

“Labour Market and Social Security”: A New Panel Study for Research on German Social Code II

Mark Trappmann, Bernhard Christoph, Juliane Achatz, Claudia Wenzig

1 Goals of the Study

The central aim of the Institute for Employment Research's (IAB) 'Panel Study Labour Market and Social Security' (PASS) is to examine the individual and social consequences of the implementation of the new Unemployment Benefit II in a household context (Achatz et al. 2007). For obtaining this goal, the most important research questions to be answered are:

- 1 Which are the pathways that lead people into unemployment and in particular: which of these lead into long-term unemployment and dependency on Unemployment Benefit II?
- 2 How does the social situation of persons and households change when they receive such benefits?
- 3 Which are the subjective/cognitive strategies employed to cope with long-term unemployment and one's individual dependence on public transfers? Will attitudes of the respondents that are constitutive for their actions change over time?
- 4 What are the institutional arrangements for getting in touch with the agencies responsible for the provision of Unemployment Benefit II? And which are the typical institutional procedures applied to accomplish a reintegration of recipients into the labour market in due time.
- 5 Which are the factors that support people in overcoming unemployment and benefit recipiency?

A study-design that is suitable to answer these research questions will face several challenges:

- i) Answering the research questions requires longitudinal data.
- ii) The respondents' household context is of essential importance.

- iii) The main research interest is on recipients of Unemployment Benefit II, a population that is particularly difficult to handle in quantitative surveys. The main problems with this population are connected to difficulties with contacting respondents (low telephone coverage), their low willingness to cooperate with the interviewer (e.g. due to negative attitudes of recipients towards public institutions) and potential language (above average percentage of migrants among recipients) problems.
- iv) Moreover, for several reasons a comprehensive scientific survey on recipients of Unemployment Benefit II requires that non-recipients are covered likewise. Among these non-recipients, persons that (except for recipient status) live in circumstances that are comparable to those of recipients are of particular interest for many research questions for they might serve as a control group or for analyses of the inflow into recipiency.

All these challenges require a study design that is tailored to the particularities of the research questions and the study population. We will present the details of such a design throughout the following paragraphs. Paragraph 2 gives an introduction to the overall sampling design and the two subsamples drawn. In paragraph 3 we will introduce the measures employed for coping with our difficult survey population. Paragraph 4 outlines a refreshment concept and paragraph 5 presents the most important modules of the survey questionnaire. Finally, in paragraph 6 a first glimpse on specific features of our control-subsample will be given.

2 The Sampling-Design of the IAB Panel Study 'Labour Market and Social Security' in the Context of Current Research on the Social Code II

The design of our study is made up by two independent subsamples, which are connected at the first sampling stage via the selection of identical primary sampling units (postcode regions) (compare figure 1). The Postcodes were selected from a postcode register using pps-sampling.¹ In order to guarantee a good mapping of regional characteristics in the sample, the pps-sampling procedure was stratified. The stratification characteristics employed were political district and municipal size (classified according to the BIK10 classification, compare Behrens 2005).

The first subsample was based on the Federal Employment Agency's Registers of Unemployment Benefit II-Recipients. From these registers a clustered random sample of so-called 'Bedarfsgemeinschaften'.² Even though these 'Bedarfsgemeinschaften' might not include all members of the household, it is always the entire household that will be interviewed. It is planned to interview 6,000 households from this register-based sample. This subsample can be used to draw inferences on the population of all

¹ In our particular case this requires an adaptation of the procedure, since the 'size' of a postcode region is usually different for both subsamples. For details on this problem compare Rudolph and Trappmann (2007).

² A so-called 'Bedarfsgemeinschaft' (which is sometimes translated as 'community of need') includes all persons in a household receiving benefits jointly (i.e. as a joint payment). In the majority of cases the 'Bedarfsgemeinschaft' and the household will be the same, which in particular applies in the case of (married or unmarried) couples and parents with children below 25 years of age. However, under specific circumstances the 'Bedarfsgemeinschaft' might not include all household members or a household in which everybody gets benefit payments might be made up by more than one 'Bedarfsgemeinschaft'. An example for the former is if a grown-up child lives with its parents and earns just enough to make its own living but has insufficient means to support his mother and father – in this case the 'Bedarfsgemeinschaft' will only include the parents. An example for the latter is a three- (or more) generation-household. Since a 'Bedarfsgemeinschaft' may only be made up of two generations, this type of household will be made up of two such 'Bedarfsgemeinschaften', one consisting of the grandparent(s) and one of the parent(s) and kid(s).

households in which at least one 'Bedarfsgemeinschaft' exists.

The second subsample was drawn from a commercial database covering all addresses at which at least one private household resides.³ It was clustered using the same spatial units (i.e. postcode regions) as in case of the first sample. The sample was disproportionally stratified by the attribute 'Status'.⁴

In order to draw the sample, first, all addresses chosen from the database were visited as part of an address-walk by interviewers of our field institute and all the names on the doorbell-panels were written down. After such units that were obviously in commercial use were excluded, in a second step one bell-sign from each doorbell panel was drawn. This was done centrally at the field institute and not by the persons doing the walk. In case that there was more than one name written on the bell-sign, a further step was taken to select a target person, whose household should be interviewed.⁵

The sample-size to be realised in this subsample is also 6,000 households. The second subsample can be used to draw inferences about all households in the Federal Republic of Germany. The disproportional stratifica-

³ We used the database MOSAIC provided by the company microm. On the coverage of this database compare Rudolph and Trappmann (2007).

⁴ The status indicator used was the variable "MOSAIC Sozio" included in the database. It comes as a nine-stage ordinal scale which was created by aggregating the occurrence of certain status-relevant attributes for small spatial units. Examples for variables that were used in this process were the frequency of high-status professions, of self-employment, of academic titles or information on automobile-possession. In addition, some characteristics were added that were available for higher-level spatial units only, like unemployment rates or purchasing-power (compare Kueppers 2005).

⁵ It can be assumed that in many cases the different names on the bell-sign refer to a common household. If this should be the case, taking the third step would be unnecessary. The advantage of this procedure is however, that it always generates a specific target-household, even for flats at which more than one household resides. Thus the contact phase of the interview is significantly shortened, because the interviewer will not have to ask the contact persons questions about other households potentially living at this address.

tion of this subsample allows a more efficient treatment of many research questions related to recipients of Unemployment Benefit II since it increases the number of cases available in various relevant subpopulations (such as persons with a high risk of becoming recipients or low income households). Thus parameter estimates will in many cases have a lower standard error.

In addition it is possible to combine both subsamples and to draw inferences on all households in Germany from the combined sample. Details on this can be found in Rudolph and Trappmann (2007).

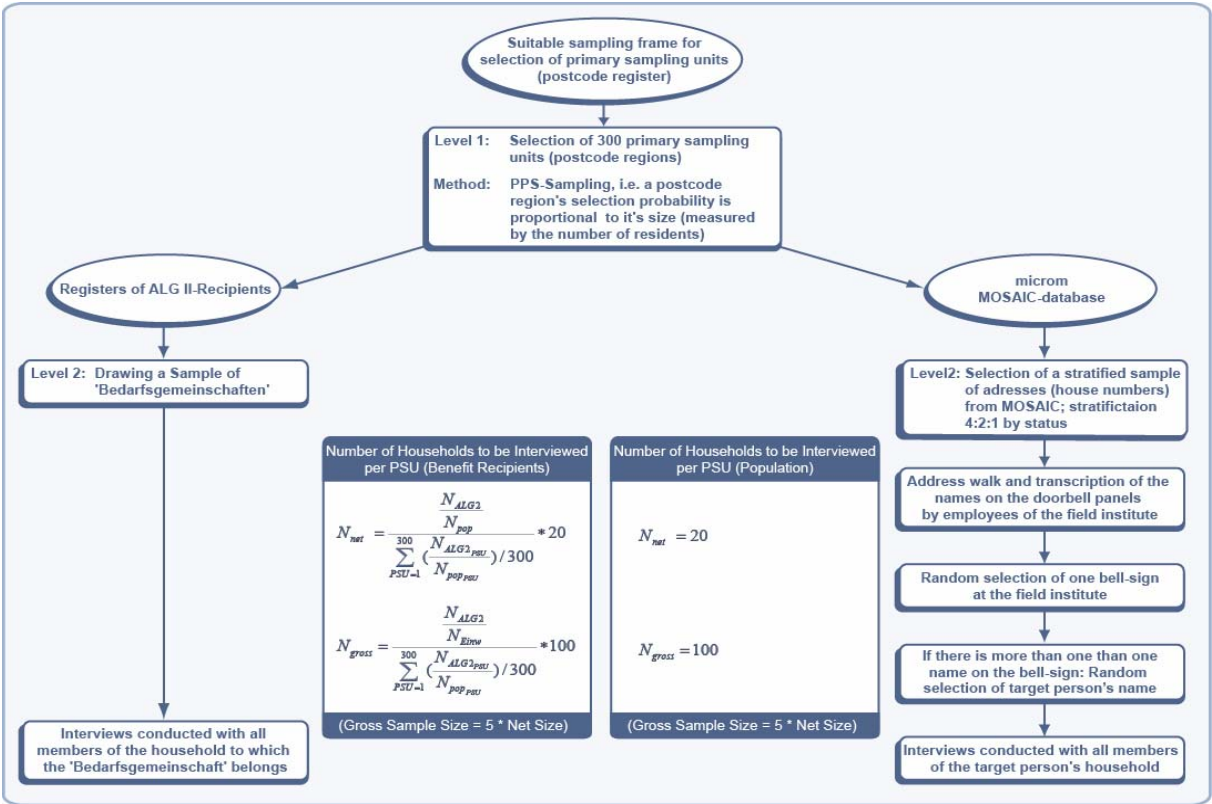


Figure 1: Sampling design of the IAB-Panel Study „Labour Market and Social Security”

3 Measures for Improving Response Rates

The planned panel study will be conducted to a particularly difficult survey population. A substantial part of this population consists of benefit recipients and, in addition, there is also a disproportionately high percentage of respondents that have a rather poor level of formal education and that have a comparably low social status.

Problems faced when trying to interview these groups are for example their higher tendency to relocate, which makes localising them rather difficult (Weiss und Bailar 2002, 87) and their generally higher tendency to refuse to participate in empirical surveys (Goyder 1987, 83).⁶ Since this is so, it is necessary to undertake special efforts in order to achieve a response rate that is at least somewhat comparable to the ones obtained in surveys of the general population. This applies even more, since the interviews conducted for our study might be mistaken for one of the inspection calls used by the employment agency to disclose cases of wrongfully obtained benefit payments. Thus recipients might (mistakenly) assume negative consequences to result from their participation in the survey (Schnell 2007).

Encouraged by an expertise written by Schnell (2007) and experiences made with a recent cross-sectional survey of benefit recipients (Infas 2006) the IAB-Team decided to take the following measures in order to address these problems.

1. Mixed Mode: in order to respond to the problems connected to difficulties in contacting target persons by phone, caused by low landline tele-

⁶ Reasons for this might be their more negative attitudes towards the organisations conducting the research (i.e. in our example the Federal Ministry of Labour and Social Affairs and the Federal Employment Agency), their less favourable opinions about the use of scientific surveys in general or that they might show rather little interest in the research topic (compare Schnell 1997, 202ff., Schnell et al. 2005b, 313). A comparable line of argument can also be found in Groves und Couper (2002, 38).

phone coverage⁷, regular changes of mobile phone numbers and frequent change of address, a mixed interview mode was employed. Persons that could not be contacted by phone were visited by an interviewer at their home, where the survey was conducted. In many U.S.-studies on benefit recipients, this measure has shown to be very effective in improving response rates (Cantor und Cunningham 2002, 70ff.). And also in our own cross-sectional survey mentioned above, this measure was effective in increasing response rates of difficult respondent groups. Even though it might have been equally or even more effective, conducting the entire survey in CAPI or PAPI-mode would not have been possible for financial reasons.

2. Incentives: For each respondent-household 10 Euro are available for buying incentives in every wave. During the first wave a large part of this Budget was used to finance an incentive to be sent along with the announcement letter. In addition, each respondent who agreed to be re-contacted for wave 2 will receive a personalised lottery ticket in-between waves.
3. Refusal-Avoidance-Training: In order to reduce refusals, the IAB required the field institute to employ a special training course for their interviewers. The programme by Schnell (Schnell and Dietz 2006), which is based on Groves and McGonagle (2001), was adapted to the specific requirements of the study.
4. Questionnaire-Translations: It is very likely that our survey population includes a disproportionately high number of migrants. Since their German language proficiency at least in some cases could be rather low, we translated our survey-instrument into Turkish, Russian and English. In the telephone-field these instruments were administered by interviewers that are native speakers of the respective language. It would

⁷ In the IAB's cross-sectional survey of Unemployment Benefit II recipients mentioned above (Infas 2006), which was also conducted as a mixed-mode survey, 20 percent of respondents stated that they did not possess a landline phone. However, since the unavailability of a landline phone in itself already constitutes an obstacle for contacting a potential respondent, this value should be considered a very conservative estimate of non-coverage in this particular population.

have been too expensive, however, to let all personal interviews be administered by native speakers. Therefore, the strategy employed in personal interviews was, to transfer respondents to the telephone field whenever possible. Where this could not be done, the CAPI-interviewers used a written foreign-language-version of the questionnaire as translation-aid.

These measures taken to improve response rates were supported by an exceptionally good database for conducting nonresponse-analyses. One of the reasons for this rather favourable situation is that we used the database MOSAIC for drawing our sample. In addition to the address-data used for sampling, MOSAIC includes a couple of other variables that are suited for predicting nonresponse. These variables were merged to both subsamples of our survey. Among these variables are attributes that can be used to predict refusals – such as e.g. social status – as well as information that can be used to predict potential loss of contact with respondents, as e.g. the rate of households moving away from the respective region over the year. All this information is available for relatively small spatial units.

A second explanation for our good database for nonresponse analyses is the fact that the sampling frame for the register-sample includes individual level variables like highest educational degree that might help predict participation

4 Sketching a refreshment concept

When creating a refreshment concept for a panel study, it is essential, do distinguish between new entries to and refreshments of the survey population. Including new population entries is essential in order to make valid inferences for the cross-sections in each wave. In contrast to that, drawing refreshments is necessary in order to adjust for the reduction in sample size caused by panel mortality. Thus it is also a measure to maintain the sample's statistical power, i.e. the probability that relationships present in the population can be revealed by using the sample.

4.1 Modelling New Population Entries

It seems reasonable to discuss the modelling of new entries to the population separately for both subsamples of the study.

In the register-based sample, the population consists of all households in Germany in which at least one 'Bedarfsgemeinschaft' resides. Drawing new entries to this population is only possible by again using the Federal Employment Agency's registers. The proper way to proceed would be to draw an additional sample from all those persons that newly claimed benefits since the original sample (or in later waves: the last refreshment sample) was drawn.

In contrast to that, the microm-sample constitutes a standard sample of the German resident population, notwithstanding the disproportional stratification of the sample. Thus it is not necessary to draw an extra sample to account for new population entries at the household level, since the households of the general population regenerate by themselves.⁸ Persons that die during the course of the panel will be replaced by persons growing old enough to be interviewed.

Further complexity is added to the concept by the fact, that households are dynamic entities. Households will split or will gain new members, not only by birth but also by new members moving into the household or new households will be created. In order to be nonetheless able to draw valid inferences from this sample, all households will be tracked in which at least one original sample member resides.

The same concept is also applied in the German Socio-Economic Panel (GSOEP), the sole German panel study with comparable sample size.

A specific problem, however, remains, which is modelling the new population entries caused by international migration. This is a tricky task, because a refreshment-mechanism like the one described above will only account for a selective part of new entries caused by migration. While it

⁸ This is a fundamental difference to specific populations like benefit-recipients, to which exactly the opposite applies: New benefit recipients will not exclusively come from current recipient households.

will take account of those migrants that enter existing panel households – e.g. by marrying a household member – migrants who do not enter existing households will not be considered. However, correctly modelling all new population entries generated by migration is not a simple task, since in Germany no single register exists, from which a potential refreshment sample could be drawn (Rendtel, Pannenberg und Daschke 1997, 272). On the other hand, using a screening procedure would be much too expensive, for the size of the original sample necessary to identify a sufficient amount of persons from such a small population – as recent migrants are – would have to be quite large indeed. The GSOEP, for example, not until 1994 drew a refreshment sample for this population, after 10 years of running the panel (Pannenberg et al. 2005, 154). Moreover, due to the difficulties involved in drawing such a sample, the one drawn by the GSOEP did – strictly speaking – not meet all the requirements usually considered necessary to be met by a probability sample.⁹

4.2 Refreshments of the Survey Population

The sole reason for drawing a refreshment sample is to compensate for reductions in sample size caused by panel mortality. In contrast to that, a correction for possible selectivity of the panel-attrition process via a refreshment sample is not feasible. Such a correction is usually accomplished by modelling attrition. In the GSOEP for example this is done by

⁹ For one, in its migrant sample, the GSOEP used persons, that had declared their willingness to participate in further studies during prior screenings conducted by the GSOEP's field institute. The chances are, that the population generated by this procedure is a selective one. In addition, since this sample still included insufficient cases, further respondents were recruited via snowball sampling, a method which does not allow for a proper calculation of inclusion probabilities (Rendtel, Pannenberg and Daschke 1997, Burkhauser, Kreyenfeld and Wagner 1997).

applying a so-called propensity weighting.¹⁰ Since significant reductions in sample size might only occur after some while, it seems not necessary to implement refreshments early in the panel. This is particularly true if new population entries are properly covered as described above.

5 Modules of the Panel Study „Labour Market and Social Security“

In each household one household questionnaire and one individual questionnaire for each person aged 15 or above is administered. Senior respondents above the age of 65 are administered a short version of the questionnaire.

The household questionnaire contains modules on household composition, dwelling, deprivation, child care, and household income including all spells of Unemployment Benefit II recipience since the introduction of this measure in January 2005 and including potential sanctions.

The individual questionnaire includes modules on education, employment, personal income, attitudes, home care, health, social networks, job search. Moreover it covers participation in programmes for Unemployment Benefit II recipients since 2005 and recipients' interaction with the agencies responsible for administering the benefits (in most cases called ARGE or Jobcenter). Information on the participants' employment history will be collected in the second wave of the panel.

¹⁰ Propensity weighting is a procedure in which in a first step participation in the following wave is predicted by one or more multivariate models. In a second step, longitudinal weights are calculated as the reciprocal of the estimated re-participation probabilities. Propensity weighting is a procedure that requires drop out processes to follow a missing at random (MAR)-mechanism, which means that drop outs are random, given the covariates included in the statistical model.

6 Results of the sampling procedure – preliminary results

Since fieldwork for wave one of the panel study has not finished yet, only preliminary results can be presented in this paper. We concentrate in our paper on the discussion of how disproportional stratification by the microm status index has worked out. The aim of the stratification was the inclusion of a larger number of people with a high risk of becoming dependent on unemployment benefits. While we can give no direct assessment of this risk using only data from the first wave of the study, we can assume that it is highly correlated with variables like education, employment status, income, wealth, or job stability.

Table 1 shows proportions (and in two cases means) for selected risk factors by microm-status (see columns 2, 4 and 6). While Column 8 gives the unweighted mean of these proportions which is equal to the expected proportion in an unstratified sample (as the three status classes are terciles of all households), column 9 gives the proportion weighted by the number of households (or persons, respectively) in the study so far. Column 10 indicates the “gain”, i.e. the additional percentage of persons with the respective risk factor in the study compared to an unstratified sample.

The results show that, while differences between the three status groups in many instances are quite pronounced, the ‘gain’ is typically between three and ten percent. The highest gain, however, is made for Unemployment Benefit II recipients since 2005. The proportion is 14.8 percent higher in the stratified than it would be in an unstratified sample. This indicates that we can be optimistic that the cumulation of risk-factors in the stratified sample will allow us to observe a considerably higher number of transitions than would be possible in an unstratified sample.

Table 1: Preliminary Results of the Sampling Procedure

HOUSEHOLD INDICATOR-VARIABLES

Trait	proportion in class 1 (low)		proportion in class 2 (med.)		proportion in class 3 (high)		unweighted	weighted	gain
	n1	n1	n2	n2	n3	n3	mean	mean	
rented house/flat	0.540	1538	0.458	1041	0.453	660	0.484	0.496	2.52%
no car	0.202	1606	0.146	1080	0.125	680	0.158	0.168	6.83%
deprivation index>6	0.181	1607	0.122	1080	0.070	683	0.124	0.139	12.18%
unempl. benefit 2 (ever)	0.157	1604	0.080	1078	0.055	674	0.097	0.111	14.83%
savings<10,000€	0.717	1398	0.622	949	0.535	535	0.624	0.652	4.37%
kids in household	0.581	1607	0.534	1080	0.458	683	0.524	0.541	3.17%
single parent hh	0.045	1607	0.039	1080	0.022	683	0.035	0.039	8.86%

PERSON INDICATOR-VARIABLES

Trait	proportion in class 1 (low)		proportion in class 2 (med.)		proportion in class 3 (high)		unweighted	weighted	gain
	n1	n1	n2	n2	n3	n3	mean	mean	
low formal education	0.308	1906	0.314	1305	0.211	764	0.278	0.291	4.96%
no vocational/tertiary degree	0.191	1908	0.157	1307	0.131	766	0.160	0.168	5.48%
no regular full- or part-time job	0.475	1999	0.425	1381	0.405	802	0.435	0.445	2.31%
fixed-term employment	0.112	919	0.096	695	0.047	384	0.085	0.094	10.54%
bad health (self-rated)	0.040	2008	0.033	1387	0.024	805	0.032	0.035	7.38%
not born in Germany	0.122	2010	0.091	1386	0.071	807	0.094	0.102	7.79%
parent or grandparent not born in Germany	0.251	1997	0.211	1375	0.195	794	0.219	0.227	3.69%
father no secondary ed.	0.051	1553	0.038	1079	0.029	653	0.039	0.042	7.59%
mother no secondary ed.	0.062	1549	0.050	1061	0.035	651	0.049	0.053	7.40%
no close friends	0.042	2010	0.039	1385	0.020	808	0.034	0.037	9.46%

METRIC VARIABLES

Trait	mean in class 1 (low)		mean in class 2 (med.)		mean in class 3 (high)		unweighted	weighted	gain
	n1	n1	n2	n2	n3	n3	mean	mean	
equiv. income (hh)	1324	1374	1594	914	1973	535	1631	1535	6.25%
job income (person)	2308	889	2607	689	3325	396	2747	2617	4.98%

7 Conclusion

The IAB panel study ‘Labour Market and Social Security’ is an important new database for research on unemployment in general and recipients of Unemployment Benefit II in particular. It combines state-of-the-art methods with some innovative features like e.g. the use of a commercial database for stratifying the general-population-subsample. This use of two different sampling frames in particular – in addition to the general-population-subsample a second subsample based on the Federal Employment Agency’s registers – allows for an efficient allocation of resources with respect to various research questions from this field.

With this paper we tried to give an overview of these methodological features and some of the general topics covered by the study. Needless to say that a panel survey is always work in progress, so there are still a lot of things that will come to PASS.

8 References

Achatz, Juliane et al. (2007): IAB-Panelbefragung von Haushalten im Niedrigeinkommensbereich. Entwurf für ein Rahmenkonzept. In: Promberger, Markus (Hrsg.) Neue Daten für die Sozialstaatsforschung: Zur Konzeption der IAB-Panelerhebung „Arbeitsmarkt und Soziale Sicherung“. Nürnberg: IAB Forschungsbericht Nr. 5/2007 (in print).

Behrens, Kurt (2005): Stadtregionen. Von Boustedt zu BIK. In: Arbeitsgruppe Regionale Standards (Hg.) (2005): Regionale Standards, Mannheim: ZUMA, S. 59-92.

Burkhauser, Richard V./Kreyenfeld, Michaela/Wagner, Gert G. (1997): The German Socio-Economic Panel. A representative sample of reunited Germany and its parts. In: Vierteljahrshefte zur Wirtschaftsforschung, 66, S. 7-16.

Cantor, David/Cunningham, Patricia (2002): Methods for obtaining high response rates in telephone surveys. In: Ver Ploeg, Michelle/Moffitt, Robert A./Citro, Constance F., S. 55-85.

Groves, Robert M./Couper, Mick P. (2002): Designing surveys acknowledging nonresponse. In: Ver Ploeg, M./Moffitt, R.A./Citro, C.F., S. 13-54.

Groves, Robert M./McGonagle, Katherine A. (2001): A theory guided interviewing training protocol regarding survey participation. In: Journal of Official Statistics, 17, S. 249-265.

Goyder, John (1987): The Silent Minority: Nonrespondents of Sample Surveys. Cambridge: Polity Press.

Infas (2006): IAB-Projekt: „Zusammenlegung von Arbeitslosen- und Sozialhilfe mit Einführung des SGB II: Übergangsmuster, Strukturen und Folgen für die Betroffenen“. Repräsentative Querschnittsbefragung. Abschlussbericht zur methodischen Umsetzung, Bonn: Infas.

Kueppers, Rolf (2005): MOSAIC von microm. In: G. Grözinger/W. Matiske (Hg.) (2005): Deutschland Regional. Sozialwissenschaftliche Daten im Forschungsverbund, München: Hampp, S. 95-104.

Pannenberg, Markus et al. (2005): Sampling and Weighting. In: Haisken-DeNew, John P./Frick, Joachim R. (Hg.) (2005): Desktop Companion to the German Socio-Economic Panel Study (SOEP). Version 8.0, Berlin: DIW, S. 153-186.

Rendtel, Ulrich/Pannenberg, Markus/Daschke, Stefan (1997): Die Gewichtung der Zuwanderer-Stichprobe des Sozio-oekonomische Panels (SOEP). In: Vierteljahreshefte zur Wirtschaftsforschung, 66, S. 271-285.

Rudolph, Helmut/Trappmann, Mark (2007): Design und Stichprobe des Panels „Arbeitsmarkt und Soziale Sicherung“ (PASS). In: Promberger, Markus (Hrsg.) Neue Daten für die Sozialstaatsforschung: Zur Konzeption der IAB-Panelerhebung „Arbeitsmarkt und Soziale Sicherung“. Nürnberg: IAB Forschungsbericht Nr. 5/2007 (in print).

Schnell, Rainer (1997): Nonresponse in Bevölkerungsumfragen. Ausmaß, Entwicklung und Ursachen, Opladen: Leske+Budrich.

Schnell, R. (2007): Alternative Verfahren zur Stichprobengewinnung für ein IAB-Haushaltspanel mit Schwerpunkt im Niedrigeinkommens- und Transferleistungsbezug. In: Promberger, Markus (Hrsg.) Neue Daten für die Sozialstaatsforschung: Zur Konzeption der IAB-Panelerhebung „Arbeitsmarkt und Soziale Sicherung“. Nürnberg: IAB Forschungsbericht Nr. 5/2007 (in print).

Schnell, Rainer/Dietz, Cornelia (2006): CATI-RAT. Multimediale Interviewer-Schulung für CATI-Interviewer. Unveröffentlichtes Manuskript, Universität Konstanz, Center for Quantitative Methods and Survey Research.

Schnell, Rainer/Hill, Paul/Esser, Elke (2005): Methoden der empirischen Sozialforschung. 7. Auflage, München: Oldenbourg.

Weiss, Charlene/Bailar, Barbara A. (2002): High response rates for low-income population in-person surveys. In: Ver Ploeg, Michelle/Moffitt, Robert A./Citro, Constance F., S. 86-104.