Survey Research Methods (2013)
Vol.7, No.2, pp. 103-114
ISSN 1864-3361
http://www.surveymethods.org

The Recruitment of the Access Panel of German Official Statistics from a Large Survey in 2006: Empirical Results and Methodological Aspects

Boyko Amarov and Ulrich Rendtel
Freie Universität Berlin, Germany

In 2004 Germany's Federal Statistical Office (Statistisches Bundesamt, Destatis) started the recruitment of an access panel (AP) from participantes in the German microcensus (MC), a large household survey. This access panel, a pool of persons willing to take part in voluntary surveys, currently serves as the sampling frame for the DE-SILC, the German subsample of the European Union Statistic on Income and Living Conditions. Sampling from panelists rather than directly from the population promised lower survey costs due to easy access to the AP participants and higher response rates. While participation in the MC is mandatory by law, joining the AP is voluntary. Approx. 10 percent of the MC households agree to enter the panel. In this work we examine the recruitment from the 2006 MC using socio-economic and demographic characteristics available in both the AP and the MC to explore the selectivity of the recruitment process. We also discuss the implications of German privacy protection legislation for this analysis. Finally we consider the longitudinal use of the AP in a methodological discussion on the question whether samples from the AP can be regarded as probability samples from the general population.

Keywords: Access Panel; Official Statistics; Recruitment Propensities

1 Introduction

At the European Union (EU) summits in Lisbon and Nice in 2000 the reduction of poverty in Europe was assigned a high priority. A key instrument in this aspect was the method of open coordination for governance in the area of social policy, defined and adopted by the Lisbon Council. With this the exchange of best practices in fighting poverty between the EU member states became an important part of the coordination. In order to evaluate social policies, the EU required timely and reliable measures of poverty that are comparable between the EU member states. Until 2001 the primary source of such data in Europe has been the European Community Household Panel (ECHP) which attempted to harmonise the entire survey process. This approach proved to be problematic in the EU and the panel was discontinued. In 2004 it was superseded by the European Union Statistics on Income and Living Conditions (EU-SILC) where harmonisation was to be achieved at the output level. This meant that while the joint quality standards adopted for EU-SILC stipulated a minimum sample size and the concepts to be measured by the survey, the member states were given considerable flexibility for its implementation (Mejer 2003). Being free to choose the best instrument for the task, some of them integrated the questionnaire into existing surveys, e.g. the British Household Panel Survey in the UK, while other states launched entirely new surveys.

Contact information: Ulrich Rendtel, Freie Universität Berlin, Germany, e-mail: Ulrich.Rendtel@fu-berlin.de

An obvious choice for Germany would have been to integrate EU-SILC into the annual microcensus (MC), a large household survey with mandatory participation. Such an integration was already done in the case of the Labour Force Survey which became part of the MC-questionnaire in 1968. However, this option was discarded because of concerns about the data quality of the MC itself as the additional respondents' burden from the EU-SILC questionnaire was considered too high.

A key requirement of the EU-SILC quality standards was that the new survey was to be selected at random according to a probability sampling design. This ruled out non-probability methods such as quota-sampling, the traditional method employed by Germany's Federal Statistical Office (Destatis, Statistisches Bundesamt) for voluntary surveys. Faced with declining response rates in household surveys (de Leeuw and de Heer 2002), Destatis decided to explore a new approach for drawing random samples from the general population that could compete with quota-sampling in terms of cost. Instead of using the MC itself, Destatis started the recruitment of an access panel (AP) from the households leaving the MC after their fourth and final year of participation (Körner et al. 2006). This AP, a pool of households willing to take part in voluntary surveys of Destatis, currently serves as the sampling frame for DE-SILC (the German subsample of EU-SILC). The recruitment of the AP started with the MC 2004 after an experimental phase in 2001 in five federal states (Nimmergut et al. 2004). Reduced costs due to higher response rates and easier access to the sampling units were a strong argument for the introduction of the AP in German official statistics. It was expected that persons who have al-

ready stated their willingness to participate in surveys would show lower nonresponse rates and shorter response times than the general population. Furthermore, the quality of the data collected from AP surveys was expected to be higher as most of the panelists have had four years of experience with the quite demanding MC interviews.

These advantages are in no case unique to this particular AP. Access panels have long been employed in market research and opinion polls, commonly in the form of online APs (Couper 2000). Despite these benefits, statistical offices still shun APs for the production of official statistics. To our knowledge, no other agency apart from Destatis has yet made use of this method for the collection of official data, although there are examples of active cooperations with academic access panels, e.g. the LISS and CentERdata panels at the university of Tilburg (van der Laan 2009; Hoogendoorn and Daalmans 2009). A major argument against their use for official statistics is the common lack of a theoretically sound recruitment plan based on probability sampling (Bethlehem 2009). Indeed, many of the APs deployed by commercial vendors rely on volunteer opt-in schemes like advertisement on Internet portals or convenience sampling where there is no information about those who do not join the panel and the possibilities to assess the selectivity of recruitment are limited.

This point was explicitly addressed by Destatis. Because of the voluntary participation in the AP it is essential to gain as good understanding as possible of the selectivity of recruitment. A crucial difference that sets this AP apart from many commercial ones is its recruitment from the MC. When a household enters the AP, a part of the variables measured in its last MC interview are stored together with its contact details. These data, called profile variables, describe socio-economic and demographic characteristics of the households and their members and are used as auxiliary variables for sampling from the AP. The availability of such auxiliary information enables the implementation of complex survey design, e.g. stratified sampling, that enhance the precision of estimation from the survey data. Apart from this, the profile variables are known for both participants and nonparticipants in the AP and thus open the possibility to control for selective recruitment through a direct comparison of the panelists with the MC. In this work we use this information to study the selectivity of recruitment in 2006, the third year of the AP.

The next section proceeds with a description of the recruitment and use of the AP until 2006. Section 3 presents a short overview of empirical results from an experimental recruitment phase of the AP in 2001 and discusses difficulties encountered in estimating models for the recruitment success and the limitations on the analysis they entail. The empirical results from a Logit model fitted to the data from the 2006 recruitment are presented in Section 4. Sections 5 and 6 conclude with a short methodological discussion on the current use and maintenance of the AP, a summary of the main results and an outlook to future work.

2 Recruitment and use of the AP

The German MC is an annual survey with a four years rotation cycle collecting data on income, employment status and living conditions. It is an area sample of dwellings as sampling units, covering approximately one percent of the households in Germany every year. Participation in the MC is mandatory, ensuring low nonresponse rates of around 5 percent, mostly due to noncontact (Statistisches Bundesamt 2008). The sample is divided into four rotation groups of approximately equal size and each one is interviewed annually during four consecutive years. After that, the group is replaced by a fresh sample.

At the end of their final interview with the MC, the persons living in private households at their primary residence are asked whether they would agree to be contacted in the future in order to take part in voluntary surveys conducted by Destatis. The invitation to join the AP does not mention specific topics of the upcoming surveys or a particular mode of data collection. Neither the households nor the interviewers were offered material incentives in case of successful recruitment in 2006.[1] How this request is presented to the households depends on the mode of the MC-interview itself. Most of the households complete a face-to-face interview but in case the interviewer fails to make personal contact, the household is left a paper MC questionnaire or is contacted by telephone (Lotze and Breiholz 2002a,b). In 2006 approximately 20 percent of the households answer a self-administered mail questionnaire and another 2.5 percent take part in a telephone interview. Households that remain undecided about the AP at the time of the interview are left information materials along with a consent form and receive up to three reminders in case of nonresponse. Only whole households are admitted to the AP, so all adult household members are required to consent.

Upon entry, a part of the information from the household's last MC interview is stored in the AP database, mainly socio-economic and demographic characteristics, referred to as profile variables. These include the household's net monthly income in EUR (0–900/900–1300/1300–2600/2600–3600/3600 and more), its size and composition (single persons/couples with children/couples without children/single parents/other), the number of children under and over 18 years, the federal state and the municipality[2] (Gemeinde) of residence as well as the year and quarter of entry in the AP. At the level of individuals the AP records gender, year and month of birth, marital status, net monthly personal income, employment status and type (white collar/blue collar/civil servants, etc.), citizenship (German/Non-German), working hours per week and the highest school and professional training degrees obtained. As in commercial access panels these profile variables are primarily used for sampling purposes as their availability enables the realisation of

[1] This has changed in the years after 2006 and currently in almost all federal states the interviewers receive between 1 and 10 Euro for a successful recruitment.

[2] In our data the municipality code was assigned randomly to ensure privacy protection, so we were unable to use it in the analysis.

complex survey designs that enhance the precision of estimation with the survey data. The profile variables also make it possible to control for selective survey nonresponse since they are known for both respondents and nonrespondents.

Roughly 10 percent of the households agree to join the AP, although the exact participation rate could not be determined because the variable identifying the rotation groups of the MC was missing in the available data due to privacy protection. This success rate appears low, but it does not necessarily entail a nonparticipation bias (Groves 2006). A recent comparison of DE-SILC (2005–2007) with the German socio-economic panel (SOEP), however, gives reasons for concern. Frick and Krell (2010) pointed out large differences between income mobility measures in both surveys. Furthermore, they demonstrate implausible variations of results within DE-SILC itself. They refer to a change in the reference distribution from the 2005 cross-sectional population to the 2005–2006 longitudinal population. The corresponding samples differ by the first rotation quarter of DE-SILC that was selected from the German Households Budget Survey (HBS), a quota sample. The HBS was used as the AP in its start was too small to supply the whole DE-SILC sample. Therefore the variation between the cross-sectional and the longitudinal results reflect the omission of the HBS-rotation group from DE-SILC. This may be seen as an indication how important the requirement of a probability sampling may be. Although the SOEP cannot be regarded as a gold standard, these discrepancies still call for an analysis of the possible causes, one of which could be the selectivity in the recruitment of the AP.

Until 2006 the AP has served as a sampling frame for two samples of DE-SILC (2005, 2006) as well as the Survey of Births, a 2006 survey of women in Germany. It also provided a part of the samples for the 2005 and 2006 Surveys of Information- and Communication Technology. Figure 1 illustrates the process of recruitment and sampling and gives approximate recruitment- and survey response rates.[3] It is hard to assess whether the AP makes good on the promise of high response rates. All three surveys used self-administered mail questionnaires for data collection and the response rates of approx. 75 percent in the DE-SILC samples sound encouraging for this mode which is often reported to be prone to high nonresponse. This success was achieved with considerable effort, though. In case of nonresponse, households received up to three reminders and the field work stretched from the beginning of March to the end of May. The questionnaire was designed according to the principles of the tailored design method (Körner et al. 2005; Dillman et al. 2008). Interestingly, the response rates are close to Dillman's prediction of about 75 percent for mail surveys that follow the total design principles (Dillman 1978). There are examples in Germany that report similar response rates (Hippler 1985) without using an AP but it is hard to compare them with DE-SILC because those response rates vary dramatically depending on the topic, its saliency and sensitivity, the length and design of the questionnaire and the number and timing of reminders among other factors.

3 A Logit Analysis of the AP's recruitment

Unlike many studies of household survey nonresponse where failure to contact the sample units is a major cause of nonresponse (Schräpler et al. 2010), the analysis of recruitment success presented here compares the distribution of characteristics in the AP and the MC and is to be understood as conditional upon successful contact. Failure to contact the sampled households is indeed an issue in the MC and the noncontact rate has increased since 2005 when the MC moved to a continuous interviewing scheme throughout the whole year, but is still relatively low at about 5 percent.

There are three possible sources of nonparticipation in the AP that cannot be separated in the data and are analysed together. First, an adult household member may refuse to cooperate with the AP or the household may fail to respond after the last reminder. Third, the interviewer may simply omit the invitation to the AP.

Many factors could potentially influence the decision to participate and the literature on survey nonresponse provides theories that suggest the direction of their effects. The rational choice theory sets the decision whether to comply with a survey request in the framework of a costs-benefits analysis (Esser 1990). Expected opportunity costs, measured by the time needed to complete the interviews and the cognitive effort to answer the questions are thought to influence the decision of whether to comply with the request or not. Households with middle-aged working persons can thus be expected to be less willing to enter the AP because of lack of time. Privacy protection and confidentiality concerns can be regarded as a form of cost of the decision and are expected to have an adverse effect on the propensity to participate. Interviewers' reports from the experimental phase of the AP indeed point to privacy concerns as a major cause of nonparticipation (Nimmergut et al. 2004). Schneekloth and Leven (2003) suggest that intellectual milieus consisting of persons with higher education would be more sceptical towards statistical surveys because of confidentiality concerns. The reciprocity theory of behaviour (Groves and Couper 1998), however suggests a negative impact of low education on survey participation.

Körner and Nimmergut (2004) analysed the probability of successful recruitment in the experimental phase of the access panel in 2001 which took place in five federal states.[4] Using logistic regression models they report evidence of selectivity with respect to employment status, gender, age, household type and household income. Women were found to be more inclined to participate than men. Persons between 60 and 69 years were most likely to cooperate. With respect to employment status, self-employed persons were found to be less likely to participate than the rest. The lower income

[3] In the case of the ICT the AP was only used to supplement the main quota sample drawn from other sources. Due to the lack of reliable response indicators no response rates can be computed.

[4] Brandenburg, Bavaria, North Rhine-Westphalia, Thuringia and Hesse.

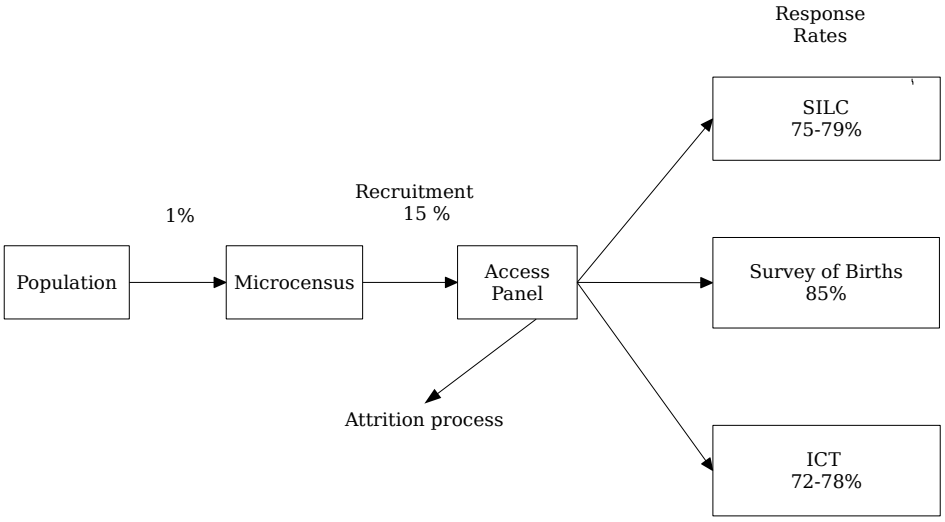


Figure 1. Recruitment and survey nonresponse in the access panel. Approximate recruitment and response rates in per cent (2005/2006).

groups (1300 EUR or less) were underrepresented. The federal state of residence also had a large impact on the participation propensity, but in their analysis the federal state was confounded with different modes of presenting the invitation to the households.

Since the above analysis the MC has undergone a major reorganisation of its field work when in 2005 it changed to continuous interviewing throughout the whole year instead of concentrating the field work to a single week (Lotze and Breiholz 2002a). The change was effected in order to accommodate a requirement of the Labour Force Survey to deliver data on a quarterly basis. By spreading the field work over the entire year the statistical agencies were able to reduce the number of interviewers and to employ the interviewer staff for longer periods. The field work was thus carried out by a generally more professional interviewer staff than prior to 2005.

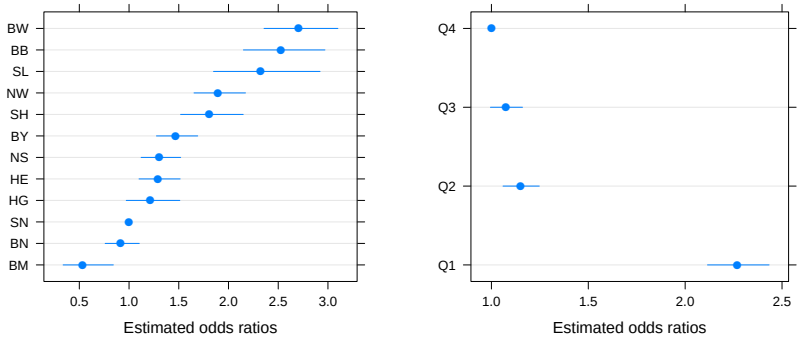
Central to the AP's design was the availability of the MC data that would enable control for selective recruitment. However, the MC is governed by a specific law, that determines the variables measured in the survey and the current edition of the law does not include a question about the agreement to participate in the AP. To complicate matters even more, German privacy legislation prevents the merging of information from voluntary surveys, including the AP itself, with data on individual level from the MC.[5] Both restrictions imply that the recruitment behaviour cannot be analysed directly with the MC microdata, so we base our Logit analysis entirely on separate frequency tables from the MC and the AP. Let X be set of all combinations of covariate values that are used in the Logit model and let N_X be the frequency counts of the cross-classification X in the MC. Now, let $N_{X,Y}$ be the frequency counts where the dependent variable Y indicates recruitment success ($Y = 1$) or failure ($Y = 0$). In our case, $N_{X,Y=1}$ is simply the number of observations in the AP for each cell of X. Standard routines are able to estimate a Logit model on the basis of the aggregated data from the two tables as input. This approach, however, is limited to using only variables that are recorded in the AP, a small subset of the information that is available in the MC.

Because of their use for sampling purposes, the profile variables are regularly updated. This is done either by retrieving the information from DE-SILC or – for those not selected or who fail to respond – from a short update questionnaire. Unfortunately, the database supplied by Destatis for this analysis contains only the up-to-date values of the profile variables. While the use of up-to-date records is of course adequate for sampling, it presents a real loss of information for a recruitment analysis as we have no access to the values of the profile variables from the time of entry in the panel. Apart from editing and data entry errors, consistency of AP and MC characteristics can only be fully guaranteed for gender and birth date, but not for variables like employment status or income. Because of this peculiarity of the data and in order to keep this effect as small as possible in our analysis we used the data from the most recent recruitment year available. These were the 2006 MC and the corresponding AP file from where we expect the smallest discrepancy between the original MC values and the AP profile variables.

An additional problem with the model estimation stems from the inability to distinguish the four rotation groups of the MC in our data (due to privacy protection). For this reason we estimated the model by repeatedly sampling one quarter of the MC at random and averaging the coefficients over the random samples.[6]

[5] The MC law is updated regularly. A variable indicating agreement to join the AP may by added in a future legislative cycle.

[6] The coefficients and confidence interval limits and standard errors reported in the figures in Section 4 and in Table 1 refer to averages over all random samples.



(a) Federal state. (b) Recruitment quarter.

Figure 2. Federal state of residence and recruitment quarter. Estimated odds ratios and 95% confidence intervals.

4 Empirical results

An exploratory analysis of the recruitment rates and the previous results of Körner and Nimmergut (2004) suggested a model at household level with main effects for federal state, net household income, employment status, weekly work hours, education level as well as the marital status of the person with the highest income in the household (HIP).[7] Furthermore, we included an indicator for households that have at least one member with a Non-German citizenship. We model the household composition through an interaction effect between the number of adults in the household (1, 2 and 3 or more) and the presence or absence of underage persons in the household as an indicator for children. Furthermore, we have included an interaction effect for the age and sex of the HIP. The estimated coefficients along with standard errors are summarised in the appendix in Table 1. The model is estimated with data on 283,367 households participating in the 2006 MC and 7,020 households recruited for the AP in the same year.[8] Positive coefficients imply higher probability of recruitment.

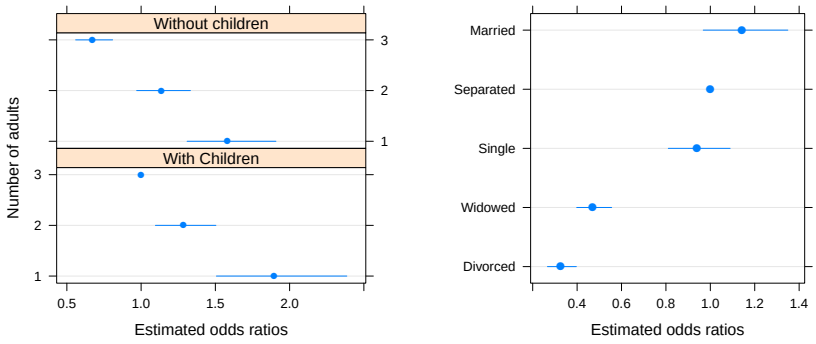
Two of the factors with the greatest influence on the participation probability are the federal state of residence and the quarter of recruitment of the household (Figures 2(a) and 2(b)). While the first variable refers to the field-work agency, which is organized at state level, the second variable refers to the period when the household is selected for the MC interview. With respect to the size of differential recruitment success one gets, for example, an odds ratio of about 5 when we compare Baden-Wuerttemberg and Bremen. The odds of successful recruitment in the first quarter of the year are more than two times higher than in the last quarter.

Although being only a speculation, we suggest that both effects reflect differences in interviewer training and motivation as well as field work organisation. While Destatis is responsible for the sampling design of the MC, the Statistical Agencies of the States (Statistische Landesämter) are in charge of the field work, including the recruitment and training of their interviewers. Köhne-Finster and Güllner (2009) present results from a survey of interviewers in the MC that shows differences in the composition of the interviewer staff across federal states. Regional differences may also play a role in the decision to participate. Interviewers' reports from the experimental phase of the AP indicate different degrees of privacy concerns in Brandenburg and Bavaria, for instance. Unpublished data on the time devoted to the training of the interviewers for AP's recruitment showed differences between the state agencies, but failed to explain the variation in the recruitment rates. Interestingly, the low recruitment success in Bremen ('BM' in 2 (a)) coincides with the frequent use of self-administered questionnaires in this state which amount to 68 percent of all MC interviews compared to a national average of 20 percent. Evidence from the experimental phase of the AP indicated that the written participation requests were much less effective compared to those made in a personal interview (Nimmergut et al. 2004). With the exception of Bremen, though, the recruitment rates do not appear to vary systematically with the share of households interviewed by self-administered questionnaires or CATI at the federal states level. Nevertheless, one would assume a dependence of the recruitment mode at the individual level. However, the mode of recruitment is not contained in the AP-database.

The significantly higher recruitment probability in the first quarter (Figure 2(b)) indicates substantial differences in the fieldwork over the course of the year. One might think of an interviewer who tries to achieve a target value of a 10 percent of recruitment rate as early as possible. However, the field agencies denied the existence of such targets. Both results show the high impact of the fieldwork's details on the recruitment success and indicate that the recruitment process is not really understood so far and that the current profile variables alone do not suffice to explain it. In order to gain better insight, the AP should record more details, especially information about the interviewers and the mode of data collection.

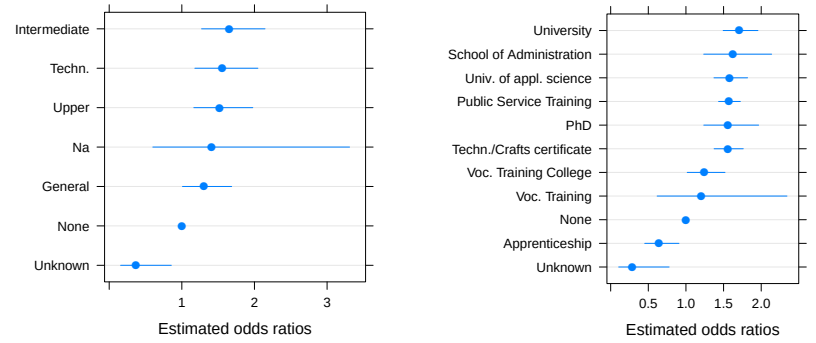
[7] The acronym HIP is motivated by the German word "Haupteinkommensbezieher", meaning the highest income person (HIP) in the household.

[8] Four federal states are excluded from this analysis because the local statistical agencies withdrew the permission to use their data: Thuringia, Mecklenburg-Western Pomerania, Saxony-Anhalt and Rhineland-Palatinate.



(a) Number of adults. (b) Marital Status.

Figure 3. Estimated odds ratios and 95% confidence intervals for the recruitment success in 2006: marital status of the HIP and household composition: number of adults with and without children.



(a) School degree. (b) Professional degree.

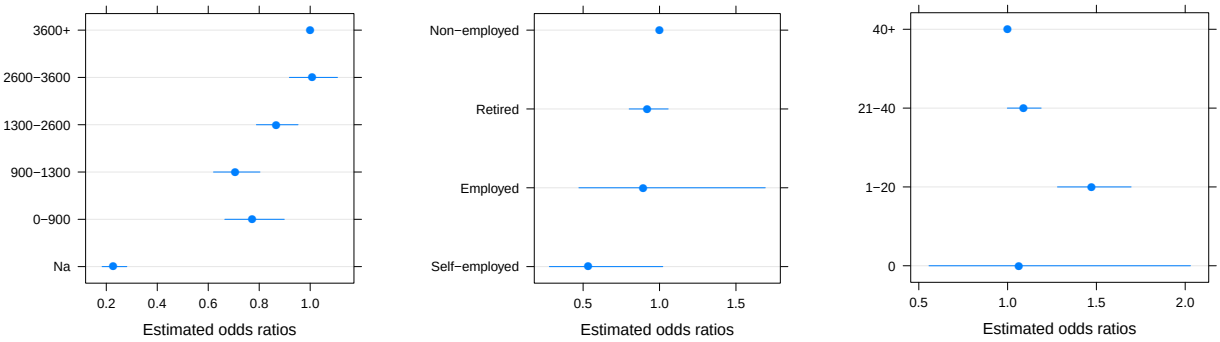
Figure 4. Estimated odds ratios and 95% confidence intervals for the type of highest school and professional degree achieved of the HIP.

Because the AP only admits whole households, i.e. where all adult household members give their consent, we expected single-person households to be overrepresented in the panel and this is indeed supported by the model estimates (Figure 3(a)). It is also in line with the findings of Körner and Nimmergut (2004) where single-person households appeared to be more likely to join. Households with children tend to be slightly less willing to join the AP compared to households without children. The Wald χ^2 test for the interaction term between the number of adults and the presence of children in the household is significant at the 5 per cent level, but Figure 3(a) shows that the largest negative effect on recruitment success is contained in the category of three or more adults with children, which is a relatively small group.

The probability of successful recruitment also appears to vary systematically between households with different marital status of the HIP (Figure 3 (b)). While there are no substantial differences between the estimated Logit coefficients for households with single, married and separated HIPs, those with a divorced HIP had an estimated 3.5 times lower odds to join than those with a married one. As we control for age the low recruitment rate for widowed and divorced persons cannot be interpreted as surrogates for old-aged persons. So one might think of persons who have recently experienced the death of a spouse or a divorce. Such events are connected with negative feelings and stress and are prone to reduce cooperation in a survey.

The participation behaviour also seems to differ with respect to the HIP's education level. Figure 4 displays the estimated coefficients and asymptotic confidence intervals for the types of school and professional training degrees of the HIP. Mainly there are – with the exception of "apprenticeship" – no substantial differences with respect to the type of schooling and professional training. Only HIPs without any school- or professional training degree seem to join the AP with a significantly lower probability. These results are to be regarded with a certain degree of caution, though. The level of education measured by these two variables can only increase over time (it is unusual to loose an already earned degree) and thus the proportion of persons with low education in any given cohort should decrease after each update of the profile variables. For both variables the category "unknown" indicates that the HIP has a school- or professional training degree but its type could not be determined. This is a small group in the population and the large confidence intervals for the coefficients reflect the uncertainty of estimation.

An implication of the reciprocity theory of respondent behaviour is that households with lower socio-economic status would be less willing to cooperate with surveys (Goyder 1987). The observed differences between the income groups



(a) Net Monthly Household Income. (b) Employment Status of HIP. (c) Weekly work hours.

Figure 5. Household income, employment status and hours of work per week of HIP. Estimated odds ratios and 95% confidence intervals.

(Figure 5(a)) tend to support this. The lowest two categories (0–1300 Euro per month) appear to be underrepresented in the AP, a finding also reported by Körner and Nimmergut (2004). In a recent nonresponse study of the SOEP, Schräpler et al. (2010) point to higher refusal rates for households in the highest income group, a result not supported by our findings with the AP. Although the profile of Logit coefficients is rather flat across income categories, the results are at odds with a standard hypothesis of an over-representation of mean income earners in surveys. The 'Na' category of the income variable results from item nonresponse in the MC itself. The large negative impact on the participation propensity is similar to results from research on panel surveys indicating that item nonresponse is a good indicator of unit nonresponse in subsequent panel waves, e.g. Rendtel (1995); Nicoletti and Buck (2004).

The participation propensity does not show significant differences between retired, employed and non-employed persons (HIP). Self-employed HIPs indeed had about 1.8 times lower estimated odds of joining the panel than non-employed or retired ones, but the effect is not significant at the 5 percent level, see Figure 5(b). The lower propensity of households with a self-employed HIP may be due to the high time pressure under which these persons tend to work in Germany. Therefore a low recruitment success would have been in line with general notions about time resources and survey participation. However, the model also includes the weekly working hours as a predictor for participation (Figure 5(c)). Households with part-time working HIP (1–21 hours per week) had the highest propensity to participate but there are almost no differences between persons with a regular amount of working hours (21–40) and persons who work even more (40+ hours). The rational choice theory would predict a significantly higher non-cooperation for the latter group. The estimated coefficient for non-working HIP is close to zero because of high collinearity with the non-employed category of the employment status variable.

With regard to age we observe a slightly higher participation propensity for persons aged 55–75. Female HIP in the group 18–25 years were more likely to join than male HIP in the same group, but the difference was not significant. A large drop in the participation propensity is found for

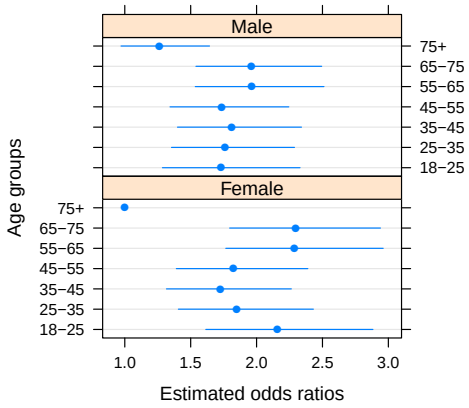


Figure 6. Estimated odds ratios and 95% confidence intervals for the probability of recruitment in 2006 for sex and age of the HIP.

HIPs older than 75. Groves et al. (1992) suggest that this behaviour might be related to health problems at this age. For femal HIPs the drop was more pronounced than for male ones (Figure 6).

Households with migration background were expected to be less cooperative partly due to language problems, a hypothesis supported by evidence from the SOEP (Schräpler et al. 2010). The AP, however, only allows to distinguish between German and a Non-German citizens, a rough indication for migration background.[9] The estimated coefficient for citizenship in Table 1 (appendix) shows markedly lower chance for successful recruitment of households with at least one foreign citizen. The estimated odds are almost two times lower than for household with only German citizens even after controlling for all the other variables in the model. In a study of the recruitment of the LISS AP van der Laan (2009) suggests language problems as a cause for the reluctance to cooperate. A study of immigrant respondents with a Turkish background in Germany, the majority of foreigners in Germany by Blohm and Diehl (2001) also supports this hypothesis.

[9] Some immigrants or their children may have obtained German citizenship.

5 Remarks on the Use and Maintenance of the AP

So far we analysed the recruitment process from the MC to the AP. This is only the first step towards the final DE-SILC sample. The next stages are the sampling design at the level of the AP and the realisation of the DE-SILC interviews after the design phase.

DE-SILC is selected from the AP according to a stratified simple random sampling design. Until 2006, Destatis used federal state of residence, net household income, household type and the employment status of the HIP as stratification variables (Horneffer and Kuchler 2008).

Let us consider the estimation of the total, t_y of a characteristic y defined for each i in the set of population elements U. Let S_1 denote the AP sample in a given year and S_2 the DE-SILC sample selected from S_1. Let $I_i^{(2)}$ be equal to 1 when $i \in S_2$ and zero otherwise. Finally, let $h = 1, \ldots, H$ index the strata. With n_h^* and N_h denoting the number of observations in the AP and the MC, respectively, the maximum likelihood estimate of the participation probability of person i in stratum h would be n_h^*/N_h. The selection probability for DE-SILC given the AP will be n_h/n_h^* where n_h is the number of elements sampled in stratum h. Finally, let π_i denote the inclusion probability in the MC of element i. The two-phase estimator (Särndal et al. 2003) for t_y will be:

$$\hat{t}_y = \sum_{i \in S_1} \sum_{h=1}^{H} \frac{y_i}{\pi_i} I_i^{(2)} \frac{N_h}{n_h^*} \frac{n_h^*}{n_h}$$

In the above, n_h^* cancels from the equation and the estimator does not depend on the random sampling of the AP. If the variables used to define the strata capture the pattern of nonparticipation, then the estimator will be unbiased. There are limitations to this approach, though. The small sample size of DE-SILC of about 14,000 households per cross-section does not allow arbitrarily deep stratification because much of the strata will be empty. Apart from this, the method can only use variables available in the AP.

The model presented in the previous section indicates the presence of other important variables such as the recruitment period, the marital status, the age and the education level of the HIP that are not among the stratification variables.[10]

The maintenance of the AP has three aspects that can affect the estimation based on survey data selected from it. First, as in every longitudinal survey, there is the phenomenon of panel attrition. In the course of time some participants become unwilling to cooperate with the panel, move to a new address and the follow-up routines of the statistical agencies fail, or simply move out of the survey area. Currently, Destatis estimates attrition probabilities for the purpose of weighting the surveys selected from the AP (Horneffer and Kuchler 2008; Körner et al. 2006). However, in order to account correctly for attrition, the database maintenance needs to properly record the movements of the panelists. Up to now, the AP does not try to distinguish between persons who have left Germany and have thus become ineligible and those who have simply dropped out of the panel. Second, the process of panel attrition needs to be examined using the information known for the attritors in order to assess its selectivity, an issue that is not addressed in published work. A rule that dismisses persons from the panel upon three consecutive refusals to take part in a survey directly links the nonresponse and attrition behaviour. At the time of writing this article we could not conduct a meaningful analysis of the attrition, because only the up-to-date information on the participants was available in the data. Even worse, some statistical agencies had deleted the profile variables of persons who had left the panel.[11] These difficulties can be overcome in the future when the statistical agencies deliver the longitudinal data of the AP.

The next question concerns the sampling from the AP. Currently, the selection of the DE-SILC sample is done from the pool of all panelists, ignoring the year of their recruitment, an approach common in commercial panels that lack regular recruitment phases. This implies a difficulty for the computation of inclusion probabilities of persons and households for DE-SILC. The key problem is that households have multiple opportunities to be selected for DE-SILC. Consider the situation in wave two of the AP. The first way to enter in DE-SILC's second wave is recruitment in wave one of the AP and non-selection for first wave of DE-SILC. The second way to enter DE-SILC is to be recruited in the AP in wave two. In order to calculate the total inclusion probability one would have to calculate both inclusion probabilities (both events are exclusive). We consider a very simplified setting with no changes in the population over time. Let A_1 and A_2 be the sets of elements recruited for the AP in waves one and two. Further, let D_1 and D_2 denote the sets of elements selected for first and second DE-SILC waves. Simple probability calculus leads to the following expressions for the selection probabilities:

$$\begin{aligned}
P(i \in D_1) =& P(i \in D_1 | i \in A_1) P(i \in A_1) \\
P(i \in D_2) =& P(i \in D_2 | i \in A_2 \cup A_1 \cap \overline{D_1}) P(i \in A_2 \cup A_1 \cap \overline{D_1}) \\
=& P(i \in D_2 | i \in A_2 \cup A_1 \cap \overline{D_1}) \times \\
& \left[(1 - P(i \in D_2 | i \in A_1)) \, P(i \in A_1) + P(i \in A_2) \right]
\end{aligned}$$

The expressions for the selection probabilities grow more complex with increasing number of waves and can become intractable.

6 Conclusions

The idea to generate a probability sample by sampling from a large national mandatory survey appears to be attractive as it offers the possibility to correct for selective recruitment schemes. For production of official data this idea was first implemented in 2004 by Germany's Federal Statistical Office on the basis of the German MC. Unfortunately, some

[10] The education level was added to the stratification variables in 2009.

[11] This problem was subsequently addressed by Destatis and should not occur in the later waves of the AP.

crucial aspects of German privacy legislation were not considered at the design stage of the AP: from the point of view of a recruitment success analysis it was a severe drawback not to include a question about the households' agreement to participate in AP in the MC questionnaire. For our analysis this implied the use of separate frequency tables from the MC and the AP. The database supplied to us for the analysis contained values of the profile variables that were partially inconsistent with the values in the MC. Although constantly updating the values of the profile variables is certainly good for sampling from the AP, it imposes problems for th recruitment analysis presented here.

Some common hypothesis about surveys participation propensity are affirmed in the recruitment success analysis: foreigners and self-employed persons seem harder to recruit than others. Persons older than 75 years seem to be less inclined to participate in the AP. Item-nonresponse in the MC is a powerful indicator of reluctance to cooperate with future surveys. Household size is an obstacle if complete participation of all household members is needed. The differences in participation propensity with respect to income, education, professional degree, family status and working hours are not so pronounced. The most important factors for the recruitment success (federal state of residence and period of recruitment) are probably associated with fieldwork. However, information data on interviewer training at Federal state level failed to reveal any clear relationship between the time devoted to the AP in the interviewers' training and the recruitment success.

Apart from rendering the estimation of models for the recruitment success cumbersome, the absence of a participation indicator in the MC microdata prevents the use of non AP-variables in the analysis. Important information would have been the mode of recruitment of the household (CAPI, CATI or self-administered) as well as more detailed data on income. Probably the most important variables influencing the decision to participate are the behaviour, experience and personality traits of the interviewers conducting the MC interviews as well as the respondents' past experience with the MC itself. Unfortunately, the microdata of the MC does not record information about the interviewers. This would not only make more extensive analyses of the AP's recruitment possible, but such information may help assess the quality of the MC data itself.[12]

The key question is whether the results from DE-SILC are reliable. If the recruitment process from the MC is only partly understood, will standard calibration routines help to end-up with unbiased population estimates? This question was addressed in a simulation experiment by Enderle et al. (2011) who mimicked the selection process of the AP. They conclude that it is essential to understand the selection process in order to correct successfully for selectivity. If important variables are omitted in the construction of the weights then the standard calibration to known MC totals will not help to compensate for such a deficiency. Therefore we recommend to put more emphasis on the statistical modeling of the recruitment process. This includes knowledge of the impact of interviewers and other details of the field work. Furthermore it should be possible to use the full range of MC-variables for the recruitment analysis.

Future research planned in cooperation with Destatis will examine the recruitment process over five years (2005–2009). The analysis over several years will help to assess the stability of the effects presented here over time.

Acknowledgements

Supported by the Deutsche Forschungsgemeinschaft (DFG) under research grant RE 1445/6-1 as a part of the Priority Programme 1292 "Survey Methodology". We also express our thanks to the Research Data Center of the Statistical Agencies of the States (Statistische Landesämter) and its branch in Berlin for providing us access to the data and for their assistance in this project. We also wish to thank the Federal Statistical Office for providing helpful comments.

References

Bethlehem, J. (2009). *Can we make official statistics with self-selection web surveys?* Proceedings of Statistics Canada Symposium 2008. Statistics Canada Catalogue no. 11-522-X. Ottawa, Ontario.

Blohm, M., & Diehl, C. (2001). When migrants interview migrants: On the survey participation of migrants. *Zeitschrift für Soziologie*, *30*, 223-242.

Couper, M. P. (2000). Review: Web surveys: A review of issues and approaches. *The Public Opinion Quarterly*, *64*, 464-494.

de Leeuw, E., & de Heer, W. (2002). Trends in household survey nonresponse: A longitudinal and international comparison. In R. M. Groves, D. A. Dillman, J. L. Eltinge, & R. J. A. Little (Eds.), *Survey nonresponse, wiley series in probability and statistics* (p. 41-54). New York: Wiley.

Dillman, D. A. (1978). *Mail and Telephone Surveys: The Total Design Method*. New York: John Wiley & Sons Inc.

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2008). *Internet, Mail, and Mixed-Mode Surveys: The Tailored Design Method* (3rd ed.). University of Michigan: Wiley.

Enderle, T., Münnich, R., & Bruch, C. (2011). *On the impact of response patterns on survey estimates from access panels.* Work in progress.

Esser, H. (1990). "Habits", "frames" und "rational choice". Die Reichweite von Theorien der rationalen Wahl (am Beispiel der Erklärung des Befragtenverhaltens). *Zeitschrift für Soziologie*, *19*, 213-247.

Frick, J. R., & Krell, K. (2010). *Measuring income in household panel surveys for Germany: A comparison of EU-SILC and SOEP.* SOEPpapers 265, DIW Berlin, The German Socio-Economic Panel (SOEP).

Goyder, J. (1987). *The Silent Minority: Nonrespondents on Sample Surveys*. Boulder, Colorado: Polity Press/Westview Press.

Groves, R. M. (2006). Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly*, *70*, 646-675.

[12] The lowest cluster information that is available is the information on the MC sampling area units. These units consist of about 8 households and there is probably only one interviewer who is responsible for several geographically close areas. Due to anonymisation in the AP it was not possible to use the regional information. However, interviewers may act in several areas. Besides, area effects may be linked to characteristics of the area.

Groves, R. M., Cialdini, R. B., & Couper, M. P. (1992). Understanding the decision to participate in a survey. *Public Opinion Quarterly*, *56*, 475-495.

Groves, R. M., & Couper, M. P. (1998). *Nonresponse in Household Interview Surveys*. New York: John Wiley & Sons.

Hippler, H. J. (1985). Schriftliche Umfragen bei repräsentativen Bevölkerungsstichproben oder: Wie erreicht man 78%. In M. Kasse & M. Küchler (Eds.), *Herausforderungen der Empirischen Sozialforschung* (p. 71-74). Mannheim: ZUMA.

Hoogendoorn, A. W., & Daalmans, J. (2009). Nonresponse in the recruitment of an internet panel based on probability sampling. *Survey Research Methods*, *3*, 69-72.

Horneffer, B., & Kuchler, B. (2008). Drei Jahre Panelerhebung EU-SILC: Erfahrungen und methodologische Weiterentwicklungen. *Wirtschaft und Statistik*, *Heft 8*, 650-661.

Körner, T., Meyer, I., Minkel, H., & Timm, U. (2005). Leben in Europa – Die neue Statistik über Einkommen und Lebensbedingungen. *Wirtschaft und Statistik*, *11*, 1137-1152.

Körner, T., & Nimmergut, A. (2004). Using an access panel as a sampling frame for voluntary household surveys. *Statistical Journal of the United Nations Economic Commission for Europe*, *21*, 33-52.

Körner, T., Nimmergut, A., Nökel, J., & Rohloff, S. (2006). Die Dauerstichprobe befragungsbereiter Haushalte. *Wirtschaft und Statistik*, *Heft 5*, 451-467.

Kühne-Finster, S., & Güllner, G. (2009). Ergebnisse der Interviewerbefragung im Mikrozensus. *Wirtschaft und Statistik*, 397-405.

Lotze, S., & Breiholz, H. (2002a). Zum neuen Erhebungsdesign des Mikrozensus. Teil 1. *Wirtschaft und Statistik*, *Heft 5*, 359-366.

Lotze, S., & Breiholz, H. (2002b). Zum neuen Erhebungsdesign des Mikrozensus. Teil 2. *Wirtschaft und Statistik*, *Heft 6*, 454-459.

Mejer, L. (2003). Harmonisation of socio-economic variables in EU statistics. In J. H. Hoffmeyer-Zlotnik & C. Wolf (Eds.), *Advances in Cross-National Comparison: A European Working Book for Demographic and Socio-Economic Variables* (p. 67-85). New York: Springer.

Nicoletti, C., & Buck, N. N. (2004). *Explaining interviewee contact and co-operation in the British and German household panels.* ISER Working Papers, Nr. 2004-06.

Nimmergut, A., Meyer, I., & Körner, T. (2004). *Haushalte Heute: Pilotstudie zur Umsetzbarkeit einer Dauerstichprobe befragungsbereiter Haushalte in der amtlichen Statistik 2001-2003 – Erhebung nach §7 BStatG*. Wiesbaden: Statistisches Bundesamt.

Rendtel, U. (1995). *Lebenslagen im Wandel: Panelausfälle und Panelrepräsentativität*. Frankfurt am Main: Campus Verlag.

Särndal, C., Swensson, B., & Wretman, J. (2003). *Model Assisted Survey Sampling*. New York: Springer.

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

Schräpler, J. P., Schupp, J., & Wagner, G. G. (2010). *Individual and neighborhood determinants of survey nonresponse: An analysis based on a new subsample of the German socio-economic panel (SOEP), microgeographic characteristics and survey-based interviewer characteristics.* SOEPpaper, Nr. 288.

Statistisches Bundesamt. (2008). *Gemeinschaftsstatistik über Einkommen und Lebensbedingungen: Leben in Europa 2006*. Wiesbaden: Statistisches Bundesamt.

van der Laan, J. (2009). *Representativity of the LISS panel.* Statistics Netherlands, discussion paper 09041.

Appendix

Table 1: Logit Model for the probability of successful recruitment in 2006. Estimated coefficients and standard errors. Positive coefficients imply higher probability of successful recruitment.

Term	Value	Estimate	Std.Err
Quarter	Q1	0.82	0.04
	Q2	0.14	0.04
	Q3	0.07	0.04
	Q4	0.00	–
Federal State	Brandenburg (BB)	0.93	0.08
	Bremen (BR)	-0.63	0.23
	Berlin (BN)	-0.09	0.10
	Baden-Wuerttemberg (BW)	0.99	0.07
	Bavaria (BN)	0.38	0.07
	Hesse (HE)	0.26	0.08
	Hamburg (HG)	0.19	0.11
	Lower-Saxony (NS)	0.27	0.08
	North Rhine-Westphalia (NW)	0.64	0.07
	Schleswig-Holstein (SH)	0.59	0.09
	Saarland (SL)	0.84	0.12
	Saxony (SN)	0.00	–
Children	No	-0.40	0.09
	Yes	0.00	–
Net Income	Na	-1.48	0.11
	0–900	-0.26	0.08
	900–1300	-0.35	0.07
	1300–2600	-0.14	0.05
	2600–3600	0.01	0.05
	3600+	0.00	–
Citizenship	Foreign	0.00	–
	German	0.58	0.06
Employment Status	Employed	-0.11	0.33
	Non-employed	0.00	–
	Retired	-0.08	0.07
	Self employed	-0.63	0.33
Number of adults*children	1	0.64	0.12
	1	No 0.22	0.12
	1	Yes 0.00	–
	2	0.25	0.08
	2	No 0.28	0.10
	2	Yes 0.00	–
	3	0.00	–
	3	No 0.00	–
	3	Yes 0.00	–
Sex	Female	0.00	–
	Male	0.23	0.13

Table 1: Continued.

Term	Value	Estimate	Std.Err
Age*sex	18–25	0.77	0.15
	18–25	Female 0.00	–
	18–25	Male -0.45	0.17
	25–35	0.61	0.14
	25–35	Female 0.00	–
	25–35	Male -0.28	0.15
	35–45	0.55	0.14
	35–45	Female 0.00	–
	35–45	Male -0.19	0.15
	45–55	0.60	0.14
	45–55	Female 0.00	–
	45–55	Male -0.28	0.15
	55–65	0.83	0.13
	55–65	Female 0.00	
	55–65	Male -0.39	0.15
	65–75	0.83	0.13
	65–75	Female 0.00	–
	65–75	Male -0.39	0.15
	75+	0.00	–
	75+	Female 0.00	–
	75+	Male 0.00	–
School degree	Basic secondary	0.27	0.13
	Intermediate secondary	0.50	0.13
	Na	0.34	0.44
	None	0.00	–
	Techn. school sec.	0.44	0.14
	Unknown	-1.00	0.43
	Upper secondary	0.42	0.13
Prof. training degree	Apprenticeship	-0.45	0.18
	None	0.00	–
	PhD	0.44	0.12
	Public Service	0.45	0.05
	School of Administration	0.49	0.14
	Techn/Crafts certicate	0.44	0.06
	UAS	0.46	0.07
	University	0.54	0.07
	Unknown	-1.25	0.51
	Voc. Training	0.18	0.34
	Voc. Training College	0.22	0.10
Marital status	Divorced	-1.12	0.10
	Married	0.13	0.08
	Separated	0.00	–
	Single	-0.06	0.08
	Widowed	-0.75	0.08
Weekly work hours	0	0.06	0.33
	1–20	0.39	0.07
	21–40	0.09	0.04
	40+	0.00	–
Area under ROC	0.65		
Hosmer-Lemeshow	19.6 on 8 df, p< 0:01		–