



Explaining Interviewee Contact and Co-Operation in the British and German Household Panels

Cheti Nicoletti and Nicholas N. Buck

**ISER Working Papers
Number 2004-6**

Institute for Social and Economic Research

The Institute for Social and Economic Research (ISER) specialises in the production and analysis of longitudinal data. ISER incorporates the following centres:

- ESRC Research Centre on Micro-social Change. Established in 1989 to identify, explain, model and forecast social change in Britain at the individual and household level, the Centre specialises in research using longitudinal data.
- ESRC UK Longitudinal Centre. This national resource centre was established in October 1999 to promote the use of longitudinal data and to develop a strategy for the future of large-scale longitudinal surveys. It was responsible for the British Household Panel Survey (BHPS) and for the ESRC's interest in the National Child Development Study and the 1970 British Cohort Study
- European Centre for Analysis in the Social Sciences. ECASS is an interdisciplinary research centre which hosts major research programmes and helps researchers from the EU gain access to longitudinal data and cross-national datasets from all over Europe.

The British Household Panel Survey is one of the main instruments for measuring social change in Britain. The BHPS comprises a nationally representative sample of around 5,500 households and over 10,000 individuals who are reinterviewed each year. The questionnaire includes a constant core of items accompanied by a variable component in order to provide for the collection of initial conditions data and to allow for the subsequent inclusion of emerging research and policy concerns.

Among the main projects in ISER's research programme are: the labour market and the division of domestic responsibilities; changes in families and households; modelling households' labour force behaviour; wealth, well-being and socio-economic structure; resource distribution in the household; and modelling techniques and survey methodology.

BHPS data provide the academic community, policymakers and private sector with a unique national resource and allow for comparative research with similar studies in Europe, the United States and Canada.

BHPS data are available from the Data Archive at the University of Essex
<http://www.data-archive.ac.uk>

Further information about the BHPS and other longitudinal surveys can be obtained by telephoning +44 (0) 1206 873543.

The support of both the Economic and Social Research Council (ESRC) and the University of Essex is gratefully acknowledged. The work reported in this paper is part of the scientific programme of the Institute for Social and Economic Research.

Acknowledgement: This work has been carried out for the Chintex Project financed by the European Commission under contract number IST-1999-11101. We thank Heather Laurie and Peter Lynn for their useful comments.

Readers wishing to cite this document are asked to use the following form of words:

Nicoletti, Cheti, and Buck Nicholas, N. (June 2004) 'Explaining Interviewee Contact And Co-Operation In The British And German Household Panels', *Working Papers of the Institute for Social and Economic Research*, paper 2004-6. Colchester: University of Essex.

For an on-line version of this working paper and others in the series, please visit the Institute's website at: <http://www.iser.essex.ac.uk/pubs/workpaps/>

Institute for Social and Economic Research
University of Essex
Wivenhoe Park
Colchester
Essex
CO4 3SQ UK
Telephone: +44 (0) 1206 872957
Fax: +44 (0) 1206 873151
E-mail: iser@essex.ac.uk
Website: <http://www.iser.essex.ac.uk>

© June 2004

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted, in any form, or by any means, mechanical, photocopying, recording or otherwise, without the prior permission of the Communications Manager, Institute for Social and Economic Research.

ABSTRACT

This paper investigates the factors affecting the contact and the co-operation of the interviewees in the British Household Panel Survey, in the German Socio Economic Panel Survey and in the European Community Household Panel for the UK and for Germany.

The differences in the contact and co-operation rates between surveys may reflect differences in the composition of the national populations and in data collection processes. The coexistence of two independent panel surveys in the UK and in Germany gives the opportunity to investigate if differentials in the contact and co-operation rates are due to differences in the data collection, personal and household characteristics and/or differences in their impact between countries or between surveys in a same country. If the differentials are explained mainly by differences in the characteristics, and above all data collection characteristics, then it is possible to reduce differentials just by harmonising the data collection. If instead differentials are attributable to heterogeneity in the response behaviour across countries or surveys in a same country, then the harmonisation of the data collection process has a more ambiguous effect.

We model the response at individual level as the occurrence of two sequential events: the contact and the co-operation. We explain the contact and the co-operation probabilities in wave t using a set of individual and household characteristics observed in wave $t-1$, and a set of variables characterising the collection process in wave t and $t-1$. Moreover, we investigate differences between surveys in the contact and co-operation probabilities by trying to disentangle the part due to differences in the distribution of the explanatory variables and the part due to differences in the model coefficients.

NON-TECHNICAL SUMMARY

The paper aims to answer the question whether the harmonization of survey design and fieldwork practice is useful in reducing differences in contact and response rates between surveys. To answer to that question we consider 4 panel surveys: two surveys designed to be harmonized surveys, the European Community Household Panel in Germany and the UK, and two surveys, the German Socio Economic Panel Survey and the British Household Panel Survey, designed in different ways, but where data has been harmonized 'ex-post'.

We analyse the differences in the contact and co-operation rates between the surveys and verify whether different fieldwork methods constitute the main explanation for those differences. To conduct this analysis we estimate models for the probability of contact and co-operation using a set of explanatory variables characterising individuals, household and fieldwork. Our results suggest that the set of fieldwork variables is more important than personal and household characteristics in explaining both the contact and the co-operation probability.

Unfortunately the question whether the fieldwork harmonization is useful to reduce differentials in response rates is not very easy to answer. This is because the ECHP input harmonization was not very successful. From our analysis it seems that differences in the contact and co-operation rates for surveys running in a same country are mainly due to differences in the distribution of explanatory variables. This implies that reducing the difference in the fieldwork should help to narrow the gap in the contact and response rates. However, when considering two surveys running in two different countries the main reason for differences in the contact and co-operation rates is due to differences in the coefficients. Therefore the impact of a fieldwork harmonization is difficult to predict in these circumstances.

1. Introduction

All household panel surveys are affected by nonresponse problems. In this paper we focus attention on unit nonresponse, which occurs when an eligible individual does not complete the questionnaire either because of a failed contact or because of a refusal or an inability to co-operate. In particular we consider attrition defined as the unit nonresponse in a given wave, of individuals responding in the previous wave. Focusing on attrition does not limit the validity of our analysis because the most frequent type of nonresponse pattern is attrition.¹

A first contribution of this paper is to unify suggestions coming from different papers on attrition in household panel surveys, in particular Hausman and Wise (1979), Ridder (1992), Fitzgerald *et al.* (1996), Lillard and Panis (1998), Campanelli *et al.* (1997, 1999, 2002) and Lepkowski and Couper (2002), to model the attrition process. As in Hausman and Wise (1979), Ridder (1992), Fitzgerald *et al.* (1996), Lillard and Panis (1998) we estimate a micro-econometric model using information on observed individual and household variables in wave $t-1$ to predict attrition in wave t . In addition, we also use variables characterising the data collection as in Campanelli *et al.* (1997, 1999, 2002) and Lepkowski and Couper (2002). The data collection variables are important to correctly specify the models, but also as potential instrumental variables, which can be used to correct for the attrition selection in regression model estimation. We specify separate models for the contact and for the co-operation probability given the contact, as suggested by Lepkowski and Couper (2002).

A second important contribution of this paper is trying to understand the causes of differences in the contact and in the co-operation probabilities across different surveys by disentangling two components: one due to differences in the population composition and survey characteristics distribution, and one due to differences in the impact of those characteristics on contact and co-operation. Disentangling the two components helps in answering to the following questions. Can we expect a similar contact and co-operation behaviour when running two different independent surveys in a same country, two similar surveys in two different countries or two different surveys in two different countries? And, after controlling for (conditioning on) personal, household and fieldwork characteristics, do the differences in the contact and co-operation rates between surveys vanish? This is important to establish to what extent an effort in harmonising the data collection process for two surveys running in two different countries really help to attenuate the differential in the response compared with running two independent surveys without any pre-harmonisation of the data collection process.

The paper is organised as follow. In the Section 2.1-2.3 we give a brief description of 4 different surveys used in the application, the British Household Panel Survey, the German Socio Economic Panel Survey and the European Community Household Panel for the UK and for Germany. In Section 3 we present some summary statistics of the personal, household and data collection variables and the rates of co-operation and contact for the 4 different surveys. In Section 4.1 we discuss the models used for the estimation of the co-operation and the contact probabilities and we show the empirical results. In Section 4.2 we analyse the differences in the rates of co-operation and contact by using the Goumulka and Stern (1990) decomposition. Finally, in the Section 5, we summarise the main findings of the paper.

2. Description of the Panel Surveys and Some Preliminary Analysis

In this section we describe the 4 surveys used in the empirical work. More precisely we consider the British Household Panel Survey (BHPS) and the European Community Household Panel (shortly Euro panel or ECHP) in the UK, and the German Socio-Economic Panel (GSOEP) and the Euro panel in Germany.

2.1 Brief Description of the BHPS

The British Household Panel Survey (BHPS) is an annual longitudinal survey carried out by the Institute for Social and Economic Research (ISER) at the University of Essex. The survey collects data on socio-economic characteristics at both the individual and the household level. It began in 1991 and, at the moment, there are eleven waves available. Further details are contained in Taylor *et al.* (2002).

¹ For an analysis of response patterns in the ECHP we refer to Nicoletti and Peracchi (2002).

The target population of the BHPS consists of all individuals resident in private households in England, Wales or Scotland South of the Caledonian Canal. Northern Ireland and Scotland North of the Caledonian Canal are excluded from the original sample. In its first wave, the BHPS covers more than 5,000 households and about 10,000 individuals. Northern Ireland began to participate later from the 7th wave. In the 7th wave, the original sample was integrated with an additional sample to provide data for the European Household Panel Survey (EHP), that also covers Northern Ireland. Finally, two extension samples for Scotland and Wales began in the 9th wave and a new sample for Northern Ireland was introduced in the 11th wave. In our empirical application we use just the original sample for the BHPS for the three-year period 1994-1996.

The sample in the first wave was drawn by two-stage sampling. The first stage consists of a stratified sample of postcode sectors, while the second stage is a systematic sample of postal addresses from each sector. All private households resident at the selected addresses are contacted and interviewed. Individuals belonging to these households are called original sample members (OSMs). OSMs include children born after the first wave to original sample members. OSMs are eligible for a personal interview if they are aged 16 or more.

All OSMs are followed and, if eligible, re-interviewed in successive waves. If they move to a new household, all members of the new household are also interviewed. Individuals that are not OSMs but become members of a household with at least one OSM are also interviewed. These new individuals are called temporary sample members (TSMs). TSMs are dropped from the panel if they stop belonging to a household with at least one OSM. The BHPS has a third category of respondents, the permanent sample members (PSMs), defined as TSMs with a strong relationship with an OSM and therefore followed even when they move to a household without OSMs. In particular, the natural parent of an OSM child is a PSM, no matter if he/she does not belong to the original sample.

All OSMs, TSMs and PSMs aged 16+ are interviewed even if they move to Scotland North of the Caledonian Canal or to an institution (except if the interview is impossible because the person is in prison, or is mentally or physically unable to respond). The mode of interview is face-to-face paper and pencil interview. If all attempts to contact a person for a face-to-face interview fail, then proxy or telephone interviews are used. From 1999 onward the face-to-face paper and pencil interviews (PAPI) have been replaced with computer assisted personal interviews (CAPI).

2.2 Brief Description of the EHP for the UK and Germany

The European Community Household Panel, shortly EHP or Euro panel, is a standardised multi-purpose annual longitudinal survey carried out for the 15 European countries belonging to the European Union (EU). It is centrally designed and co-ordinated by the Statistical Office of the European Communities (Eurostat). A more detailed description of the EHP can be found in Peracchi (2002) and in the Eurostat (2002) EHP-UDB manual.

The target population of the EHP consists of all individuals living in private households within the EU. In its first (1994) wave, the EHP covered about 60,000 households and 130,000 individuals aged 16+ in 12 countries of the EU (Belgium, Denmark, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain and the UK). Austria, Finland and Sweden began to participate later.

The EHP distinguishes between sample and non-sample persons. Sample persons are all individuals belonging to the national samples drawn from the target population in the first wave. Sample persons also include children, born after the first wave, that have at least one parent sample person. Non-sample persons are all other individuals.

Sample and non-sample persons may or not be eligible for interview in a given wave. Sample persons are eligible if they are aged 16 or older and belong to the target population (that is, they live in a private household within the EU). Non-sample persons are eligible if, in addition, they live in a household containing at least one sample person. There are therefore two main different causes of ineligibility: natural demographic events (being born less than 16 years ago, being dead) and being “out of scope” (that is, homeless, institutionalised or outside of EU). A sample person who is “out of scope” is “traced” and interviewed again if he/she returns to the target population.² Ineligible non-sample persons are not traced. A

² The exception is when a person is institutionalised or moves outside the EU and there is no information available for tracing.

household completely nonresponding for two consecutive waves is dropped from the sample as well as people with a definite refusal.

For budgetary reason the following rules were not respected in the UK, where only households with complete personal questionnaires in wave t were followed and interviewed again in the subsequent wave $(t+1)$. Some of the follow-up rules were not applied in Germany too. People non-contacted, physical incapable to fill the questionnaire or refusing to answer in wave 1 were not traced in wave 2 in Germany.³

The ECHP questionnaire is composed of several forms: a household register, a household questionnaire and a personal questionnaire.

The interviewing method recommended by Eurostat is face-to-face personal interviewing. Other interviewing methods have been used, however, such as self-completed by the respondent, telephone, and proxy (for personal interviews only) when all attempts to contact a person for a face-to-face interview failed. For the UK the method of interviewing most used has been the face-to-face computer assisted (CAPI), but proxy and telephone interviews have been also used. For Germany the main mode of interview has been the face-to-face paper and pencil interview (PAPI), but there have also been face-to-face computer assisted, self-completed by the respondent, by telephone and proxy interviews.

Sampling frames are not standardised across countries. In most countries, the sampling frame is either the population register or a master-sample created from the latest population census.

The organisation responsible for collecting the ECHP data for the UK for the 1994-1996 wave was the Office for the National Statistics (ONS). The organisation responsible for the ECHP data for Germany for the 1994-1996 wave was Statistisches Bundesamt.

2.3 Brief Description of the GSOEP

The German Socio Economic Household Panel Survey (GSOEP) is an annual longitudinal survey carried out by the German Institute for Economic Research (DIW) in Berlin. The survey collects data on socio-economic characteristics at both the individual and the household level⁴.

The original target population of the GSOEP consists of all individuals resident in private households in West-Germany in 1984, with a household head who does not belong to the following main foreign groups: Turkish, Yugoslavian, Italian, Greek and Spanish. A separate sample was drawn in 1984, in West Germany, for households with a head belonging to one of the above main foreign groups. Starting from the 1990 the GSOEP has been extended to the East-Germany. Furthermore, there are the following additional 3 samples:

1. a sample to represent recent immigrants starting in 1994,
2. a refreshment sample starting in 1998,
3. a new sample to represent individual resident in Germany in 1998.

The West-German sample covered about 4500 households in the first wave (1984), while the East-German sample covered about 2200 households in the first wave (1990). In our empirical application we use just the original samples for West and East Germany for the three-year period 1994-1996.

A multistage sampling was used for all the above samples with clusters corresponding to regions. A household questionnaire is collected for all households belonging to the sample, while a personal interview is carried out for all individuals aged 16 or more and belonging to the sampled households. The standard mode of interview is paper-and-pencil face-to-face (PAPI), but self-completed or telephone interviews are used when a face-to-face interview fails. From 1998 onward the computer assisted personal interview (CAPI) mode is also used.

The GSOEP distinguishes between sample and non-sample units as in the ECHP. The following rules are different from the ones adopted by the ECHP because people moving to a foreign EU country are not followed and because, from 1990, nonsample people are followed even if they move into a household without sample people.

³ For more details about the exception to the follow-up rules in the ECHP we refer to Eurostat (1997).

⁴ For a more detailed description of the GSOEP, we refer to Haisekn-DeNew and Frick (2002).

3. Main differences in the observed variables across the 4 surveys

In this section we show some summary statistics of variables characterising the data collection process to emphasise potential differences in the way of conducting the 4 surveys. Moreover we compare the averages of some personal and household characteristics, which may explain the differences in the co-operation and the contact rates observed across surveys. Finally we define and compute the contact and the co-operation rates.

For comparability reasons we decided to use the data published in the ECHP-UDB 2002 first 3 waves for the BHPS, the GSOEP, the ECHP-UK (ECHP for the UK) and the ECHP-D (ECHP for Germany). We focus only on the waves 1994-1996 because the ECHP-UK and the ECHP-D are available only for the first three waves of the ECHP. We use the derived versions of the BHPS and the GSOEP published in the ECHP-UDB because this should solve the problems of comparability in terms of variables and sample definitions.⁵ In addition, we use some variables available in the ECHP producer database, in particular the interviewer identifier.

The subsamples used are given by all eligible individuals (we exclude all individuals aged less than 16 or out of the target population) in wave t (1995 or 1996) and unit responding in the wave $t-1$ (1994 or 1995). When speaking of co-operation⁶ probability (rate), we mean the probability of an individual being respondent (giving back a personal questionnaire) given contact and eligibility; whereas when speaking of contact⁷ probability, we mean the probability of an individual being contacted given eligibility. We call instead the probability to be respondent for all individuals, either contacted or not contacted, the response probability, which is given by the product of the co-operation probability and the contact probability. Both the contact and the co-operation are defined at the individual level.

In Table 1 we report the co-operation, the contact and the response rates separately for the 4 surveys and for waves 1995 and 1996. Looking to the contact rates in Table 1, it is evident that something strange happened in the ECHP-UK. The contact rate in this survey is very low and this is due to the decision to follow up only the households with complete personal questionnaires. Comparing the GSOEP and the BHPS rates it seems that the BHPS present high co-operation rates and low contact rates while the opposite is true for the GSOEP. It is quite difficult to understand if these differences are due to differences:

- in defining co-operation and contact,
- in the data collection, personal or household variables,
- in the ease of contact and in the propensity to respond everything else equal, or
- in the unobserved variables (unobserved heterogeneity).

In the Section 4.2 we will investigate the differences between the co-operation and the contact rates observed for different surveys.

In Table 2 we report the averages of fieldwork variables (top part) and of personal and household characteristics (bottom part).

The fieldwork variables we consider refer to interviewed persons, interviewed households or interviewers.

The fieldwork variables at the personal level are:

- the duration in minutes of the personal interview in wave $t-1$ (Min personal interview),
- 4 dummy variables indicating the mode of interview, which can be face to face computer assisted personal interview, CAPI, self-administered, proxy and telephone interview and face to face paper and pencil interview mode, PAPI.

The fieldwork variables at household level are:

- the duration in minutes of the household interview in wave $t-1$ (Min household interview),

⁵ Nevertheless for the definitions of contact and co-operation in the BHPS we used the BHPS dataset instead of the derived version published in the ECHP.

⁶ Co-operation fails when people refuse to respond, fail to return a self-completed questionnaire, are unable to respond because of physical or language problems or the interview is not completed for some other unknown reasons.

⁷ Contact fails when it is impossible to locate the address, the person(s) is(are) temporary away or the person(s) is(are) absent at all callbacks for unknown reasons.

- the household item nonresponse index in $t-1$,⁸
- the number of visits to the household in $t-1$
- a dummy indicating the use of the same interviewer for the same interviewee in t and $t-1$.

The fieldwork variables at the interviewer level are the number of individuals attributed to each interviewer in t (work load per interviewer) and the averages by interviewer of:

- the contact dummy in wave $t-1$ for people allocated to an interviewer, who was present also in wave $t-1$ (contact rate by interviewer in $t-1$),
- the co-operation dummy in wave $t-1$ for people allocated to an interviewer, who was present also in wave $t-1$ (co-operation rate by interviewer in $t-1$),⁹
- the household item nonresponse by interviewer in wave t .

Finally we report the length of the fieldwork measured as the duration in years between the first and the last interview collected.

The average duration of personal interviews is longer than the household interview average duration in all surveys. The duration of personal interviews is on average 30 min, except in the ECHP-UK where it has an average duration of 20 min, whereas the average duration of household interviews is between 11 and 22 min. The use of the same interviewer across wave is obviously more frequent for the two oldest panel surveys, the GSOEP and the BHPS. The most used interview mode in the ECHP-UK is the face to face computer assisted personal interview (CAPI), while in all other surveys it is the face to face pen and pencil interview (PAPI). Telephone, proxy and self-administered interviews should be adopted only when the face to face interview is not possible. The not face to face interview modes are not very frequent except in the GSOEP, where 40% of the interviews are self-administered, and in the ECHP-D, where 12% of the interviews are proxy.

The household item nonresponse index observed in $t-1$ does not vary much across surveys and about 5% of the total household income is imputed. The average number of visits to the household is about 2.5 in ECHP-D and in GSOEP, it is instead about 3 in ECHP-UK and in BHPS.

The interviewer workload differs a lot across surveys. The most burdensome workload is observed for the ECHP-UK and it is likely to be due to the limitations in the budget, which affected this survey. The average contact and co-operation rates by interviewer in $t-1$ are more or less equal to the observed average contact and co-operation rates (see Table 2), except for the contact rate for the ECHP-UK and for the co-operation rate in the BHPS. These exceptions may be due to heterogeneity of the contact and the co-operation probabilities across interviewers.

The length of the fieldwork varies between 6 months for the ECHP-UK and 11 months for the GSOEP.

In the bottom part of Table 1 we report the averages for the following personal and household variables:

- a dummy indicating people aged 35 or less and another dummy for people aged 60 or more,
- a sex dummy,
- a dummy indicating individual living without a spouse (single dummy),
- the number of children,
- the number of adults,
- a dummy for homeownership,
- two dummies for the labour status position, one to indicate unemployed people and one to indicate inactive people,

⁸ The household income is computed by summing up the personal incomes reported in the personal questionnaires of the individual belonging to the same household. The personal income is in its turn computed by summing up the different income components declared by an individual in the personal questionnaire. Therefore the household income is partially observed when there are item nonresponse to the income variables in the personal questionnaires or individuals unit nonresponding. In that case the missing components of the household income are imputed so that the total final household income is given by the sum of the reported income components and of the imputed components. The item nonresponse index is an imputation index given by the ratio between the sum of the imputed income components over the total final household income.

⁹ For interviewers present in t and not in $(t-1)$ the average of variables in $(t-1)$ is replaced with the mean average observed for the interviewers present in $(t-1)$.

- two dummies for the level of education (first level for an education less than second stage or secondary school and third level is for an education level higher than second stage or secondary school),
- the number of years since last change of address.

The average variables that present the biggest variation across surveys are the percentage of people homeowner and the level of education dummies. It is evident that the percentage of homeowners is much higher in the UK than in Germany.

4. A Micro-Econometric Analysis of the Response Processes

This section is organised in two subsections. The first part describes the estimation results of the micro-econometric models used to explain the probabilities of contact and co-operation. The second part presents instead a decomposition of the differences in rates of contact and co-operation between surveys by disentangling between two components: one due to changes in the explanatory variables distributions and one due to changes in the impact of these variables between surveys.

4.1 Estimation of the contact and the co-operation processes

We model the response at individual level as the occurrence of two sequential events: the contact and the co-operation. This is the same procedure followed by several authors, among them, Campanelli *et al.* (1997), Groves and Couper (1998) and Lepkowski and Couper (2002). We specify two separate probit models for the probability of contact and for the probability of co-operation given the contact in the current wave, say wave t . We use as explanatory variables a set of individual and household characteristics observed in wave $t-1$, and a set of variables characterising the collection process in wave t and $t-1$.

We estimate two probit models, one for the contact and one for the co-operation, for the 4 surveys described above, BHPS, ECHP-UK, ECHP-D and GSOEP. Conditioning on a set of explanatory variables, x_i , the probit model for the contact is given by

$$\Pr(r_i = 1 | x_i; \alpha) = \Phi(x_i' \alpha),$$

where r_i is a binary variable taking value 1 if the individual i has successfully been contacted and 0 otherwise, Φ is the cumulative standardised Gaussian distribution function and α is a vector of coefficients.

Conditioning instead on a set of explanatory variables, z_i , the probit model for the co-operation given the contact is given by

$$\Pr(s_i = 1 | z_i, r_i = 1; \beta) = \Phi(z_i' \beta),$$

where s_i is a binary variable taking value 1 if the individual i has given back a filled questionnaire and 0 otherwise, Φ is the cumulative standardised Gaussian distribution function and β is a vector of coefficients.

Notice that our probit models do not consider the random effects for the interviewer and/or the area effects. Nevertheless, the two probit models are still consistently estimated if we assume that the random effects in the two models are uncorrelated, and identically and independently distributed among individuals with a Gaussian distribution with zero mean and constant variance for both contact and cooperation models (see Robinson 1982 and Maddala 1987). The independence of the random effects in the contact and in the co-operation models is a sensible assumption as Campanelli and *et al* (1997) showed. Campanelli *et al* (1997) specified a separate model for the contact and the co-operation given the contact considering two random effects one associated with areas and another one associated with interviewers. They found out that only the area effect is relevant in the contact model whereas only the interviewer effect is important for the co-operation model. These results support our assumption of independence between random effects in the contact and the co-operation model.

We have also estimated separate random effects probit models for contact and cooperation to evaluate the importance of the random effects. More specifically we have considered interviewer random

effects by using the interviewer identifiers. We find that the percentage of the error variance explained by random effects is always below 30% except for the contact model using the GSOEP where it reaches 50%. We do not report the results for the random effect probit models, which coefficients are directly comparable, after a proper rescaling with the ones of simple probit models, see Arulampalam (1999) for details.

Some of the variables describing the data collection process change across surveys but they have no variation or very little variation within the same survey, so we are compelled to exclude some of the variables from the probit models. To allow a comparison of the effect of the explanatory variables across surveys we maintain the same specification of the model for the 4 surveys. To keep the specification of the models as parsimonious as possible, we eliminate some of the potential explanatory variables whose coefficients are not significantly different from zero.

We use, as explanatory variables, data collection characteristics at household level and at interviewer (area) level. At household level we use the number of visits to the household in wave $t-1$, the dummy indicating the use of the same interviewer for the contact of the household in $t-1$ and t , say the interviewer continuity dummy, and the household item nonresponse index in wave $t-1$. At interviewer level we use instead the interviewer workload in wave t (number of households per interviewer), the average household item nonresponse index in wave t , the average contact rate in wave $t-1$ and the average co-operation rate in wave $t-1$. We can think the data collection characteristics at interviewer level as area variables describing differences across areas, which may explain differences in contact and co-operation.

We consider also personal and households variables, in particular two age group dummies (one for people aged 35 or less, and one for people age 60 or more), the sex dummy, the dummy indicating individual living without a spouse, the number of children, the number of adults, two dummy for the labour status (unemployed and inactive) and the dummy indicating the home ownership.

Finally we take account of possible changes between waves 1995 and 1996 by introducing a wave dummy variable.

The probit models estimates are presented in Tables 7-14. The explanatory variables in the contact model relevant for all surveys are:

- the dummy indicating the home tenancy, which has a negative effect on the contact probability,
- the age dummies, which indicate an easier contact for people over 35,
- the dummy for the interviewer continuity,¹⁰ which has a positive coefficient,
- the dummy for people living without a spouse, who are more difficult to be contacted,
- the wave dummy.

The explanatory variables in the co-operation model relevant for all surveys are:

- the number of adults in the household, which is negatively related to the co-operation probability,
- the dummy for people inactive, who tend to be more co-operative,
- the age dummies, which indicate that the co-operation is more difficult for people younger than 35 and older than 60 except in the GSOEP where there is no difference in co-operation for people over 60,
- the dummy for the interviewer continuity, which has a positive effect on co-operation,
- the number of visits to the household in the previous wave, which is negatively related to the co-operation probability.
- the workload of the interviewer, which is negatively related to the response probability except in the ECHP-D where it is not significant,
- the dummy indicating people living without a spouse, which has a negative effect,
- the wave dummy.

In Tables 3 and 4 we present some tests to verify the significance of 4 different factors affecting the co-operation and the contact:

- the data collection variables,
- the personal/household variables,

¹⁰ This dummy variable may be endogenous. As suggested in Campanelli and O'Muircheartaigh (2002), the endogeneity problem is likely to be ascribed to an interviewer nonrandom attrition, which has a higher probability in areas where the co-operation and the contact probabilities are lower. Contrasting results have been found about the impact of the interviewer effect from previous studies, Campanelli *et al* (1997, 1999, 2002) and Laurie *et al* (1999). We have tried to estimate the models without the interviewer continuity dummy and the coefficients do not change significantly, so that we infer that the endogeneity does not affect enormously our results. Nevertheless, the coefficient of the interviewer continuity dummy should be interpreted with caution.

- the wave dummy, which is supposed to capture individual invariant unobserved heterogeneity across waves.

It seems that none of these factors can be disregarded in explaining the co-operation and the contact probabilities except the wave dummy, which is not always significant at 1% level. In both the contact and the co-operation models the data collection variables are the set of characteristics which the higher significance level. Checking the structural stability of the models across surveys with Chow tests, we strongly reject the assumption of equal coefficients. Moreover it seems that the coefficients of the data collection characteristics change more across surveys than the coefficients of the personal and household characteristics. Even when comparing two surveys running in a same country the coefficients are significantly different.

In the next section we try to investigate differences between surveys in the co-operation and in the contact probabilities by trying to disentangle the part due to differences in the models coefficients and the part due to differences of the explanatory variables distributions.

4.2. Decomposition of the differences in the contact and the co-operation rates.

In this section we decompose the differences in the observed contact and co-operation rates between pair of surveys into two additional components: (1) a component due to differences in the distribution of the personal, household, data collection variables and wave dummies, (2) a component associated with differences in the impact of the variables on the propensity to respond and to be contacted.¹¹

Let us consider again the probit model, defined in last section, for the binary variable r_i taking value 1 if the individual i has been contacted and 0 otherwise, that is

$$\Pr(r_i = 1 | x_i; \alpha) = \Phi(x_i' \alpha).$$

Let us assume that the coefficients α vary across surveys. Let α_0 and α_1 be two vectors of coefficients associated with two different surveys, say 0 and 1, $f_0(x_i)$ and $f_1(x_i)$ be the density/probability distribution functions for the explanatory variables in surveys 0 and 1, and d_i^0 and d_i^1 be two dummy variables indicating if an individual i belongs to the survey 0 and to the survey 1. Then the marginal probability of contact for the survey 0 is given by

$$\Pr(r_i = 1 | d_i^0 = 1) = \int \Phi(x_i' \alpha_0) f_0(x_i) dx_i,$$

and the analogous probability for the survey 1 is given by

$$\Pr(r_i = 1 | d_i^1 = 1) = \int \Phi(x_i' \alpha_1) f_1(x_i) dx_i.$$

The difference between the two above marginal contact probabilities can be decomposed in the following way:

$$\Pr(r_i = 1 | d_i^1 = 1) - \Pr(r_i = 1 | d_i^0 = 1) = \int \Phi(x_i' \alpha_0) (f_1(x_i) - f_0(x_i)) dx_i + \int f_1(x_i) (\Phi(x_i' \alpha_1) - \Phi(x_i' \alpha_0)) dx_i.$$

The last equation shows how the difference between the contact probabilities, observed for two different surveys, can be decomposed into two components. The first component, given by the first addend in the right hand side of the last equation, represents the effects of changes in the distribution of the explanatory variables between the two surveys; whereas the second component, given by the second addend in the right hand side, represents the effects of changes in the coefficients. Notice that the decomposition results would not change if we considered random effects probit models. Assuming that random effects are normally distributed for all surveys, differences between contact and cooperation rates are not at all explained by differences in random effects distribution.

¹¹ See Goumulka and Stern (1990) for a first example of application of the decomposition.

The marginal contact probability for a specific survey, say 0 (or 1), can be estimated just by replacing the coefficients α_0 (α_1) with their estimates and by considering the sampling average instead of the integral in the following way

$$\hat{p}_0 = \sum_i d_i^0 \Phi(x_i' \hat{\alpha}_0) / \sum_i d_i^0, \quad \hat{p}_1 = \sum_i d_i^1 \Phi(x_i' \hat{\alpha}_1) / \sum_i d_i^1,$$

where the \sum_i is over all individuals belonging to the surveys 0 or 1.

The two terms of the decomposition can be instead estimated as follow:

$$\hat{p}_1 - \hat{p}_0 = \left(\frac{\sum_i d_i^1 \Phi(x_i' \alpha_0)}{\sum_i d_i^1} - \frac{\sum_i d_i^0 \Phi(x_i' \alpha_0)}{\sum_i d_i^0} \right) + \left(\frac{\sum_i d_i^1 \Phi(x_i' \alpha_1)}{\sum_i d_i^1} - \frac{\sum_i d_i^1 \Phi(x_i' \alpha_0)}{\sum_i d_i^1} \right).$$

The decomposition can be performed obviously for both the contact and the co-operation probabilities. We present the results of the estimation of the two components in Table 5 and in Table 6, where we report also a Wald test to verify if each component is significantly different from zero.¹² In particular we look at 4 differences in the contact and in the co-operation probabilities: between the ECHP-UK and the BHPS, between the ECHP-D and the GSOEP, between the ECHP-UK and the ECHP-D, and between the BHPS and the GSOEP.

The difference between contact and co-operation probabilities is almost exclusively due to changes in the coefficients when considering surveys running in two different countries. This means that even if the explanatory variables distribution were equal between two surveys running in two different countries, the contact and the co-operation rates would not be equal because of a different impact of the variables. In other words the ease of contact and the propensity to co-operate, every explanatory variable being equal, are different across surveys running in different countries. When instead we compare two different surveys running in a same country (this is the case when comparing ECHP-UK and BHPS in the UK or ECHP-D and GSOEP in Germany), the difference between contact and co-operation probabilities is almost exclusively due to changes in the explanatory variables.¹³ Changes in the fieldwork characteristics (workload, performance by interviewer measured by the percentages of successful contact and co-operation, average item nonresponse index by interviewer, personal item nonresponse index in previous wave, interviewer continuity dummy and number of visits) are likely to be the main explanation for differences in the contact rates between ECHP-UK and BHPS and between ECHP-D and GSOEP and differences in the response rates between ECHP-UK and BHPS.

We computed the same type of decomposition by excluding the fieldwork variables. Notice that splitting the set of variables x into two subvectors, x_1 for the personal and household characteristics and x_2 for the fieldwork variables, we can rewrite the difference between the probability of contact (or co-operation) between two surveys as the sum of three addends as follows:

$$\Pr(r_i = 1 | d_i^1 = 1) - \Pr(r_i = 1 | d_i^0 = 1) = \int \Phi(x_i' \alpha_0) (f_1(x_{2,i} | x_{1,i}) - f_0(x_{2,i} | x_{1,i})) f_0(x_{1,i}) dx_i + \int \Phi(x_i' \alpha_0) (f_1(x_{1,i}) - f_0(x_{1,i})) f_1(x_{2,i} | x_{1,i}) dx_i + \int f_1(x_i) (\Phi(x_i' \alpha_1) - \Phi(x_i' \alpha_0)) dx_i.$$

If the distribution of the fieldwork characteristics given the personal/household variables is the same between survey 0 and survey 1, i.e. if $f_1(x_{2,i} | x_{1,i}) = f_0(x_{2,i} | x_{1,i})$, then the first addend of the above sum vanishes. Let $\Phi(x_i' \alpha_{10})$ be a probit model for the contact (or the co-operation) omitting the fieldwork variables and let α_{10} be the set of parameter associated to the personal and household variables for the survey 0. If $\int \Phi(x_i' \alpha_0) f_1(x_{2,i} | x_{1,i}) dx_{2i} \approx \Phi(x_i' \alpha_{10})$, then the decomposition considering and not considering the fieldwork variables should produces similar results. This assumption is rejected when comparing co-

¹² Details on the computation of the two components and their standard errors can be required to the authors.

¹³ The only exception is for the co-operation model when comparing ECHP-D and GSOEP.

operation and contact between ECHP-UK and BHPS and when comparing contact between ECHP-D and GSOEP. We conclude therefore that the distribution of the fieldwork characteristics given the personal/household variables changes a lot between ECHP-UK and BHPS for both contact and response and between ECHP-D and GSOEP for the contact.¹⁴

When comparing the harmonized ECHP between the UK and Germany the difference in the contact and response rates explained by differences in the coefficients and by differences in the variables is higher than expected. We would expected a smaller difference with respect when comparing two non-harmonized surveys running in two different countries (GSOEP and BHPS), but our results do not support this expectation. It is obvious that this unexpected result may be due to an inadequate harmonization of the fieldwork for the ECHP-UK where contact is exceptionally low with respect to all other EU countries and co-operation is exceptionally high. It seems that the distribution of fieldwork characteristics, conditional on the personal/household variables, are slightly more similar when comparing BHPS and GSOEP than when comparing ECHP-UK and ECHP-D. We would infer from this that the fieldwork pre-harmonization in the ECHP was not very successful.

In conclusion, when we compare two panel surveys running in parallel for a same country co-operation and contact probabilities are more similar, any explanatory variable being equal. When instead we compare two surveys running in different countries the change in the coefficients is higher and it is the main explanation for differences in the contact and in the response rates. Harmonization of the fieldwork can potentially help in reducing the difference in the contact and in the response rates, but this is not the case for the harmonization of the ECHP in the UK and in Germany. Notice that even if the fieldwork harmonization in the ECHP had been successful, difference in contact and co-operation rates due to coefficients' changes between countries would have been unavoidable.

5. Conclusions

In this paper we have tried to explain differences in the contact and the co-operation processes for four household panel surveys, the ECHP-D, the ECHP-UK, the GSOEP and the BHPS. A very clear-cut result of the paper is that the fieldwork harmonization of the ECHP in the UK and in Germany was not very successful. Unfortunately, for budgetary reasons it has not been possible for the UK to adopt all the fieldwork rules suggested by Eurostat to the countries taking part to ECHP.

A second clear-cut result of the paper is that there are at least two important factors affecting the co-operation and the contact probabilities. These are the data collection characteristics and the personal/household variables. Furthermore, it seems that the set of data collection variables is more important than personal and household characteristics in explaining both the contact and the co-operation probability.

When comparing two surveys in two different countries the differences in the co-operation and contact rates are mainly due to changes in the impact of the explanatory variables (changes in the coefficients) rather than to changes in the variables distribution. Changes in the distribution of the explanatory variables are instead the main reason for differences in the rates when comparing two surveys running in parallel in a same country.

The question whether the fieldwork harmonization is useful to reduce the contact and the response rates gap between surveys is not easy to be answered. From our analysis it seems that differences in the contact and co-operation rates for surveys running in a same country are mainly due to differences in the explanatory variables distribution. This implies that reducing the difference in the data collection variables should help to narrow the gap in the contact and response rates. However, when considering two surveys running in two different countries the main reason for differences in the contact and co-operation rates is due to differences in the coefficients. The impact of a fieldwork harmonization is difficult to predict in these circumstances.

¹⁴ Let notice that if we think to x_2 were an area (interviewer) random effect, then the decomposition results with and without considering the random effects would be the same. The only difference is that the computation of a proper decomposition with random effect requires to properly rescaling the coefficients for each surveys to take account of a different random effect variance across surveys.

References

- Arulampalam W. (1999), A note on estimate coefficients in random effects probit models, *Oxford Bulletin of Economics and Statistics*, 61, 597--603.
- Campanelli, P., O'Muircheartaigh, C. (1999), Interviewers, Interviewer Continuity, and Panel Survey Nonresponse, *Quality & Quantity*, 33, 1, 59-76.
- Campanelli, P., O'Muircheartaigh, C. (2002), The importance of experimental control in testing the impact of interviewer continuity on panel survey nonresponse, *Quality & Quantity*, 36, 129-144.
- Campanelli, P., Sturgis, P., and Purdon, S. (1997), *Can You Hear Me Knocking: An Investigation into the Impact of Interviewers on Survey Response Rates*, Social and Community Planning Research, London.
- Goumulka J., Stern N. (1990), The employment of married women in the United Kingdom 1970-83, *Economica*, 57, 171-99.
- Eurostat (2002), *ECHP UDB Manual*, Eurostat, Luxembourg.
- Fitzgerald J., Gottschalk P., Moffitt R. (1996), An analysis of sample attrition in panel data: the Michigan Panel Study of Income Dynamics, *The Journal of Human Resources*, 33, 251-299.
- Groves R.M., Couper M.P. (1998), *Nonresponse in Household Interview Surveys*, Wiley, New York.
- Haisekn-DeNew J.P., Frick J. (2002), *Desktop Companion to the German Socio Economic Panel Study*, DIW, Berlin.
- Hausman J.A. and Wise D.A. (1979), Attrition bias in experimental and panel data: the Gary-income maintenance experiment, *Econometrica*, 47, 271-320.
- Laurie H., Smith R., Scott L. (1999), Strategies for reducing nonresponse in a longitudinal panel survey, *Journal of Official Statistics*, 15, 2, 269-182.
- Maddala G.S. (1987), Limited dependent variable models using panel data, *Journal of Human Resources*, 22, 307-337.
- Lepkowski J.M., Couper M.P. (2002), Nonresponse in the second wave of longitudinal household surveys', in R.M. Groves, D.A. Dillman, J.L. Eltinge, and R.J.A. Little (eds.), *Survey Nonresponse*, Wiley, New York.
- Lillard L.A., Panis C.W.A. (1998), Panel attrition from the Panel Study of Income Dynamics, *The Journal of Human Resources*, 33, 439-547.
- Peracchi F. (2002), The European Community Household Panel: A review, *Empirical Economics*, 27, 63-90.
- Ridder G. (1992), An empirical evaluation of some models for non-random attrition in panel data, *Structural Change and Economic Dynamics*, 3, 2, 337-355.
- Robinson P.M. (1982), On the asymptotic properties of estimators of models containing limited dependent variables, *Econometrica*, 50, 27-42.
- Taylor M.F. (ed). with John Brice, Nick Buck and Elaine Prentice-Lane (2002), *British Household Panel Survey User Manual Volume A: Introduction, Technical Report and Appendices*, University of Essex, Colchester.

Tab. 1 Co-operation, contact and response rates

	ECHP-D	ECHP-UK	GSOEP	BHPS
Co-operation rate wave 95	0.951	0.993	0.946	0.965
Co-operation rate wave 96	0.982	0.972	0.946	0.983
Contact rate wave 95	0.982	0.773	0.995	0.982
Contact rate wave 96	0.980	0.819	0.991	0.992
Response rate 95	0.934	0.768	0.941	0.947
Response rate 96	0.962	0.796	0.937	0.975

Tab. 2 Averages of fieldwork, personal and household characteristics

	ECHP-D	ECHP-UK	GSOEP	BHPS	
<i>Fieldwork variables</i>					
<i>Fieldwork variables at personal level</i>					
Min personal interview in t-1		30	22	35	35
Same interviewer used in t and t-1		0.447	0.595	0.697	0.793
Dummy telephone interview in t-1		0.005	0.009	0.002	0.012
Dummy CAPI in t-1		0.044	0.954	0.000	0.000
Dummy proxy in t-1		0.123	0.079	0.002	0.000
Dummy self-administered in t-1		0.032	0.000	0.399	0.000
<i>Fieldwork variables at household level</i>					
Min household interview in t-1		22	16	18	11
Household item nonresponse index in t-1		0.050	0.043	0.050	0.060
No. of visits in t-1		2.422	3.355	2.423	2.865
<i>Fieldwork variables at interviewer level</i>					
Contact rate by interviewer in t-1		0.989	0.981	0.985	0.955
Co-operation rate by interviewer in t-1		0.955	0.989	0.955	0.926
Household item nonresponse by interviewer in t		0.052	0.043	0.046	0.060
Interviewer work load in t (median)		21	52	45	38
<i>Length of the fieldwork in t</i>		.	0.468	0.917	0.615
<i>Personal and household characteristics</i>					
Dummy age 35 or less		0.251	0.319	0.382	0.347
Dummy age 60 or more		0.256	0.281	0.206	0.240
Sex dummy		0.465	0.521	0.521	0.540
Single dummy		0.236	0.339	0.311	0.355
No. of children		0.611	0.643	0.625	0.635
No. of adults		2.165	2.198	2.371	2.207
Dummy homeownership		0.485	0.727	0.408	0.802
Dummy unemployed		0.046	0.049	0.067	0.047
Dummy inactive		0.375	0.410	0.361	0.399
Dummy first level education		0.281	0.478	0.206	0.446
Dummy third level education		0.234	0.199	0.220	0.363
Time since last change of address		11.3	10.0	10.4	9.5

Tab. 3 Significance tests for the co-operation probability model

Significance test	BHPS	ECHP-UK	ECHP-D	GSOEP
Data collection variables, Wald test $\chi^2(6)$	683.6	63.3	105.5	549.4
p-value	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Personal and household variables, Wald test $\chi^2(8)$	39.4	83.6	47.7	125.1
p-value	(0.0000)	(0.0000)	(0.0000)	(0.0000)
1996 wave dummy	4.9	86.2	4.7	4.8
p-value	(0.0273)	(0.0000)	(0.0302)	(0.0285)

Tab. 4 Significance tests for the contact probability model¹⁵

Significance test	BHPS	ECHP-UK	ECHP-D	GSOEP
Data collection variables, Wald test $\chi^2(5)$	489.9	1151.9	442.5	267.5
p-value	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Personal and household variables, Wald test $\chi^2(8)$	99.9	340.1	138.4	117.8
p-value	(0.0000)	(0.0000)	(0.0000)	(0.0000)
1996 wave dummy	13.3	13.4	349.6	3.9
p-value	(0.0003)	(0.0002)	(0.0000)	(0.0478)

Tab. 5 Decomposition of differences in the co-operation probabilities

Comparing Against	ECHP-UK BHPS	ECHP-D GSOEP	ECHP-UK ECHP-D	BHPS GSOEP
Difference in the co-operation probability	0.01188	0.02015	0.01994	0.02821
Percentage	100%	100%	100%	100%
Difference caused by changes in the variables distribution	0.02509	0.00449	-0.00525	0.00369
Percentage	211%	22%	-26%	13%
Wald significance test	28.6	4.1	-0.8	4.5
p-value	0.0000	0.0001	0.4246	0.0000
Difference caused by changes in the coefficients	-0.01321	0.01566	0.02519	0.02452
Percentage	-111%	78%	126%	87%
Wald significance test	-14.4	7.1	3.7	13.8
p-value	0.0000	0.0000	0.0002	0.0000

¹⁵ The sample of the ECHP-UK does not include the households with unit nonresponses in wave t-1.

Tab. 6 Decomposition of differences in the contact probabilities¹⁶

Comparing Against	ECHP-UK BHPS	ECHP-D GSOEP	ECHP-UK ECHP-D	BHPS GSOEP
Difference in the contact probability	-0.15969	-0.01188	-0.15302	-0.00521
Percentage	100%	100%	100%	100%
Difference caused by changes in the variables distribution	-0.14655	-0.00848	-0.01194	0.00222
Percentage	92%	71%	8%	-43%
Wald significance test	-207.3	-4.6	-4.4	6.2
p-value	0.0000	0.0000	0.0000	0.0000
Difference caused by changes in the coefficients	-0.01314	-0.0034	-0.14108	-0.00743
Percentage	8%	29%	92%	143%
Wald significance test	-16.1	-1.5	-44.6	-8.7
p-value	0.0000	0.1387	0.0000	0.0000

Tab. 7 Co-operation probit model for the ECHP-UK without households with unit nonresponse in t-1

Variables	Coefficients	Std. Err.	Test	p-value
Wave 1996	-0.53754	0.05790	-9.28	0.000
Workload	-0.00393	0.00058	-6.74	0.000
Item nonresponse by interviewer	1.64016	0.87252	1.88	0.060
Co-operation rate by interviewer	-1.43787	1.74026	-0.83	0.409
Age 35 or less	-0.18150	0.06801	-2.67	0.008
Age 60 or more	-0.55059	0.08538	-6.45	0.000
Female	0.02228	0.05565	0.40	0.689
Living without a spouse	-0.07275	0.06298	-1.16	0.248
No. of children	-0.03734	0.02915	-1.28	0.200
No. of adults	-0.17194	0.02978	-5.77	0.000
Unemployed	0.15203	0.13298	1.14	0.253
Inactive	0.41314	0.07328	5.64	0.000
No. of visits	-0.01891	0.01203	-1.57	0.116
Same interviewer	0.23181	0.05985	3.87	0.000
Item nonresponse	-0.34440	0.18707	-1.84	0.066
Constant	2.13094	0.27444	7.76	0.000
Wald joint significance test	302.22	0.00000	No. obs.	15057

¹⁶ The sample of the ECHP-UK does not include the households with unit nonresponses in wave t-1.

Tab. 8 Contact probit model for the ECHP-UK without households with unit nonresponse in t-1

Variables	Coefficients	Std. Err.	Test	p-value
Wave 1996	-0.23134	0.06311	-3.67	0.000
Workload	-0.00324	0.00012	-26.72	0.000
Item nonresponse by interviewer	1.33430	0.88217	1.51	0.130
Contact rate by interviewer	5.13897	1.07681	4.77	0.000
Age 35 or less	-0.38668	0.07412	-5.22	0.000
Age 60 or more	0.41902	0.13866	3.02	0.003
Female	0.19440	0.06293	3.09	0.002
Living without a spouse	-0.45405	0.07022	-6.47	0.000
No. of children	0.02563	0.03267	0.78	0.433
No. of adults	-0.30312	0.03255	-9.31	0.000
Unemployed	-0.06132	0.11329	-0.54	0.588
Inactive	0.16245	0.08170	1.99	0.047
Home tenant	-0.13590	0.07143	-1.90	0.057
No. visits	-0.01264	0.01303	-0.97	0.332
Same interviewer	1.46687	0.07932	18.49	0.000
Constant	1.32588	0.18483	7.17	0.000
Wald joint significance test	15268.96	0.00000	No. obs.	18353

Tab. 9 Co-operation probit model for the BHPS

Variables	Coefficients	Std. Err.	Test	p-value
Wave 1996	0.17108	0.07750	2.21	0.027
Workload	-0.01619	0.00073	-22.04	0.000
Item nonresponse by interviewer	-3.08725	0.82625	-3.74	0.000
Co-operation rate by interviewer	1.62772	0.33539	4.85	0.000
Age 35 or less	-0.05109	0.08772	-0.58	0.560
Age 60 or more	-0.01904	0.12842	-0.15	0.882
Female	0.20994	0.07583	2.77	0.006
Living without a spouse	-0.15878	0.08345	-1.90	0.057
No. of children	-0.03666	0.03821	-0.96	0.337
No. of adults	-0.06812	0.04049	-1.68	0.092
Unemployed	-0.38718	0.12897	-3.00	0.003
Inactive	0.16281	0.09911	1.64	0.100
No. of visits	-0.02887	0.01240	-2.33	0.020
Same interviewer	0.61158	0.07865	7.78	0.000
Item nonresponse	0.04194	0.21125	0.20	0.843
Constant	1.54751	0.21200	7.30	0.000
Wald joint significance test	2068.9	0.00000	No. obs.	14265

Tab. 10 Contact probit model for the BHPS

Variables	Coefficients	Std. Err.	Test	p-value
Wave 1996	-0.44521	0.12208	-3.65	0.000
Workload	-0.00901	0.00060	-15.06	0.000
Item nonresponse by interviewer	-2.37872	1.44385	-1.65	0.099
Contact rate by interviewer	-1.27724	1.13916	-1.12	0.262
Age 35 or less	-0.38590	0.12017	-3.21	0.001
Age 60 or more	0.52905	0.20614	2.57	0.010
Female	0.16497	0.10209	1.62	0.106
Living without a spouse	-0.25058	0.11080	-2.26	0.024
No. of children	-0.05523	0.04621	-1.20	0.232
No. of adults	0.01546	0.04856	0.32	0.750
Unemployed	-0.34195	0.15532	-2.20	0.028
Inactive	-0.01448	0.12156	-0.12	0.905
Home tenant	-0.46502	0.10679	-4.35	0.000
No. visits	-0.01958	0.01825	-1.07	0.283
Same interviewer	0.56312	0.13018	4.33	0.000
Constant	4.03541	0.37124	10.87	0.000
Wald joint significance test	1152.79	0.00000	No. obs.	14353

Tab. 11 Co-operation probit model for the ECHP-D

Variables	Coefficients	Std. Err.	Test	p-value
Wave 1996	0.23906	0.11027	2.17	0.030
Workload	0.00020	0.00021	0.98	0.329
Item nonresponse by interviewer	1.48300	0.43094	3.44	0.001
Co-operation rate by interviewer	0.93650	0.34928	2.68	0.007
Age 35 or less	-0.18576	0.04889	-3.80	0.000
Age 60 or more	-0.10657	0.06541	-1.63	0.103
Female	-0.05894	0.04128	-1.43	0.153
Living without a spouse	-0.17310	0.04958	-3.49	0.000
No. of children	0.03860	0.02485	1.55	0.120
No. of adults	-0.07413	0.02360	-3.14	0.002
Unemployed	-0.01172	0.09286	-0.13	0.900
Inactive	0.07514	0.05463	1.38	0.169
No. of visits	-0.05249	0.01271	-4.13	0.000
Same interviewer	0.27312	0.11249	2.43	0.015
Item nonresponse	-0.83429	0.09664	-8.63	0.000
Constant	1.40693	0.23088	6.09	0.000
Wald joint significance test	274.80	0.00000	No. Obs.	15975

Tab. 12 Contact probit model for the ECHP-D

Variables	Coefficients	Std. Err.	Test	p-value
Wave 1996	-1.31674	0.07042	-18.70	0.000
Workload	0.00027	0.00028	0.96	0.339
Item nonresponse by interviewer	-1.00308	0.47542	-2.11	0.035
Contact rate by interviewer	-3.92433	1.56810	-2.50	0.012
Age 35 or less	-0.23249	0.06319	-3.68	0.000
Age 60 or more	0.03082	0.09495	0.32	0.746
Female	0.06947	0.05708	1.22	0.224
Living without a spouse	-0.39378	0.06427	-6.13	0.000
No. of children	-0.06842	0.03020	-2.27	0.023
No. of adults	-0.01064	0.03618	-0.29	0.769
Unemployed	-0.06491	0.11743	-0.55	0.580
Inactive	-0.09397	0.07267	-1.29	0.196
Home tenant	-0.43404	0.06435	-6.74	0.000
No. visits	-0.05328	0.01762	-3.02	0.003
Same interviewer	1.92822	0.09301	20.73	0.000
Constant	5.26515	0.18126	29.05	0.000
Wald joint significance test	778.27	0.00000	No. obs.	16399

Tab. 13 Co-operation probit model for the GSOEP

Variables	Coefficients	Std. Err.	Test	p-value
Wave 1996	-0.07502	0.03426	-2.19	0.029
Workload	-0.00048	0.00004	-12.01	0.000
Item nonresponse by interviewer	-1.13951	0.32671	-3.49	0.000
Co-operation rate by interviewer	-0.12221	0.21261	-0.57	0.565
Age 35 or less	-0.14724	0.03180	-4.63	0.000
Age 60 or more	0.06548	0.04880	1.34	0.180
Female	0.01883	0.02766	0.68	0.496
Living without a spouse	-0.12657	0.03136	-4.04	0.000
No. of children	-0.00801	0.01545	-0.52	0.604
No. of adults	-0.06242	0.01329	-4.70	0.000
Unemployed	-0.07632	0.05169	-1.48	0.140
Inactive	0.05952	0.03562	1.67	0.095
No. of visits	-0.07752	0.00844	-9.19	0.000
Same interviewer	0.18248	0.03626	5.03	0.000
Item nonresponse	-0.27232	0.08683	-3.14	0.002
Constant	1.86627	0.09912	18.83	0.000
Wald joint significance test	645.09	0.00000	No. obs.	25748

Tab. 14 Contact probit model for the GSOEP

Variables	Coefficients	Std. Err.	Test	p-value
Wave 1996	0.13930	0.07023	1.98	0.047
Workload	0.00073	0.00010	7.32	0.000
Item nonresponse by interviewer	-2.74374	0.46917	-5.85	0.000
Contact rate by interviewer	-1.40183	1.51160	-0.93	0.354
Age 35 or less	-0.22762	0.07856	-2.90	0.004
Age 60 or more	0.04805	0.10612	0.45	0.651
Female	0.15242	0.06298	2.42	0.016
Living without a spouse	-0.17653	0.06982	-2.53	0.011
No. of children	0.08853	0.04182	2.12	0.034
No. of adults	-0.09545	0.03020	-3.16	0.002
Unemployed	-0.40883	0.09691	-4.22	0.000
Inactive	-0.34222	0.07786	-4.40	0.000
Home tenant	-0.52179	0.07870	-6.63	0.000
No. visits	-0.00585	0.02144	-0.27	0.785
Same interviewer	1.20599	0.07897	15.27	0.000
Constant	2.29227	0.21541	10.64	0.000
Wald joint significance test	475.05	0.00000	No. obs.	25976