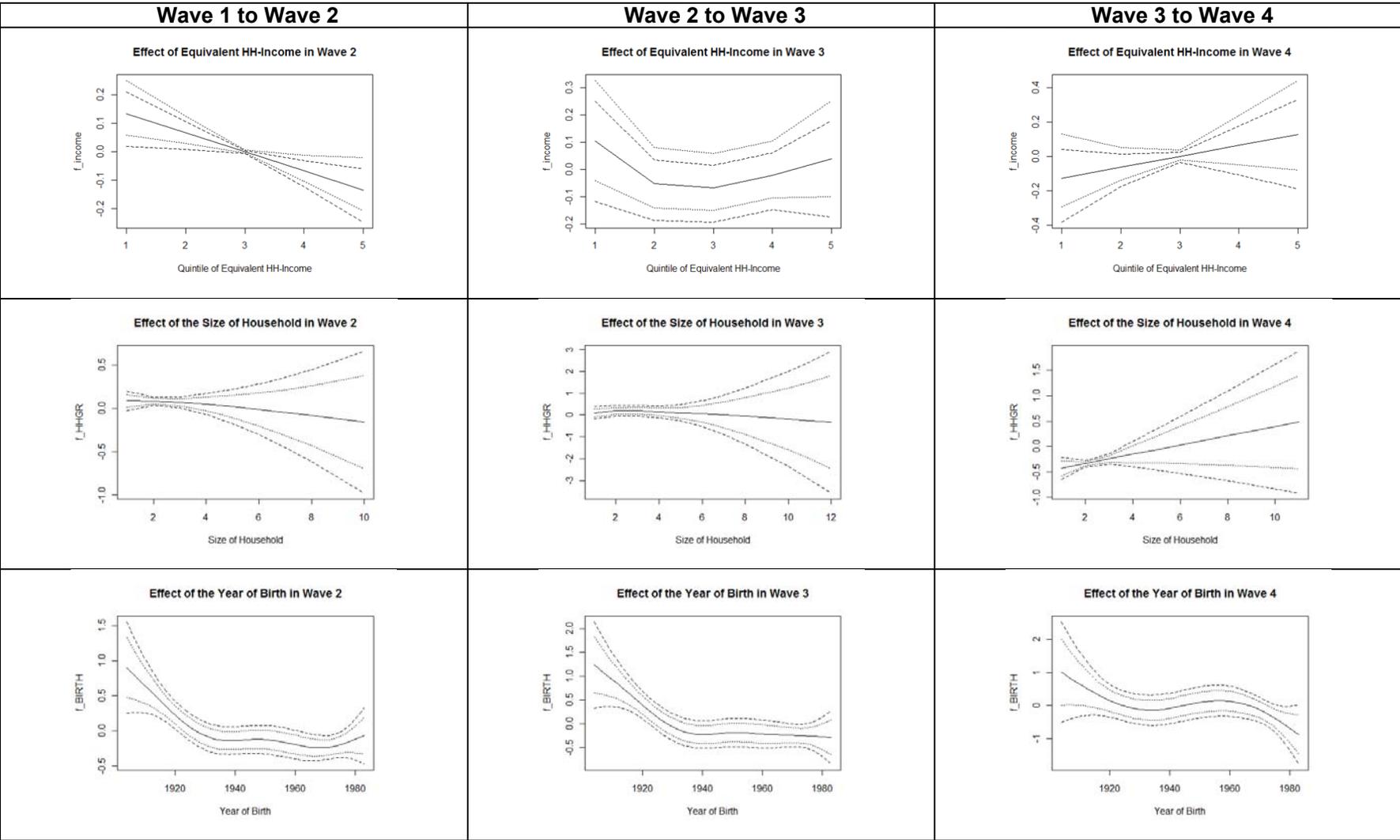
Sources: SOEP (v26), Sample F, waves 1 to 4; authors' calculations.

Figure 7: Structured spatial effect of the logit models for the final dropout classifications in Sample F

Wave 1 to Wave 2	Wave 2 to Wave 3	Wave 3 to Wave 4
Significant Areas (80% confidence interv.)	Significant Areas (80% confidence interv.)	Significant Areas (80% confidence interv.)
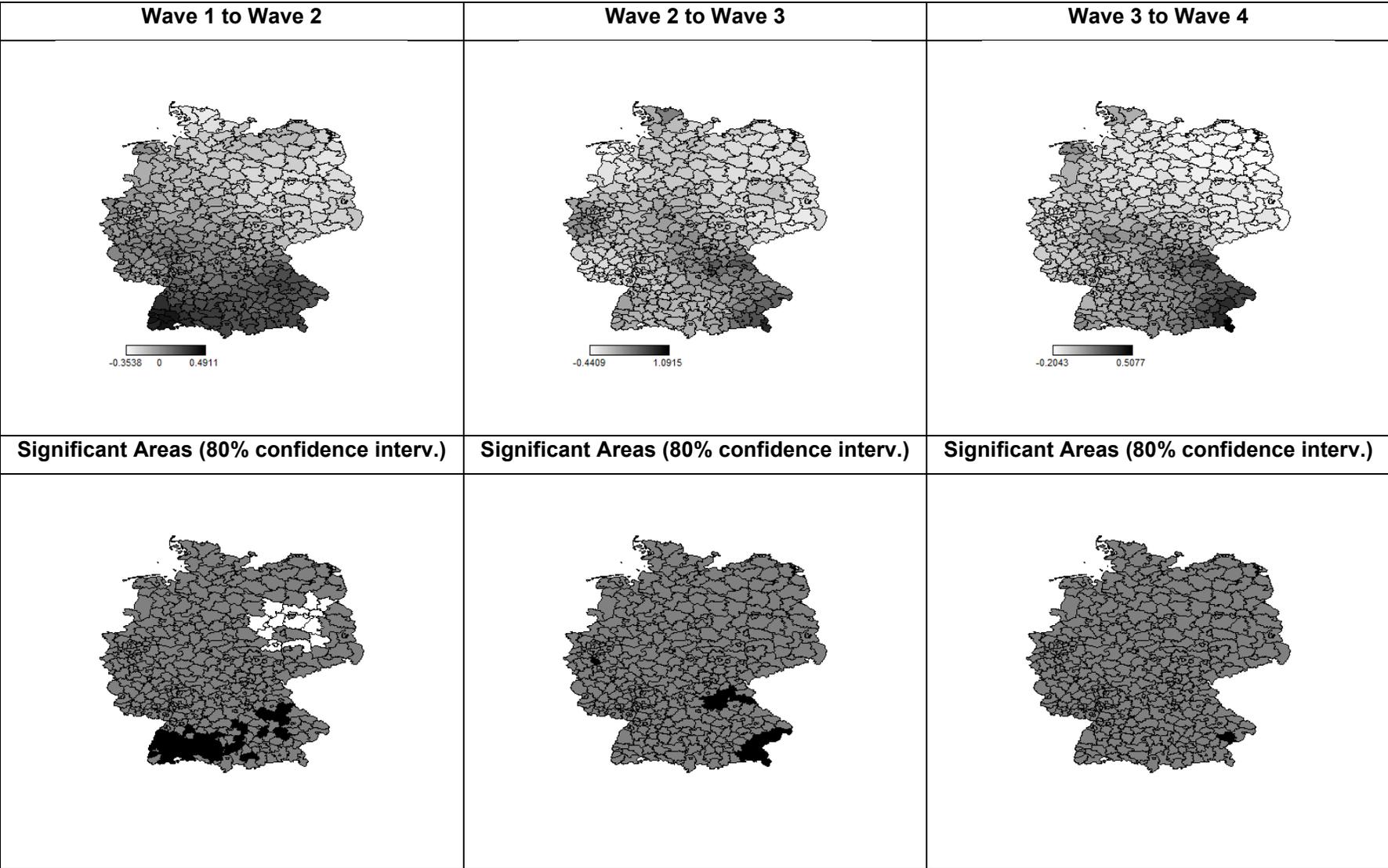
Sources: SOEP (v26), Sample F, waves 1 to 4; authors' calculations.

4.3 Panel attrition models with and without inactive respondents who are willing to participate again

Now we intend to explore whether there is a change in the explanation of the panel attrition process due to dropout classifications in Sample F if formerly inactive respondents who are willing to participate again are treated as respondents. We pool the data from the first four waves of sample F and estimate several longitudinal logit models with random effects.

In the first model, we explore the probability of a respondent being classified as a dropout vs. becoming a participant again in the next wave ($t + 1$) in sample F. In the second and third models, we examine if there are significant predictors that could identify the converted participants and reveal differences between converts, participants, and final dropouts in the next wave. We therefore estimate the probability of being a converted non-respondent vs. regular respondent and the probability of being a converted non-respondent vs. a final dropout in the next wave.

In the fourth model, we switch the non-respondents who declared in the resurvey their willingness to participate again in the SOEP in our response variable from the group of non-respondents ($Y = 1$) to the participants ($Y = 0$). Hence, we estimate the probability of adjusted non-response vs. participation in the next wave and explore if the conversion of inactive survey members has an effect. We change some parameters of the cross-sectional models slightly: For the longitudinal models, we assume that life satisfaction has a nonlinear effect on the response variables and that the predictor "size of the household" can be reduced to an indicator variable "SINGLE" with 1 – one-person household, 0 – otherwise.

4.4 Estimates of panel attrition models for the first four waves

The estimates for the parametric (Table 10, columns 2 and 3) and nonparametric terms (Figure 8a, row 1) of the first logit model in Table 10 confirm the results of the cross-sectional models in the previous section. The probability of a dropout classification increases with respondents' migration status (β = 0.306) and mail interview mode (β = 1.002) and decreases slightly with life satisfaction. Low household income (first quintile) and being over 80 years of age increase the dropout classification rates substantially. Again, we find a significant structured geographic effect (Figure 9, row 1). The dropout classification rates are significantly lower (light grey) than expected in the Berlin area and higher (dark grey) in the south of Germany. In the cross-sectional models, we see that this effect occurs mainly from wave 1 to 2.

The converted non-respondents have different characteristics than both participants and the final dropouts. The parametric estimates (Table 10, columns 4 and 5) show that previously inactive respondents who are willing to participate again in the SOEP are mainly

Germans living in multi-person households, and the nonparametric terms (Figure 8a, row 2 and Figure 8b, row 1) suggest that "converts" have higher life satisfaction and higher household incomes than participants and final dropouts. In both cases, the function for year of birth f(BIRTH) reaches its maximum at 1940, implying that the probability of being a converted survey member is highest for persons aged 66. Furthermore, Figure 9 shows that conversion rates are higher than expected in the south of Germany in the area around Freiburg but that the effect is not significant. Overall, these results allow us to assume that the converted respondents differ both from final dropouts and from participants.

In the last logit model, we switched the inactive respondents who declared their willingness to participate again to the group of active respondents (Y = 0). Here, we treat the converted respondents as if they were not dropouts because they are again active respondents in the SOEP. Hence, we explore the probability for the adjusted dropouts vs. participants. It is evident in this model that the positive migration effect and the nonlinear negative life satisfaction effect are slightly stronger than in the "unadjusted" case because converted respondents are mainly non-migrants, have a higher level of life satisfaction, and are treated here as participants. The negative relationship between equivalent household income and dropping out is significant and now almost linear.
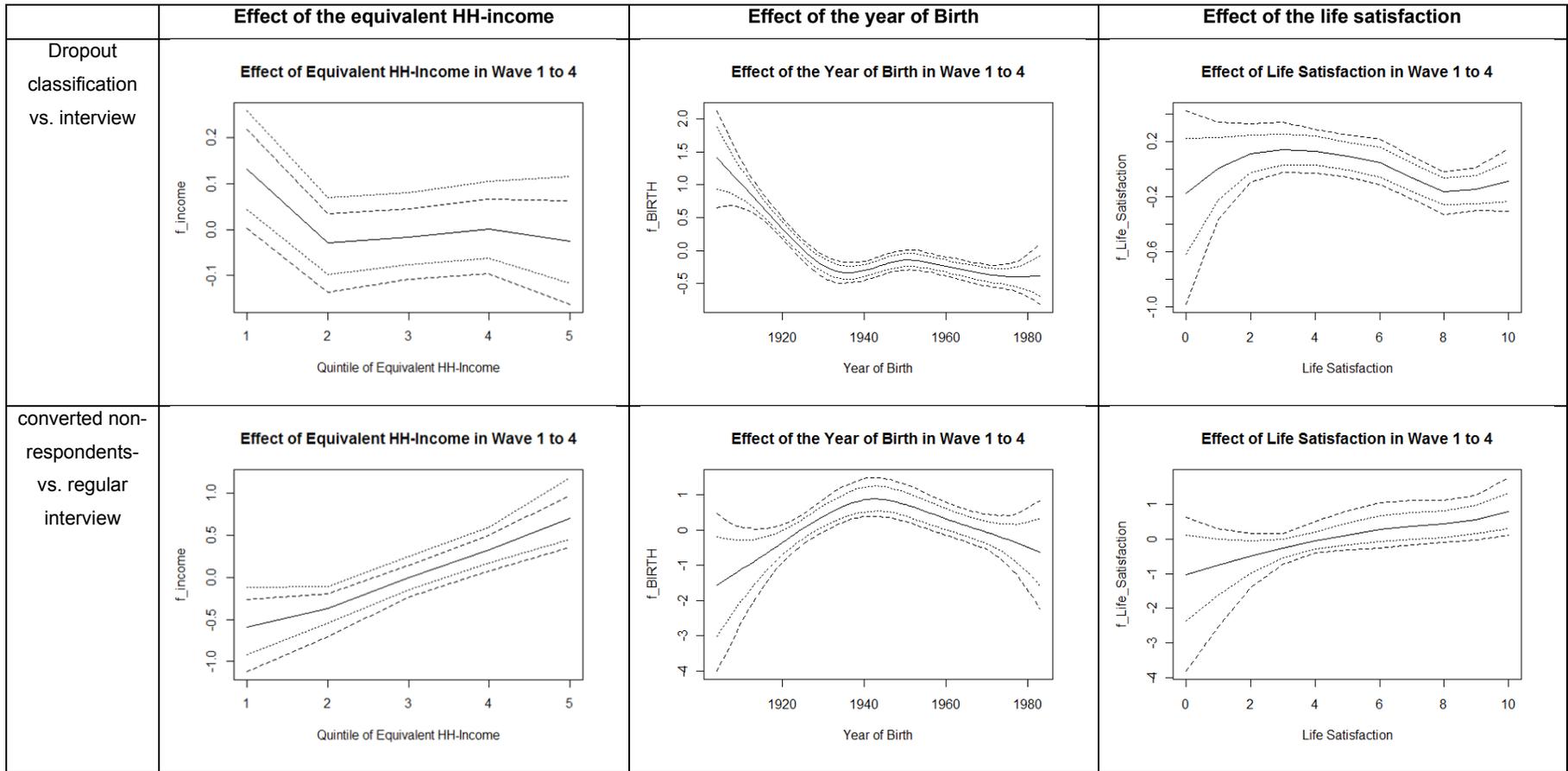
Table 10: Estimates of the parametric terms of the semiparametric panel logit models (household level)

	Dropout Classification vs. interview		Converted non-respondents vs. regular interview		Converted non-respondents vs. final dropouts		Adjusted dropout Classifications vs. regular interview	
	(*t* + *1*)		(*t* + *1*)		(*t* + *1*)		(*t* + *1*)	
	β	Std.dev.	β	Std.dev.	β	Std.dev.	β	Std.dev.
Fixed Effects								
Const.	-1.581*	0.14	-5.907*	1.56	-2.675*	0.45	-2.120*	0.40
Sex (1 – male)	-0.041	0.07	-0.029	0.24	-0.107	0.22	-0.066	0.09
Migration	0.306*	0.09	-0.173	0.34	-0.796*	0.32	0.440*	0.14
Single	-0.098	0.08	-1.235*	0.34	-0.983*	0.30	0.017	0.09
Mail	1.002*	0.17	1.862*	0.51	0.498	0.47	0.966*	0.23
change interviewer	0.206	0.14	0.223	0.44	-0.054	0.49	0.217	0.16
w1 (ref)								
w2	-0.887*	0.07	-0.921*	0.25	-0.170	0.25	-0.751*	0.14
w3	-1.724*	0.10	-1.645*	0.34	-0.208	0.33	-1.575*	0.17
w4	-1.770*	0.10	-1.851*	0.39	-0.321	0.38	-1.549*	0.20
Random								
Sigma²_e	π²/3		π²/3		π²/3		π²/3	
Sigma²_v	0.037	0.12	3.130	4.42	-	-	1.601	1.39
district_spatial	0.222	0.11	0.498	0.54	0.084	0.17	0.319	0.22
district_random	0.075	0.05	0.311	0.32	0.176	0.21	0.064	0.06
-2LL	8,459.4		1,302.1		842.9		6,745.0	
DIC	8,299.1		1,057.8		804.1		5,738.7	
N	18,215		17,072		1,301		18,215	

Note:*0 ∉ $CI[95\%]$

Sources: SOEP (v26), Sample F, waves 1 to 4, household and individual questionnaire; authors' calculations.

Figure 8a: Nonparametric functions of the metric covariates of the semiparametric panel models

	Effect of the equivalent HH-income	Effect of the year of Birth	Effect of the life satisfaction
Dropout classification vs. interview			
converted non-respondents- vs. regular interview			

Confidence bands: dashed line: 95% and dotted line 80%.

Sources: SOEP (v26), Sample F, waves 1 to 4; authors' calculations.

Figure 8b: Nonparametric functions of the metric covariates of the semiparametric Panel models
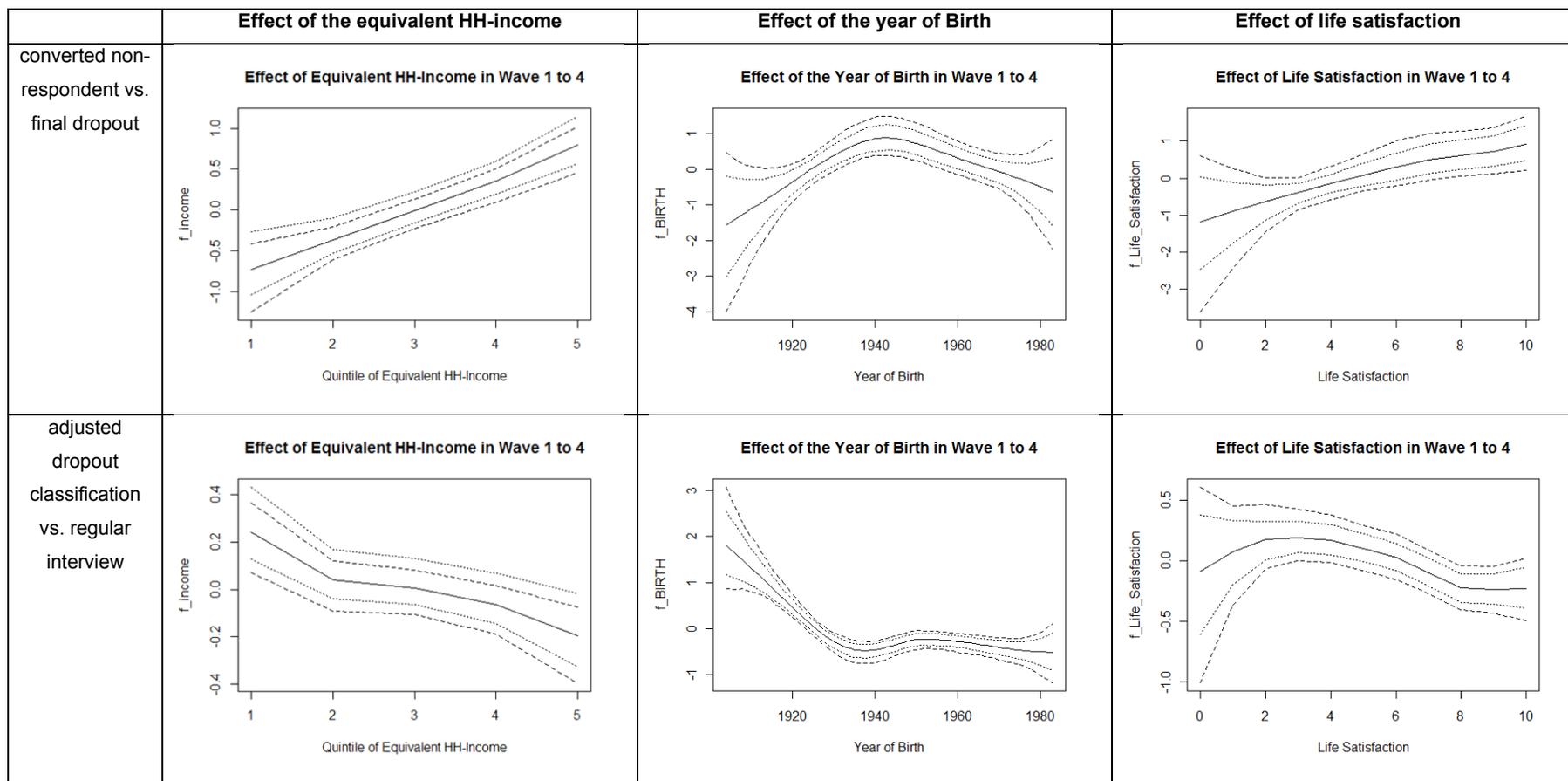
	Effect of the equivalent HH-income	Effect of the year of Birth	Effect of life satisfaction
converted non-respondent vs. final dropout			
adjusted dropout classification vs. regular interview			

Confidence bands: dashed line: 95% and dotted line 80%.

Sources: SOEP (v26), Sample F, waves 1 to 4; authors' calculations.

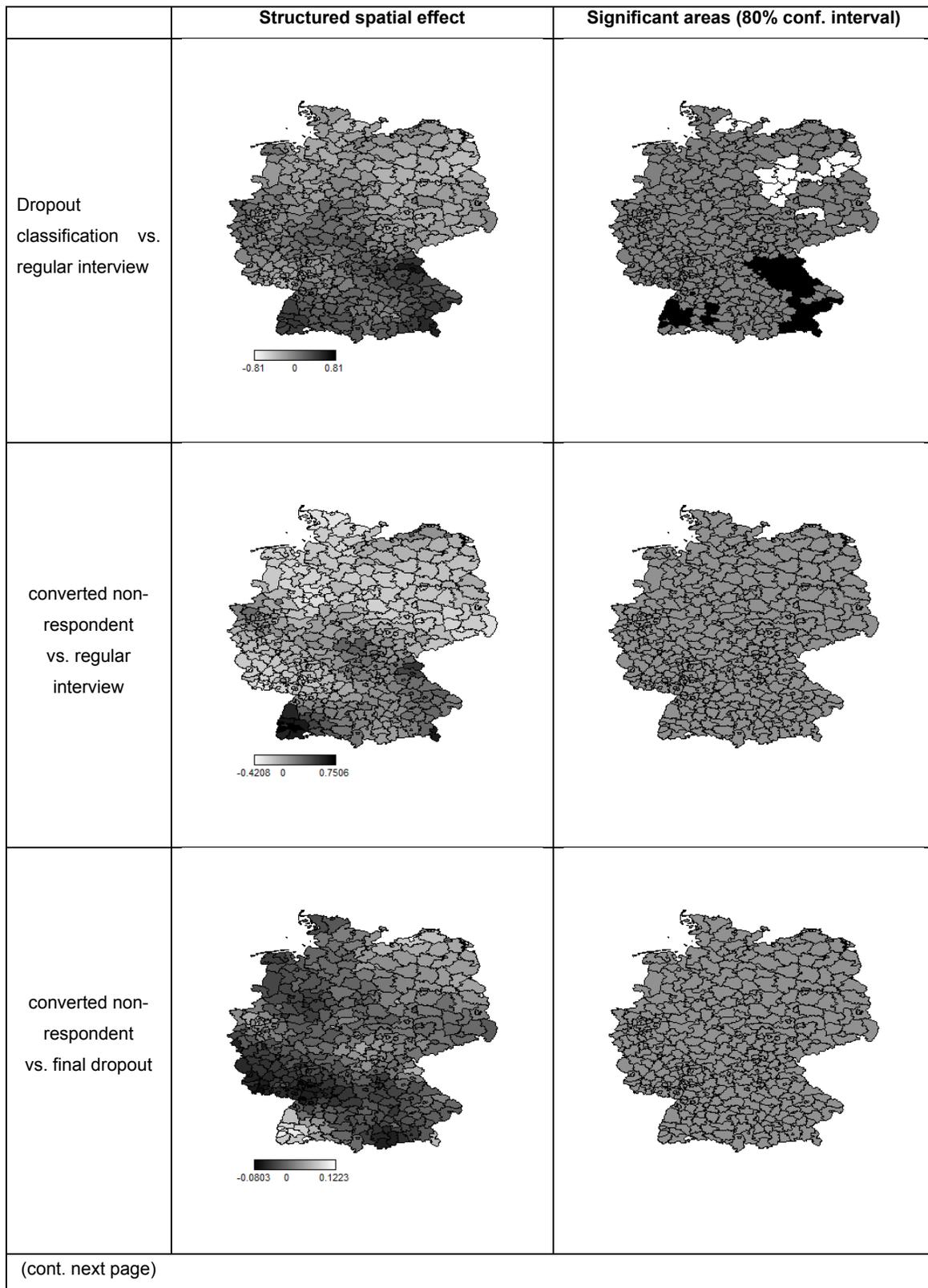Figure 9: Structured spatial effect of the semiparametric panel models for sample F

	Structured spatial effect	Significant areas (80% conf. interval)
Dropout classification vs. regular interview		
converted non-respondent vs. regular interview		
converted non-respondent vs. final dropout		
(cont. next page)		

(continued)		
adjusted dropout classification vs. regular interview		
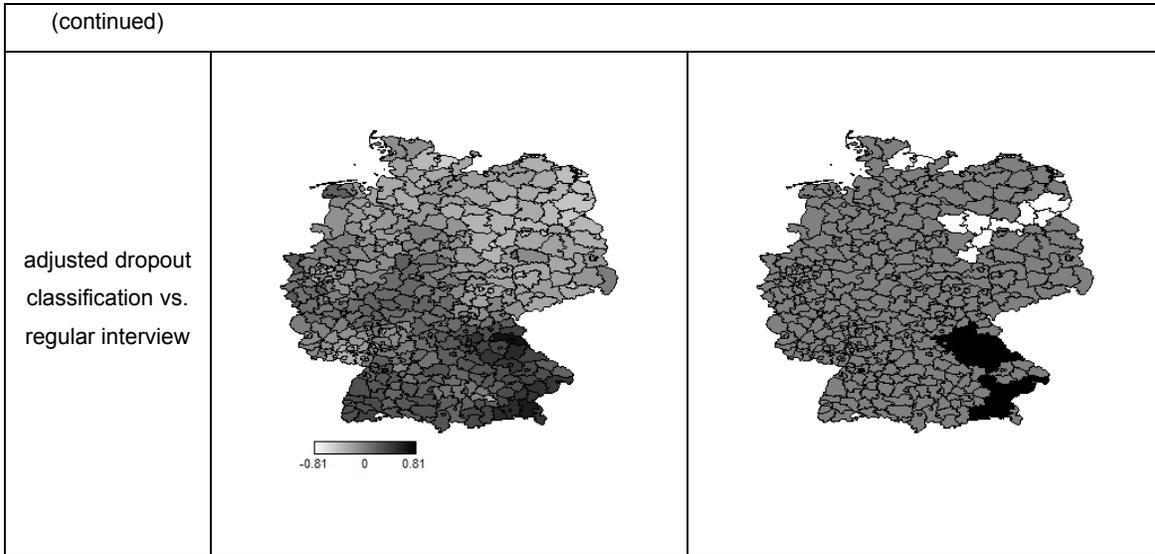
Sources: SOEP (v26), Sample F, waves 1 to 4; authors' calculations.

5. The effect of including converted non-respondents in the SOEP on regression estimates for life satisfaction

In recent decades, life satisfaction has become one of the most widely used indicators of well-being. Research on this topic has uncovered internal and external determinants of life satisfaction (Schimmack and Lucas 2007). Whereas the internal determinants are rooted in the human genetic makeup and biological dispositions, the external determinants refer to living conditions such as income, employment, and social relationships. It can be shown that some external determinants such as divorce, unemployment, and disability produce lasting changes in life satisfaction (Schimmack and Lucas 2007, Lucas et al. 2004). The main SOEP, as well as the resurvey of non-participants, provides information about some of these important indicators. Our goal is to explore not the determinants of life satisfaction but the effects of including converted non-respondents into the normal SOEP in 2006. To this end, we examine the estimates of linear regression models for life satisfaction with and without the inactive respondents in subsample F in wave 6 (year 2006). If there are minor changes or no changes at all, we can conclude that the effort to convert final dropouts is not worthwhile, at least for this analysis.

For our analysis we again use semiparametric regression models. Life satisfaction is measured on an 11-point scale and can be treated as a continuous variable. The regressors are demographic variables like "year of birth," "sex," and "migration background" as well as "family status" and "subjective health," which is measured on a 5-point scale. Again, the

effects of year of birth as well as subjective health are estimated by nonparametric functions. All other variables are categorical variables.

The parametric estimates are shown in Table 11. The second and third columns contain the estimates for the sample without the converted non-respondents. We can see here that life satisfaction decreases significantly if the respondent is unemployed, but that gender and migration background have no effect. For family status, we use as reference the category "single" and we can see that spouses who live together have the highest, and married respondents who are separated the lowest, life satisfaction. Separated respondents have similar satisfaction values to singles, while widows and widowers have significantly higher values.

In the last two columns, we find the estimates for the sample containing converted non-respondents. Overall, we see only small differences. If we compare the mean of the sample with and without the converted respondents, in both cases we get the same values up to two decimal places (6.879 and 6.872). We can therefore conclude that adding the converted non-respondents who are willing to participate in the study again makes no or only a very minor difference.[7] The differences between female and male respondents are now significant: female respondents have slightly higher satisfaction values than males. The other values are in line with the previous model.

---

[7] This is remarkable because it is well-known that the reported satisfaction with life decreases with the number of participations in the SOEP survey (e. g. Frick et al. 2006).

Table 11: Estimates of the parametric terms of the semiparametric linear regression models for life satisfaction
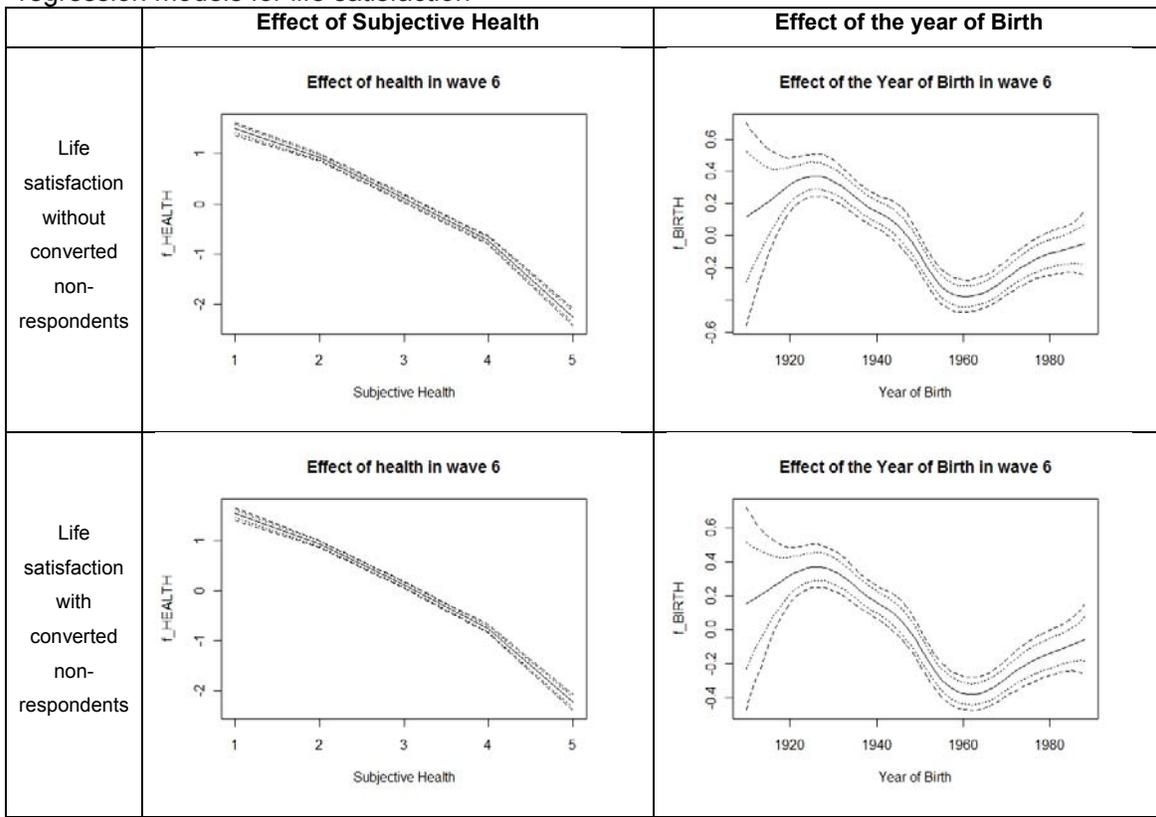
	Life satisfaction without converted non-respondents		Life satisfaction with converted non-respondents	
	B	Std.dev	β	Std.dev
Fixed Effects				
Const.	6.423*	0.09	6.396*	0.09
Sex (1- male)	-0.073	0.04	-0.084*	0.04
Migration	-0.077	0.06	-0.058	0.06
unemployed	-1.025*	0.08	-1.000*	0.07
single (Ref.)				
married (living together)	0.331*	0.06	0.316*	0.07
married (living seperated)	-0.378*	0.15	-0.357*	0.15
Separated	-0.005	0.09	-0.028	0.09
Widowed	0.241*	0.10	0.219*	0.10
Random				
Sigma²_e	$\pi^2/3$		$\pi^2/3$	
district_spatial	0.099	0.05	0.090	0.04
district_random	0.122	0.03	0.118	0.024
-2LL	6836.8		7203.4	
DIC	6634.4		7100.4	
Restricts	411		411	
N	6838		7298	

Note:*$0 \notin CI[95\%]$

Sources: SOEP (v26), Sample F, household and individual questionnaire, wave 6 (year 2006), Resurvey 2006; authors' calculations.

The nonparametric functions (figure 10) show strong significant influences of subjective health (1 – very good, 5 – very bad) and year of birth. The first is negative and almost linear; the last nonlinear. The lowest satisfaction values are predicted for respondents aged 46 (year of birth 1960), and younger or older respondents show higher satisfaction values. Hence, we get a u-shaped function. We can see that the nonparametric functions are quite similar and that there are no differences between estimates with and without converted non-respondents.

Figure 10: Nonparametric functions of the metric covariates of the semiparametric linear regression models for life satisfaction
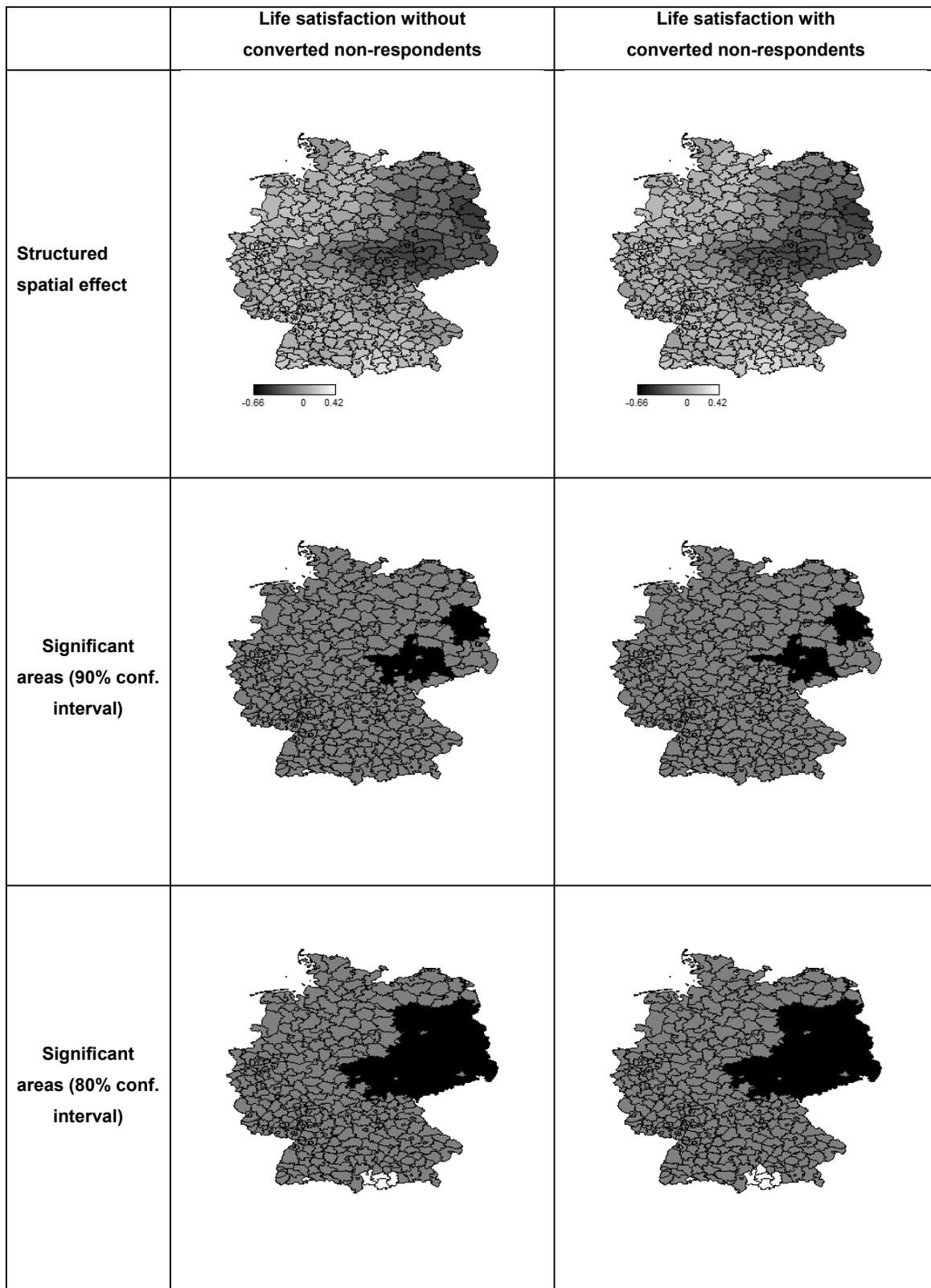
	Effect of Subjective Health	Effect of the year of Birth
Life satisfaction without converted non-respondents	Effect of health in wave 6	Effect of the Year of Birth in wave 6
Life satisfaction with converted non-respondents	Effect of health in wave 6	Effect of the Year of Birth in wave 6

Confidence bands: dashed line: 95% and dotted line 80%.

In Figure 11, we show the structural spatial effect of life satisfaction. In East Germany, the satisfaction values are significantly lower than in the western and southern parts of Germany. This strong effect is more or less independent of the converted respondents. On the 90% confidence interval, we find three small local districts (*Kreise*) which are no longer significant in the model with the converted respondents. On the 80% confidence level, there are no differences at all.

Figure 11: Structured spatial effect of the linear regression models for life satisfaction

	Life satisfaction without converted non-respondents	Life satisfaction with converted non-respondents
Structured spatial effect		
Significant areas (90% conf. interval)		
Significant areas (80% conf. interval)		

Sources: SOEP (v26), Sample F, household and individual questionnaire, wave 6 (year 2006), Resurvey 2006; authors' calculations.

6. Conclusion

SOEP's conversion study was done to convert the survey members who were classified as "dropouts" in the years 2001 to 2004 to again active participants. The fieldwork organization TNS Infratest Sozialforschung had sent all 4,992 non-respondents a short questionnaire by mail in the year 2006. The response rate of this mail survey was about 18%. Our findings show that most of the participants in the conversion study are German and around 60 years of age on average. The participation rate increased with the quintiles of equivalent income. Of these participants, 79% (14% of all recontacted non-respondents) were willing to participate in an active manner in the SOEP again, at least in a shortened version of the questionnaire.

Our research question was to explore whether in an ongoing panel study such a non-respondent conversion matters. We focused our analysis on the attrition process and on regression estimates of life satisfaction.

First we examined whether the inclusion of converted respondents changes the attrition process in a positive way, and we switched the converted respondents from the group classified as dropouts to the group of active respondents.

The cross-sectional as well as the longitudinal analysis of the panel attrition process in SOEP's Sample F showed that mail interviews, a migration background, as well as age have a positive, and income and life satisfaction a negative effect on the probability of becoming a final dropout. Furthermore, the dropout classifications are geographically selective and the share in two southern German states (Bavaria and an area around Freiburg) is significantly higher than expected.

The converted respondents had different characteristics than the regular respondents as well as the final dropouts. Hence, there was hope that inclusion of these respondents would reduce the selection effect due to attrition. However, the converted respondents are mainly German, more often live in multiperson households, have on average a high life satisfaction and have on average a higher equivalent household income than the final dropouts. Therefore, we conclude that we converted the "wrong" sample members. In addition, the conversion took place more often than expected (but not in a significant manner) in southwestern Germany, particularly in an area around Freiburg where a significantly higher share of non-respondents live. Indeed, the good news is that according to our results, conversion reverses this higher share of non-respondents in this area (whereas the significantly higher share of non-respondents in Bavaria still remains). The bad news is that the conversions do not level out the higher share of dropout classifications with a migration

background; instead, inclusion increases the migration bias because the converted non-respondents are mainly Germans.

In our last analysis, we explored the influence of the conversion on regression estimates for life satisfaction. The relevant determinants for this response variable are available in the conversion study as well as in the regular SOEP. Overall we find only small differences in the estimates with and without the converted non-respondents. The parametric as well as the non-parametric estimates are quite similar and in line with our expectations. A small change in the parameters for gender lead to a significant value, and a few districts in East Germany are no longer lower than expected because of the slightly higher satisfaction of the converted non-respondents who are now included. Nevertheless, the mean value for life satisfaction is up the same to two decimal places with and without converted non-respondents.

In sum, we can conclude on the basis of our analyses that the effort to convert non-respondents in the SOEP may slightly reduce the geographic selectivity of the attrition process but that it has no major effect on estimates of regression parameters for life satisfaction.

Finally we conclude that the "SOEP resurvey" shows that special efforts to convert inactive into active survey participants are relatively ineffective in two respects. First, the rate of successful conversions is low (about 15 percent). Second, the composition of the longitudinal file does not improve with respect to individual characteristics. The same groups that have a higher likelihood to be non-respondents in the regular SOEP (low-income respondents and respondents with migration background) show lower conversion rates as well. As a result, the final active file is even more selective than the regular file of SOEP respondents.

Due to the low conversion rate, not even the efficiency of the SOEP sample (the sample size) can be improved significantly by conversion efforts. The resurvey was an interesting and for survey methodological research important experiment, but it does not offer an effective means of improving the SOEP fieldwork.

References

Beullens, K., Billet, J. and Loosveldt, G. (2010): The effect of the elapsed time between the initial refusal and conversion contact on conversion success: evidence from the 2nd round of the European social survey. Quality & Quantity, Vol. 44, 1053-1065.

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

Brezger, A., Kneib, T., and Lang S. (2005): BayesX manuals. Available from http://www.stat.uni-muenchen.de/~bayesx

Burton, J., Laurie, H. and Lynn, P. (2006): The Long-Term Effectiveness of Refusal Conversion Procedures on Longitudinal Surveys. Journal of the Royal Statistical Society Series A (Statistics in Society). Vol. 169.

Calderwood, L., Plewis, I., Ketende, S.C. and Taylor, R. (2010): Experimental testing of refusal conversion strategies in a large-scale longitudinal study. CLS Cohort Studies, Working paper 2010/9. Centre for Longitudinal Studies. London.

Curtin, Richard, Stanley Presser and Singer, E. (2000). The Effects of Response Rate Changes on the Index of Consumer Sentiment. Public Opinion Quarterly 64(4): 413-428.

DeMaio, T. J. (1980). Refusals: Who, Where and Why. *Public Opinion Quarterly* 44: 223-233.

Eilers, P.H.C. and Marx, B.D. (1996): Flexible Smoothing with B-splines and Penalties. *Statistical Science*, 11(2): 89-121.

Fahrmeier, L. and Lang, S. (2001): Bayesian inference for generalized additive mixed models based on Markov random field priors. *Journal of the Royal Statistical Society C (Applied Statistics)*, 50, 201-220.

Fahrmeier, L., Kneib, T. and Lang, S. (2004): Penalized structured additive regression for space-time data: a Bayesian perspective. *Statistica Sinica* 14: 731-761.

Fahrmeier, L., Sagerer, F. and Sussmann, G. (2007). Geoadditive regression for analyzing small-scale geographical variability in car insurance. Blätter der Deutschen Gesellschaft für Versicherungs- und Finanzmathematik e.V. (DGVFM). 28: 47-65.

Fahrmeier, L., Sager, T. and Lang, S. (2009). Regression. Modelle, Methoden und Anwendungen. 2. Auflage. Heidelberg, Springer.

Frick, J.R., Goebel, J., Schechtman, E., Wagner, G.G. and Yitzhaki S. (2006). Using Analysis of Gini (ANoGI) for detecting whether two sub-samples represent the same universe. *Sociological Methods & Research* 34(4), 427-468.

De Leeuw E. and de Heer, W.(2002). Trends in Household Survey Nonresponse: A Longitudinal and international Comparison. In: Groves R.M., Dillman, D.A., Eltinge, J.L. and Little, R.J.A.: Survey Nonresponse. New York, Wiley. 41-54.

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

Groves, R.M. (2006). Nonresponse Rates and Nonresponse Bias in Household Surveys. *Public Opinion Quarterly*, Vol. 70, No.5, Special Issue, pp. 646-675.

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

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

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

Kroh, M. and Spieß, M. (2008). Documentation of Sample Sizes and Panel Attrition in the German Socio-Economic Panel (1984-2007). DIW Data Documentation 39.

Keeter, S., Kennedy, C., Dimock, M., Best, J. and Craighill, P. (2006). Gauging the Impact of Growing Nonresponse on Estimates from a National RDD Telephone Survey. Public Opinion Quarterly 70(5), 759-779.

Lucas, R.E., Clark, A.E., Georgellis, Y. and Diener, E. (2004): Unemployment Alters the Set Point for Life Satisfaction. *Psychological Science* 15(1), 8-13.

Miller, J. M. and Wedeking, J. (2003) The harder we try, the worse it gets? Examining the impact of refusal conversions and high callback attempts on the quality of survey data. Paper presented at the 58th Annual Meeting of the American Association for Public Opinion Research, Nashville.

Retzer, K., Schipani, D. and Cho, Y. (2004): Refusal Conversion: Monitoring the Trends. Paper presented at the annual meeting of the American Association for Public Opinion Research, Pointe Hilton Tapatio Cliffs, Phoenix, Arizona, May 11, 2004. 2009-05-26 http://www.allacademic.com/meta/p116057_index.html

Schimmack, Uli (2014), Separating Panel Effects from Age and Period Effects on Satisfaction Judgments in the German Socio-Economic Panel (SOEP), SOEPpapers on Multidisciplinary Panel Data Research (in preparation), DIW Berlin.

Schimmack, U. and Lucas, R. (2007): Marriage Matters: Spousal Similarity in Life Satisfaction. Schmollers Jahrbuch, *Journal of Applied Social Science Studies*, 127, p.1-7.

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

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

Schräpler, J.-P., Schupp, J. and Wagner, G.G. (2010). Individual and Neighborhood Determinants of Survey Nonresponse. SOEPpapers on Multidisciplinary Panel Data Research No. 288. DIW Berlin.

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

Seiler, C. (2010): Dynamic modelling of Nonresponse in Business Surveys. Technical Report Number 093. Department of Statistics, University Munich.

Singer, E. (2002). The use of incentives to reduce nonresponse in household surveys. In: Groves et al. (eds.): Survey Nonresponse, 163-177. Wiley.

Singer, E. (2006). Introduction: Nonresponse Bias in Household Surveys. *Public Opinion Quarterly*, Vol. 70, No.5, Special Issue, pp. 637-645.

Stoop, I.A. (2004). Surveying Nonrespondents. *Field Methods*, Vol. 16, No. 1, 23-54.

TNS Infratest Sozialforschung (2011). SOEP 2006 – Methodenbericht Wiederbefragung von Panelausfällen zum Befragungsjahr 2006 (Welle 23) des Sozio-oekonomischen Panels (SOEP 2006 Innovationsprojekte). SOEP Survey Papers 55: Series B. Berlin: DIW/SOEP.

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