

Respondent, Survey Design and Interviewer Characteristics as Determinants of Respondents' Propensity to Consent to Data linkage

Emanuela Sala¹, Jonathan Burton², Gundi Knies²

¹ Dipartimento di Sociologia e Ricerca Sociale, Università di Milano Bicocca, Piazza dell'Ateneo Nuovo, 20126, Milano, Italy.

² Institute for Social and Economic Research, University of Essex, Wivenhoe Park, Colchester, Essex, CO4 3SQ, UK.

Abstract

In the UK, in order to link individual-level administrative records to survey responses, respondents need to give their consent. This paper explores whether characteristics of the respondent, the survey design or the interviewer influence consent. We use the British Household Panel Survey (BHPS) combined with a survey of interviewers to model the probability that respondents consent to adding health and social security records to their survey responses. A clear pattern of consent emerges. Consent is related to respondents' attitudes to privacy, community-mindedness and to the salience of the data linkage request. Some survey design features such as survey "fidelity" and interview sequence within the household are found statistically significant. By contrast, interviewer characteristics, including personality and attitudes to persuasion, are not associated with consent. Only the interviewer's task-specific experience matters. Implications of the findings are discussed and areas of future research are identified.

Keywords: Consent, Data linkage, BHPS, Administrative records, Interviewer characteristics, Survey design features.

1. Introduction

Linkage of administrative data to survey data is becoming increasingly popular both in the UK and elsewhere. Major social surveys have linked their data with a wide range of administrative data including benefit receipt, adolescent's school performance and health and morbidity (e.g., the US Current Population Survey, the Longitudinal Survey of Young People in England, the UK Millennium Cohort Study). Data linkage is regarded as a powerful tool to overcome some of the main challenges currently facing survey practitioners. Administrative records offer a wealth of information which could

significantly enhance research opportunities, help improve data quality, reduce survey costs and ease respondent (and interviewer) burden.

One of the challenges to this end is gaining respondents' informed consent to their data being linked. Willingness to give consent is not universal, reducing the number of observations and potentially introducing bias. In fact, previous studies have shown that consenters and non-consenters vary on socio-economic characteristics (for recent reviews see Dunn et al., 2004; Huang et al., 2007; Tate et al., 2006), and there is some empirical evidence that consent is also associated with features of the data collection process and with study characteristics; including the survey topic, the domain of the data linkage (Jenkins et al., 2006; Singer et al., 2003) and who is asking for consent (Armstrong et al., 2007). Systematic research on these issues is as yet scant, even in the survey methodology literature. In particular, there are very few studies that look at consent bias on general population surveys, or that explore potential differences across different domains of data linkage (cf. Jenkins et al., 2004). Moreover, no study has investigated the role specific interviewer characteristics and survey design features play in the consent process.

This paper aims to identify the correlates of obtaining informed consent to data linkage. It makes an important contribution to the existing literature by systematically examining consent bias not only with respect to respondent characteristics and survey design features but also with respect to interviewer characteristics. We use an innovative study design drawing on the British Household Panel Study (BHPS) combined with a rich dataset from a survey of the BHPS interviewers.

1.1 Factors that affect consent to data linkage

1.1.1 Respondent characteristics

The research that explores respondent characteristics and consent to data linkage has produced two major findings. First, consent has been found to be associated with respondents' socio-demographic characteristics (age, gender, socio-economic status and ethnicity) and their health (Gerber et al., 2007; Olson, 1999). However, the nature of the relationship between the different respondent characteristics and the propensity to consent remains unclear as characteristics that are associated with higher consent in one study are negatively associated with consent in another (see, for example, Kho et al. 2009).

Second, consent is related to respondent's perception of risk, altruism and community-mindedness. Consent is lower among people who refuse to provide information on income or wealth (Jenkins et al., 2006; Olson, 1999; Woolf et al., 2000), who believe that the data may be used for fraud detection (Gray et al., 2008) and higher among those who perceive that the wider society can benefit from the data linkage (Dunn et al., 2004; Jenkins et al., 2006). Consent is also lower for respondents who fear that information may not be kept confidential (Armstrong et al., 2007).

1.1.2 Survey design features

Research into the impact of survey design features on a wide range of survey outcomes shows, for example, that interview length and topic do play a role in obtaining

respondents' consent to co-operate (for a review see Groves and Couper, 1998). This is an area of research that is very much under-researched with respect to consent to data linkage. We are aware of only a single study that explores this. Jenkins et al. (2006) found that consent to data linkage is positively associated with the quality of the interviewer-respondent rapport (as measured by interview length and the interviewer evaluation of the "smoothness" of the interview). It is, then, not implausible that other characteristics of the interview process such as number – and order – of interviews in the household also affect consent.

1.1.3 Interviewer characteristics

A further gap in the empirical literature on patterns of consent to data linkage exists with respect to interviewer characteristics. However, this research field is quite promising: qualitative epidemiological studies suggest that patients' propensity to consent varies with the status of the medical staff who is asking for consent (i.e., consent rates are higher when GPs, rather than receptionists ask for it, see, e.g., Armstrong et al., 2007; Baker et al., 2000). Also survey methodology research has consistently documented the occurrence of interviewer effects in a wide range of survey outcomes including (non)response and data quality (amongst others, Fuchs, 2009; O'Muircheartaigh and Campanelli, 1998; Pickery and Loosveldt, 2000).

Unfortunately, the former research strand has not investigated what particular characteristics of the medical staff (e.g., age, gender etc.) are likely to be associated with patients' consent and the latter, while looking at a broad range of interviewer characteristics, has failed to clearly identify what interviewer characteristics are driving these interviewer effects (Esbensen and Menard, 1991; Link, 2006; Lipps, 2007; O'Muircheartaigh and Campanelli, 1999; Pickery and Loosveldt, 2000; Pickery and Loosveldt, 2001; Pickery and Loosveldt, 2004). The role of interviewer experience remains particularly unclear. For example, Hansen (2007), Pickery and Loosveldt (2000) and Jäckle et al. (2010) find that more experienced interviewers achieve higher response rates while Kennickell (1999) and, more recently, Durrant et al. (2010) show that long-term interviewers can perform less well than those with less experience. Some empirical evidence shows that the interviewer personality and attitudes, rather than their socio-demographic characteristics, are associated with a number of survey outcomes. For example, interviewer confidence and attitudes towards persuasion have a positive impact on survey response (Kennickell, 1999; Lehtonen, 1996). Preliminary work by Jäckle et al. (2010) finds that interviewer personality, measured using the so-called 'Big Five' instrument (John and Srivastava, 1999), is associated with co-operation. That study also finds, however, that interviewer attitudes to persuading respondents are *not* associated with respondents' co-operation.

Against the backlight of this literature, this paper explores the role of a broad range of respondent characteristics, survey design features and interviewer characteristics on respondents' propensity to consent to administrative data linkage. We expect to find that respondent's propensity to consent is associated with indicators of respondents' risk aversion and community-mindedness as well as with survey design features such as interviewer-respondent rapport and household-interview specific characteristics (the number – and order – of interviews in the household). We also speculate that consent is associated with interviewers' personality and their attitudes to persuading respondents. Given the mixed and sometimes inconsistent findings yielded by previous empirical

research, we are not expecting to find any particular associations between consent and standard respondent and interviewer socio-demographic characteristics.

2. Data

We use the British Household Panel Study (BHPS) combined with information gathered in a survey of the interviewers who collected the data.

2.1 The British Household Panel Survey (BHPS)

The BHPS is one of the most important research resources in the UK and is one of the longest running household panel studies in the world. Run by the Institute for Social and Economic Research (ISER), it started in 1991 with a sample of nationally-representative stratified, clustered sample of 5,500 households and roughly 10,000 individuals interviewed face-to-face, with interviewers calling on respondents in their homes. In 1999, booster samples of around 1,500 households each were added in Scotland and Wales, and in 2001, a sample of 2,000 households was added in Northern Ireland. In Wave 18, the most recent of the survey, 12,971 full interviews were completed. Annual waves of data collection provide a wide range of information including household composition and conditions, education and training, health and use of health services, labour market behaviour, socio-economic values and different income sources.

2.1.1 Data linkage

Asking for consent to data linkage to health, social security benefits and educational administrative records was implemented at Wave 18 of the BHPS. The data linkage module was administered at the end of the individual questionnaire. In the UK informed consent must be obtained from respondents in order to link administrative data at the individual level to survey data. If the respondent verbally agreed to give consent, the interviewer then handed them a form that the respondent was asked to read and sign. All adults were asked for their consent to link to their own health and benefit records. None of the consents were conditional on other consents being given, so if someone refused to give consent to one data linkage they were still asked about the next data linkage. Forty-one percent of adult respondents gave consent to health data linkage whereas 32 percent consented to the linkage to economic records.

2.2 The BHPS interviewer survey

The interviewer survey consisted of a self-completion questionnaire administered during Wave 18 BHPS briefings. Researchers from ISER attended a number of these in-person briefings and administered a questionnaire to all 180 interviewers present at those briefings (68 percent of all interviewers at Wave 18). At briefings at which a researcher was present, all interviewers (100 percent) completed the questionnaire and returned it to the researcher in a sealed envelope. We used a dataset of interviewers provided by the survey agency to check for bias between interviewers who completed the interviewer survey and those who did not. Although we did not find any evidence for bias with respect to interviewer age, we did find that men were more likely to be overrepresented in the interviewer survey.

The interviewer questionnaire collects five types of information: basic information on socio-demographic characteristics (age, sex, educational qualification, presence of children at home, household composition), interviewer experience (type and duration), interviewers' views on different aspects of their job, interviewers' personality traits measured by the "Big Five" taxonomy (John and Srivastava, 1999), their attitudes to persuading and contacting respondents measured by five items from the Lehtonen scale (Lehtonen, 1996) and three items used by Blohm et al. (2007).

The level of item non-response was very low and varied from 1 percent or less for the questions on interviewer experience to about 2 percent on questions on personality traits and attitudes to persuading respondents. The highest level of item non response was for education (5 percent).

3. A Statistical Model of Consent

3.1 Model specification

There are two consent outcomes which are available for all adult respondents to the BHPS; the consent to link administrative health records and the consent to link administrative benefit records. The data space allows us to focus on three different probabilities, i.e., the probability to consent to health data linkage only, the probability to consent to the benefit data linkage only, and the probability to consent to both data linkage requests. From the point of view of substantive analysis all three probabilities are interesting because they tell us how much bias we may expect if we use BHPS linked with health records only, with benefit records only or with records from both domains (assuming there exists a record for each consenter and it can be linked successfully), respectively.

In our empirical analysis we will focus on the probability to consent to both health and benefit data linkage, mainly because we are concerned not only with consent bias but also with modelling consent in the most comprehensive way possible, and achieving a very high degree of generalisability. We will estimate respondent's propensity to consent on the basis of both outcomes using multivariate bivariate probit models, which can be written as:

$$\begin{cases} y_{1i}^* = \beta_1' x_{1i} + u_{1i} \\ y_{2i}^* = \delta_2' z_{2i} + u_{2i} \end{cases}, i = 1, 2, \dots, n$$

where y_{1i}^* and y_{2i}^* are latent variables so that the observed dichotomous outcomes y_{1i} , i.e., the health record linkage request, and y_{2i} , i.e., the benefit record linkage request, are given by:

$$\begin{cases} y_{ki} = 1 \text{ if } y_{ki}^* > 0 \\ y_{ki} = 0 \text{ if } y_{ki}^* \leq 0 \end{cases}, k = 1, 2$$

In the model, x_{1i} and z_{2i} are vectors of observed exogenous variables that have been suggested to affect consenting, and β_1 and δ_2 are the respective parameter vectors. The error terms in this model are distributed as standard bivariate normal variables with correlation coefficient Rho (ρ). More detailed information on this standard model can be found, for instance, in Greene (2003). To our end it is important to note that Rho may be interpreted as the respondents' unobserved propensity to consent (see Jenkins et al., 2006). If the parameter is statistically significant, modelling the consent outcomes jointly

is more efficient than using univariate probit models (or indeed univariate linear probability models), which otherwise yield the same substantive results. Estimation of the model is straightforward using Stata's *biprobit* command (StataCorp, 2009). We adjust standard errors for clustering on interviewers.

3.2 Choice of predictor variables

The BHPS offers plenty of information on respondents, their households, and the interview situation, both for the present and the past. Our choice of variables is guided by the literatures on consent bias, survey co-operation and interviewer effects on data quality. We organise the variables in three blocks, i.e., respondent characteristics, survey design features and interviewer characteristics.

3.2.1 Respondent characteristics

Like most other research on consent bias, our models consider respondent demographic characteristics (age, gender, and ethnicity) and their socio-economic characteristics (education, household income, and household context) as well as region of residence. Moreover, we include a number of characteristics which we believe tap into the respondent's perceptions of the risk of data linkage (e.g., information on refusing to provide information on income from investments to proxy for the respondent's general attitudes to sharing information), indicators of data linkage salience (whether or not respondents have been to hospital in the previous 12 months, had any of 15 types of health problems; whether they currently receive income support payments from the government, and how many means-tested benefits they receive) and of 'community-minded' attitudes (dummy variables for (i) whether or not the respondent supports a left-wing/liberal party, (ii) whether or not they do voluntary work without receiving pay, and (iii) whether or not they generally trust others).

3.2.2 Survey design features

With respect to survey design features potentially affecting consent, we include a number of proxy measures for rapport (the number of years the respondent has been participating in the BHPS, and whether or not the interviewer in the current wave interviewed the respondent in the previous year). To capture potential influences of others we include a dummy for whether or not others were present at any time during the interview (i.e., not specifically when the consent was asked). Moreover, we consider how many interviews had already taken place in the household for the present BHPS wave, and the number of consents that had already been given by other household members at the time the respondent is asked. This exploits information about the time of the interviews with other members of the household and the respective consent outcomes. We believe the measures will pick up what we might refer to as 'household contagion', i.e., the influence of the respondent's and the interviewer's knowledge of how easy/difficult it has been to get consents from the people already interviewed in this household.

3.2.3 Interviewer characteristics

With respect to interviewer characteristics, we use interviewer socio-demographic characteristics (sex, age, education). In addition, we include three different measures of

‘experience’; (i) job experience, i.e., the number of years that the person has been an interviewer, (ii) survey experience, i.e., the number of interviews on this survey the interviewer has carried out this wave and (iii) task experience, i.e., the information about what has already occurred when asking for consent within interviews this wave. The idea here is that interviewers may accumulate not only knowledge about how easy or difficult it is to obtain consents within a given household, but also across households. This could pan out either positively, namely, if interviewers learn from their past task-specific experience and manage to adjust the way in which they ask consents, or negatively, if they do not.

To explore any other influence that interviewers may have on obtaining respondents’ consent, we also include the Lehtonen scale that measures interviewer attitudes towards persuasion as well as a version of the “Big Five” personality traits. All these variables have been collected in the BHPS interviewer survey (described above).

4. Results

Table 1 reports the results of joint estimation of consent to both health and benefit data linkage, controlling for respondent characteristics, survey design features and interviewer characteristics. First, note that the cross-equation correlation Rho is highly statistically significant suggesting that there is an unobserved factor that affects both decisions. We interpret this as the respondents’ unobserved propensity to consent. It cannot be ruled out, however, that despite the large number of regressors included in our analysis, there are other unobserved characteristics that influence both consents.

Table 1 about here

4.1 Respondent characteristics

When looking at respondents’ characteristics that influence consent, a clear pattern emerges: respondents’ propensity to consent to data linkage does not seem to be strongly associated with their demographic or socio-economic characteristics but it appears to be related to their attitudes to privacy, community-mindedness and data linkage salience.

Table 1 shows that, on the whole, demographic and socio-economic respondent characteristics are only mildly associated with consent. Members of UK minority ethnic groups as well as older respondents are less likely to consent. These findings are consistent with previous studies (see, e.g., Hockley et al., 2008; Tate et al., 2006; Woolf et al., 2000). However, we do have to add the caveat that the ethnic minority sample size in the BHPS is quite small, even when groups are combined. Characteristics describing the respondent’s household context (including household income) do not appear to be associated with consent.

On the other hand, indicators of respondents’ attitudes to privacy and community-mindedness show a marked association with consent. In particular, refusing to answer the question on income from investment is a strong predictor of not giving consent, while generally trusting others positively affects consent. Indicators of saliency of the data linkage such as being in the hospital in the last 12 months or receiving a larger number of means-tested benefits are positively associated with consent.

The pattern of consent that we have just described is true for both types of consent (though some of the effects seem to have more influence on consent to benefit record

linkage). For example, for the salience indicators, the statistical associations with consent vary for the health and benefit consent outcomes. Whilst the health measure is only significant for the health linkage, the benefit measure is significant on both.

4.2 Survey design features

As Table 1 clearly shows, respondents' propensity to consent to data linkage also is strongly associated with some survey design features; the sequence of interviews within the household, the effects of previous consent requests ('household contagion') and survey "fidelity". The direction of some of these relationships, however, is not always as expected. Interviewer-respondent rapport as measured by the variable "same interviewer as previous wave" does not seem to play a role in the consent process.

There is a negative association between household interview sequence and consent. Later interviewees are less likely to consent to data linkage compared to household members who are interviewed earlier. We interpret this variable as an indicator of survey resistance. However, it could also be interpreted as an indicator of interviewer burden (e.g., due to time pressure interviewers rush through the later interviews).

Interestingly, respondents' probability to consent is positively associated with the number of household members who have already consented to data linkage. This is evidence for a 'household contagion' effect (i.e., household members consult each other and take joint decisions). Note that mere presence of others during the interview does not appear to be associated with consent.

Contrary to our expectations, respondents' consent to data linkage is negatively associated with the survey "fidelity" indicator: respondents' propensity to consent decreases with the number of years they have been in the panel.

We also do not find strong evidence to support the claim that interviewer-respondent rapport has an impact on consent. Respondents' consent is only very weakly associated with having been interviewed by the same interviewer in the previous year.

As with respondent characteristics, the pattern of consent that we have just described is true for both types of consent.

4.3 Interviewer characteristics

Table 1 shows that interviewer socio-demographic characteristics are not associated with respondents' propensity to consent. Furthermore, interviewer personality traits and attitudes to persuading respondents show no statistically significant association with consent. The only interviewer characteristic that appears to matter is experience. Albeit, it is not the length of time they have worked as interviewers that matters, but rather their more specific survey experience in the current wave and their task-specific experience. The more BHPS interviews an interviewer has already carried out during the wave, the less likely she/he is to obtain respondents' consent – this may reflect that households who may be more reluctant or harder to interview will be interviewed later in the fieldwork period. However, the more successful she/he has been in obtaining respondents' consent in one particular domain, the more likely she/he is to gain respondents' consent.

5. Discussion

Despite the increasing number of British and international surveys that ask respondents for permission to data linkage, there is currently very little knowledge on what drives consent. Research on consent to data linkage is mainly constituted by descriptive analyses of the variation in respondents' propensity to consent and the potential resulting respondent bias. The assumption underlying these studies is that the mechanisms that govern consent are located mainly in survey respondents (or, in case of medical studies, patients) and, in particular, in their socio-demographic characteristics. However, factors relating to the survey process may be equally important. Interviewers and their characteristics, the survey topic as well as the point in time during the life of a panel survey and the survey mechanics within the household could play a major role in the respondents' decision whether or not to consent.

This paper advances the knowledge about consent and consent bias in important ways. It is the first empirical analysis to explore the role of respondents, survey design features and the interviewer in obtaining respondents' consent to perform data linkage. Consent bias is examined for different types of administrative data, comparing consent to link economic records and health records. The research looks at the issue of consent to data linkage using a general population sample, rather than a specialised medical-based sample. Moreover, it draws on a large pool of interviewer-level characteristics, making this the first research investigating whether interviewer attitudes to persuading respondents and interviewer personality affect respondents' consent to data linkage. It is also the first time that the mechanics of interviewing within a household context are explored.

Although further studies are needed to draw general conclusions on the mechanisms that lead survey respondents to consent to link administrative records to their survey data, our analysis shows a clear pattern of consent. We find that respondent socio-demographic characteristics are mildly associated with consent, but there is a much greater effect of their attitudes towards privacy and community-mindedness. Contrary to other research, we do not find that survey design features, such as the length of the interview are significantly associated with consent (Jenkins et al., 2006). We also find that the length of time in the panel was significant, but not positively, as we expected; those who had been in the panel for longer are *less* likely to give consent, other things being equal. This is an interesting finding. We speculate that BHPS respondents who have been parts of the panel for longer may feel they have provided so much information already over the past (up to 18 years) that they do not see why access to administrative data may be needed. Another reason may be that they are suspicious of a survey innovation which comes about after so many years.

We find that interviewer characteristics, including interviewer attitudes and personality traits, which have been suggested as a source for variation in survey response in the survey methodology literature, are not associated with consent to data linkage. We do, however, find a significant effect of the intra-household dynamics on consent, suggesting that the decision by an individual is located within the interaction between the individual, the interviewer and the wider household-context. For example, we find a negative relationship between the number of BHPS interviews an interviewer has already carried out during a given wave and the likelihood to obtain respondents' consent. However, when an interviewer is successful in obtaining consent to perform data linkage in one domain, she/he is also very likely to gain another respondent's consent in that domain.

We theorise that an interviewer who has already asked the consent questions, and the consent has been withheld, is likely to feel less optimistic about asking an additional person in the same household. This may affect the effort made by the interviewer to explain the purpose of the data linkage, and may even affect whether or not the interviewer actually asks the question – rather than just assuming a refusal. A respondent who is aware that others in the household have already withheld their consent may find it easier to withhold their own consent – and thus save themselves a couple of minutes in the interview – and may even feel some ‘peer pressure’ to refuse to maintain a consistent household response (and not appear inconsistent).

So what are the implications of this and how might future research shed further light on these issues? Given the relatively low consent rates, data analysts should give a careful consideration of statistical power while addressing specific research questions (e.g., when looking at health or economic conditions of minority groups). They also should be aware that adjustments may be needed to take into account the sources of bias that may be introduced while asking for respondent’s consent. These adjustments may vary depending on what – and how many – administrative sources are being used (see Sala et al. 2010, Table 3). To address these two issues (why only some respondents consent and why different types of respondents consent), future studies could aim to provide a better understanding of the reasons why respondents consent or do not consent to data linkage, for example by carrying out qualitative studies on groups of consenters and/or non-consenters or adding a follow up question in a quantitative study that collects respondents’ reasons for withholding consent.

Our research findings also have implications for survey designers. On the one hand, survey designers may tailor the consent question differently on the basis of answers given to prior survey questions that are strongly correlated with propensity to consent (e.g., missing data on income from investments). On the other hand, survey designers and survey agencies may develop an interviewer training programme which also includes a discussion on how household members take survey decisions and provides guidelines on how to deal with difficult cases. To this extent, sections of the interview could be recorded and behaviour-coded. We might find indications of household-level interactions which lead to order-effects, or indications of time pressure on the interviewer which may lead to a tendency to skip past the consent question, particularly if other members of the household have already withheld their consent. Paradata on timing and contact attempts are needed to test the order-effect hypothesis. Only when a substantial body of empirical research has been collected, a theoretical model that explains the complex processes that lead respondents to consent can be formulated.

Acknowledgments: We are grateful to colleagues at the Institute for Social and Economic Research and to participants at the Institute of Education-ISER Joint Seminar on data linkage for their comments on previous drafts of the paper. Special thanks go to Mark Bryan, Stephen Jenkins, Peter Lynn and Stephanie McFall. This project is funded by the ESRC grant RES-586-47-0002.

REFERENCES

- Armstrong V., Barnett J., Cooper H., Monkman M., Moran-Ellis J., Shepherd R. (2007) Public Perspectives on the Governance of Biomedical Research: A qualitative study in a deliberative context Wellcome Trust, London.
- Baker R., Shiels C., Stevenson K., Fraser R., Stone M. (2000) What proportion of patients refuse consent to data collection from their records for research purposes? *British Journal of General Practice* 50:655-556.
- Blohm M., Hox J.J., Koch A. (2007) The influence of interviewers' contact behavior on the contact and cooperation rate in face-to-face household surveys. *International Journal of Public Opinion Research* 19:97-111.
- Dunn K.M., Jordan K., Lacey R.J., Shapley M., Jinks C. (2004) Patterns of Consent in Epidemiologic Research: Evidence from Over 25,000 Responders. *American Journal of Epidemiology* 159:1087-1094.
- Durrant G.B., Groves R.M., Staetsky L., Steele F. (2010) Effects of Interviewer Attitudes and Behaviors on Refusal in Household Surveys. *Public Opinion Quarterly* 74:1-36.
- Esbensen F. A., Menard S. (1991) Interviewer-Related Measurement Error in Attitudinal Research: A Nonexperimental Study. *Quality and Quantity* 25:151-165.
- Fuchs M. (2009) Gender-of-Interviewer Effects in a Video-Enhanced Web Survey Results from a Randomized Field Experiment. *Social Psychology* 40:37-42.
- Gerber Y., Jacobsen S. J., Killian J. M., Weston S., Roger V. L. (2007) Participation bias assessment in a community-based study of myocardial infarction, 2002-2005. *Mayo Clinic Proceedings* 82:933-938.
- Gray M., Constantine R., d'Ardenne J., Blake M., Uhrig S.C.N. (2008) Cognitive testing of Understanding Society. The UK Household Longitudinal Study Questionnaire. Understanding Society Working Paper 04-2008.
- Greene W.H. (2003) *Econometric Analysis*. Prentice Hall, Upper Saddle River, New York.
- Groves R.M., Couper M.P. (1998) *Nonresponse in household interview surveys* Wiley, New York.
- Hansen K.M. (2007) The Effects of incentives, interview length, and interviewer characteristics on response rates in a CATI study. *International Journal of Public Opinion Research* 19:112-121.
- Hockley C., Quigley M.A., Hughes G., Calderwood L., Joshi H., Davidson L.L. (2008) Linking Millennium Cohort data to birth registration and hospital episode records. *Paediatric and Perinatal Epidemiology* 22:99-109.
- Huang N., Shih S.F., Chang H.Y., Chou Y. (2007) Record linkage research and informed consent: who consents? *BMC Health Services Research* 7:18.
- Jäckle A., Lynn P., Sinibaldi J., Tipping S. (2010) The Effect of Interviewer Personality, Skills and Attitudes on Respondent Co-operation with Face-to-Face Surveys. mimeo.

- Jenkins S.P., Lynn P., Jäckle A., Sala E. (2004) Linking household survey and administrative record data: what should the matching variables be?, ISER Working Paper, University of Essex, Colchester.
- Jenkins S.P., Cappellari L., Lynn P., Jäckle A., Sala E. (2006) Patterns of consent: evidence from a general household survey. *Journal of the Royal Statistical Society, Series A* 169:701-722.
- John O.P., Srivastava S. (1999) The Big Five Trait Taxonomy: History, Measurement, and Theoretical Perspectives, in: O. P. John and L. A. Pervin (Eds.), *Handbook of Personality: Theory and Research*, Guilford Press, New York, 102-138.
- Kennickell A.B. (1999) Analysis of Nonresponse Effects in the 1995 Survey of Consumer Finances. *Journal of Official Statistics* 15:283-303.
- Lehtonen R. (1996) Interviewer attitudes and unit nonresponse in two different interviewing schemes., in: S. Laaksonen (Ed.), *Sixth International Workshop on Household Survey Nonresponse*, Statistics Finland, Helsinki.
- Link M.W. (2006) Predicting the Persistence and Performance of Newly Recruited Telephone Interviewers. *Field Methods* 18:305-320.
- Lipps O. (2007) Interviewer and Respondent Survey Quality Effects in a CATI Panel. *Bulletin de Methodologie Sociologique* 95:5-25.
- O'Muircheartaigh C., Campanelli P. (1998) The relative impact of interviewer effects and sample design effects on survey precision. *Journal of the Royal Statistical Society, Series A* 161:63-77.
- O'Muircheartaigh C., Campanelli P. (1999) A multilevel exploration of the role of interviewers in survey non-response. *Journal of the Royal Statistical Society, Series A* 162:437-446.
- Olson J.A. (1999) Linkages With Data From Social Security Administrative Records in the Health and Retirement Study. *Social Security Bulletin* 62: 73-85.
- Pickery J., Loosveldt G. (2000) Modeling Interviewer Effects in Panel Surveys: An Application. *Survey Methodology* 26:189-198.
- Pickery J., Loosveldt G. (2001) An exploration of question characteristics that mediate interviewer effects on item non-response. *Journal of Official Statistics* 17:337-350.
- Pickery J., Loosveldt G. (2004) A Simultaneous Analysis of Interviewer Effects on Various Data Quality Indicators with Identification of Exceptional Interviewers. *Journal of Official Statistics* 20:77-89.
- Singer E., Van Hoewyk J., Neugebauer R.J. (2003) Attitudes and behaviour. The impact of privacy and confidentiality concerns in the 2000 Census. *Public Opinion Quarterly* 67:368-384.
- StataCorp. (2009) *Statistical Software: Release 11.0* Stata Corporation, College Station, TX.
- Tate A.R., Calderwood L., Dezateux C., Joshi H. (2006) Mother's Consent to Linkage of Survey Data with her Child's Birth Records in a Multi-ethnic National Cohort Study. *International Journal of Epidemiology* 35:294-298.
- Woolf S.H., Rothemich S.F., Johnson R.E., Marsland D.W. (2000) Selection Bias From Requiring Patients to Give Consent to Examine Data for Health Services Research. *Archives of Family Medicine* 9:1111-1118.

Table 1. Propensity to consent as a function of respondent, interview and interviewer characteristics (bivariate probit regressions).

	Coefficients	
	Consent to health data linkage equation	Consent to benefit data linkage equation
Respondent characteristics		
Male	0.04	0.09**
Ethnicity (British/Irish White)		
<i>Other White</i>	-0.29**	-0.25*
<i>Mixed</i>	-0.29	-0.24
<i>British Asian/Black</i>	-0.38**	-0.37*
<i>Other ethnicity</i>	-0.13	-0.09
Age group (16-24 years old)		
<i>25-39 years old</i>	-0.21*	-0.29**
<i>40-49 years old</i>	-0.17	-0.33***
<i>50-59 years old</i>	-0.12	-0.25*
<i>60+ years old</i>	-0.07	-0.30**
Education degree or beyond	0.09	0.15**
Household type (Single)		
<i>Couple, no children</i>	0.01	-0.04
<i>Couple with children</i>	0.00	-0.10
<i>Lone parent</i>	-0.01	-0.06
<i>Other household type</i>	0.16	0.03
Household size	-0.01	-0.02
England	0.03	0.03
London/Southeast	0.16*	0.15*
Household income (log)	0.02	0.04
Refused question: Income from investment	-0.62***	-0.75***
Generally trusts others	0.22***	0.19***
Supports leftwing/liberal party	0.17***	0.19***
Does unpaid voluntary work	0.12*	0.11*
Has health problems	0.06	0.08
Has been to hospital	0.14*	0.01
Receives any state benefits	-0.01	0.02
Number of means-tested benefits received	0.04	0.07**

(continues)

Table 1. (continued)

	Coefficients	
	Consent to health data linkage equation	Consent to benefit data linkage equation
Survey design features		
Interview sequence within household	-0.75***	-0.71***
Number of previous health consents in household	1.14***	0.60***
Number of previous benefit consents in household	0.40***	0.93***
Others present during interview	0.03	0.08*
Number of years in the BHPS	-0.02**	-0.02***
Same interviewer as previous wave	0.09	0.15
Interviewer characteristics		
Male interviewer	-0.03	-0.11
Interviewer age group (40-49 years old)		
50-59 years old	0.05	0.05
60-69 years old	0.12	0.09
70+ years old	-0.04	-0.13
Interviewer has degree or above	0.07	0.03
Interviewing experience in years	0.00	-0.02
Number previous interviews by interviewer	-0.02***	-0.02***
Number of health consents already obtained	0.03**	0.01
Number of benefit consents already obtained	0.01	0.03**
Attitudes to persuading		
All can be persuaded	-0.05	-0.09
Should persuade	-0.01	-0.01
Should respect privacy	-0.07	-0.11
Should accept refusal	0.06	0.07
Emphasise voluntary nature	-0.06	-0.04
Personality traits		
Agreeableness	0.03	0.03
Conscientiousness	0.02	0.00
Openness	-0.02	0.00
Extraversion	-0.05	-0.05
Neuroticism	-0.03	-0.02
Cross-equation correlation	0.94	
Log(pseudo)Likelihood	-4541.3	
N	5825	

Source: BHPS Wave 18 linked with interviewer survey.