Effects of Interviewer Attitudes and Behaviors on Refusal in Household Surveys

Gabriele B. Durrant[1], Robert M. Groves[2], Laura Staetsky[3] and Fiona Steele[4]

[1] Corresponding author
Southampton Statistical Sciences Research Institute
University of Southampton
Southampton, SO17 1BJ
United Kingdom
Tel: +44 (0)23 8059 5481
Fax: +44 (0)23 8059 5763
g.durrant@southampton.ac.uk

[2] Survey Research Centre
University of Michigan
426 Thompson Street
Ann Arbor, MI 48106-1248
USA
Tel: +1 734 764 8365
Fax: +1 734 764 8372
grovesb@umich.edu

[3] Southampton Statistical Sciences Research Institute
University of Southampton
Southampton, SO17 1BJ
United Kingdom
Tel: +44 (0)23 8059
Fax: +44 (0)23 8059 5763
lst@southampton.ac.uk

[4] Graduate School of Education
University of Bristol
35 Berkeley Square
Bristol, BS8 1JA
United Kingdom
Tel: +44 (0) 117 331 0828
Fax: +44 (0) 117 925 5412
Fiona.Steele@bristol.ac.uk

Word count of main text: 6490 (without references, tables etc)

Word count of abstract: 242

GABRIELE B. DURRANT is lecturer at the Southampton Statistical Sciences Research Institute, University of Southampton. ROBERT M. GROVES is professor and director, University of Michigan Survey Research Center, and research professor at the Joint Program in Survey Methodology. LAURA STAETSKY is research fellow at the Southampton Statistical Sciences Research Institute, University of Southampton. FIONA STEELE is professor at the Centre for Multilevel Modelling, Graduate School of Education, University of Bristol. Address correspondence to Gabriele B. Durrant, e-mail: g.durrant@southampton.ac.uk.

Acknowledgements.

The research was funded by the Economic and Social Research Council (ESRC), UK, titled 'Hierarchical analysis of unit nonresponse in sample surveys', grant number: RES-062-23-0458, award holders: Gabriele B. Durrant (Principal Investigator), Fiona Steele, Chris Skinner and John Micklewright. The authors appreciate comments on an earlier draft from Chris Skinner and John Micklewright. The authors would like to thank the UK Office for National Statistics (ONS) for allowing access to the dataset. Any views expressed in this paper are those of the authors.

Abstract

Interviewers play a crucial role in gaining cooperation from a sample unit. This paper aims to identify the interviewer characteristics that influence survey cooperation. Of principal interest to survey practitioners are interviewer attributes associated with higher cooperation rates, particularly among sample members whose characteristics are traditionally associated with a lower probability of response. Our data source is unusually rich, in that it contains extensive information on interviewers including their attitudes and behaviors which is linked to detailed information on both responding and nonresponding sample units. An important value of the data is that they permit examining a host of as yet unanswered questions about whether some interviewer attributes stimulate cooperation among some respondents but not others. In short, we investigate if some sample units react favorably to certain interviewer characteristics. A multilevel cross-classified logistic model with random interviewer effects is used to account for clustering of households within interviewers due to unmeasured interviewer attributes, and for the cross-classification of interviewers within areas. The model allows for statistical interactions between interviewer and household characteristics.

We find that interviewer confidence and attitudes towards persuading reluctant respondents play an important role in explaining between-interviewer variation in refusal rates. We also find evidence of interaction effects between the interviewer and householder, for example with respect to gender and educational level, supporting the notion of similarity of interviewers and respondents generating higher cooperation. The results are discussed with respect to potential implications for survey practice and design.

Introduction

As survey response rates continue to decline, efforts to understand how they affect the quality of survey estimates abound (Keeter et al., 2000; Groves and Peytcheva, 2008). One challenge is clear - the field must first understand the mechanisms that affect the decision to participate in a survey, and then learn how those mechanisms relate to the survey variables we wish to measure. However, efforts to understand survey participatory decisions have been plagued by inadequate data resources informative about nonrespondents.

In face-to-face and telephone surveys one of the survey participatory mechanisms under partial control of the researcher is the interviewer. Survey methodology has long recognized the essential role of the interviewer in the quality of survey estimates (Feldman, et al. 1951-52; Hanson and Marks, 1958). The interviewer is the key agent of the researcher: interviewers define the purposes of the survey to the sample persons, provide them with a key rationale for participating in the survey, address any concerns and find convenient times for the respondent to be interviewed.

The interviewer activities are crucial determinants of measurement error and nonresponse. Research has primarily focused on the measurement error aspect and has shown that when interviewer characteristics are thematically linked to the constructs of the measurement, measurement error variance associated with interviewers can be induced (e.g. for gender-induced effects see Kane and Macauley, 1993; Flores-Macias and Lawson, 2008). In statistical terms, these effects might be viewed as fixed-effects, i.e. they are conceptualized as systematic influences repetitive over realizations of the survey. Another literature examines variation without explicitly identifying its source. "Interviewer variance" (due to unobserved characteristics) is the variability in survey estimates expected to arise when survey estimates vary depending on which interviewers conduct the data collection. Interviewer variance has been shown to have magnitudes

that rival sampling variance (Bailey et al. 1978) and there is evidence of larger interviewer variance as a function of interviewer training and experience (Freeman and Butler, 1976; Fowler and Mangione, 1990).

While the above literature is large, it focuses on *measurement* error properties. In contrast, there is relatively little with focus on nonresponse. Interviewers with more experience and more confidence about their performance prior to the survey tend to have higher cooperation rates (Groves and Fultz, 1985; Hansen, 2007; Groves and Couper, 1998). There is also evidence of higher cooperation when attributes of interviewers and respondents are matched (Moorman et al., 1999). However, many of the prior studies have the weakness of being post-hoc analyses, fraught with nonobservation of those who are not interviewed. They often have only limited information available on interviewers, are only based on a small number of interviewers and analyze effects in only one survey with a specific design and survey topic, which make general conclusions more difficult (e.g. O'Muircheartaigh and Campanelli, 1999; Pickery et al., 2001).

This paper addresses a unique match of survey data with census data. A key strength of this study is the rich information available on the interviewers and both the responding and nonresponding sample units, including information about each call made to the household. This unique type of data permits the investigation of interaction effects between characteristics of sampled units and those of interviewers which for many other studies has not been possible so far due to data limitations (Groves and Couper, 1998; O'Muircheartaigh and Campanelli, 1999; Pickery and Loosveldt, 2002, 2004). This study also takes account of several surveys simultaneously.

The paper asks the following questions:

1. What are the attributes of interviewers that affect cooperation rates?

2. Are different interviewer attributes important for the participation of different sample units?
3. Do interviewer-level influences on cooperation depend on survey designs?

To address these questions we use multilevel cross-classified models with random interviewer effects to account for individual interviewer characteristics, the clustering of households within interviewers and the cross-classification of interviewers within areas. The structure of the paper is as follows. First, the study design and data are described. Then, the analysis method and the results are presented. The final section discusses implications for survey practice.

Study Design and Data

This study benefits from detailed information about interviewers obtained via a survey of face-to-face interviewers employed by the UK Office for National Statistics (ONS) in 2001 (Interviewer Attitude Survey, IAS; Freeth et al., 2002). The timing of the survey was chosen to coincide with the 2001 UK Census. The design of the questionnaire was conceptually based on earlier work by Groves and Couper (1998) and Hox and De Leeuw (2002). The survey provided information on socio-demographic characteristics, interviewer experience, work background, workload planning and organization, interviewing strategies, doorstep approaches, behaviors and attitudes, including attitudes and feelings towards refusal, persuasion of reluctant respondents, gaining cooperation, working at different times and days of the week and travel preferences. The questionnaires were filled in prior to working on the sampled cases included in this study. The questionnaire was filled in by the interviewers using a Computer Assisted Self-Interviewing instrument (CASI). Participation in the survey was voluntary and interviewers who participated in the survey were paid one hour for their time. The survey was not anonymous because identifying information was needed to link the

resulting data to other sources. Two reminders were sent. In total about 84% of interviewers replied to the survey.

The main advantage of this study is that these interviewer data have been linked to other sources of information, including the response outcomes of six major household surveys. Detailed information about both responding and nonresponding households has been obtained from the UK 2001 Census for each person in the sample housing units. These data provide socio-economic and demographic characteristics of the individuals living in the sampled households, as well as a wide range of household characteristics. In addition, for every call to the household the interviewer recorded observations made via an interviewer observation (IO) questionnaire, referred to as field-process or paradata (Couper, 1998). These observations include basic information about the household, characteristics of the accommodation, household composition, the quality of housing and observations about the neighborhood, such as presence of children, gender and age of the householder talked to at each contact, presence of physical impediments and if the interviewer feels safe in the area after dark. Area-level data from aggregated census data were also linked and include variables such as population density, percentage of the population living in houses, flats and communal establishments, percentage belonging to a certain age, ethnic or economic activity group or religion, and unemployment rate.

The linkage of the different data sources was carried out by ONS, primarily based on the address of the household, the interviewer id and, if necessary, further identifying information. About 95% of all households were successfully linked to their census record. All linkage was quality assured by ONS based on the distribution of key variables before and after the linkage (for further information see Durrant and Steele 2009; White et al. 2006). Although the data have been expensive to collect they only recently became available for research purposes and further analysis.

The six cross-sectional household surveys included in the study cover a wide range of topics and vary in their design and requirements. The six surveys are: the Expenditure and Food Survey (EFS), the Family Resources Survey (FRS), the General Household Survey (GHS), the Omnibus Survey (OMN), the National Travel Survey (NTS) and the Labour Force Survey (LFS). The survey topics range from expenditure to travel habits. Two of the surveys (EFS and NTS) require keeping a diary and have long questionnaires. Two surveys (OMN and LFS) have comparatively short questionnaires. The surveys also vary in their length of data collection period, interviewer workload and the interviewer training that each interviewer receives prior to the survey. An overview of the differences in the design of the six surveys that may affect cooperation is given in Table 1.

[Table 1]

The outcome of the six surveys, and the dependent variable in our analysis, is defined as an indicator of refusal versus cooperation, conditional on the interviewer having made contact with the household. Household cooperation is defined as obtaining an interview from *at least one* member of the household, including both full (i.e. *all* household members respond) and partial cooperation (i.e. *at least one* but not all respond). Five of the six surveys require that all persons 16 years and older take part in the survey; in the Omnibus survey only one person from each household is selected for interview. (For analysis purposes, we regard this as a special case of full response.) There is substantial variation in refusal rates across surveys from just over 30% for the EFS to about 14% for the LFS, which may be explained by the differences in survey topics, interview length, length of data collection period, and additional requirements such as a diary.

The analysis sample includes all households selected for interview and successfully contacted in one of the six surveys during May-June 2001, the months immediately following the 2001 census. Certain cases have been deleted from the analysis as described in detail in Durrant and Steele (2009), such as non-eligible persons, vacant homes and re-issues (i.e. cases where there was a change in interviewer after the first interviewer had been unsuccessful in gaining cooperation from a household), resulting in a dataset of 17701 contacted households, 564 interviewers and 392 areas. (For exact wording of questions and distribution of the variables included in final multilevel model see Tables A1-A2 of Appendix).

When analyzing interaction effects between the householder and the interviewer ideally one would like to investigate the characteristics of the person at the doorstep. This, however, is difficult since usually this person cannot be identified in the data available. Here, we use information obtained from interviewer observations on gender and approximate age of the person who the interviewer talked to at first contact. Also, to be able to use variables measured at the individual level in our models on household-level nonresponse, we define some variables for the household reference person (HRP) to represent the household as a whole (for further justification see Durrant and Steele, 2009).

The interpretation of interviewer effects may be complicated by their confounding with area effects. In clustered survey designs an interviewer is normally assigned to a primary sampling unit (PSU) and their workload consists of all sampled households in that PSU. Occasionally an interpenetrated design is employed, where interviewers are allocated at random to households, enabling, at least to some extent, a separation of interviewer and PSU effects (O'Muircheartaigh and Campanelli, 1999; Schnell and Kreuter, 2005). More often, however, no such design is employed. Area effects are then usually ignored in analyses or area information is not available (e.g.

Pickery and Loosveldt, 2004). Due to the high costs involved, the surveys included in this study also did not employ interpenetrated sampling designs. Consequently, it is not possible to fully separate interviewer from PSU effects. However, a complete confounding of area and interviewer effects was avoided: most interviewers work on several surveys and some mobile interviewers exist, leading to interviewers working across PSUs; we allow for area effects in our models where areas are defined at the local authority district level, a geographical area slightly larger than a PSU. As a result, interviewers and areas are cross-classified, i.e. an interviewer may work in several areas and an area may be covered by several interviewers. We do not claim to be able to disentangle area and interviewer effects precisely. We argue that the interviewer characteristics of primary interest, such as attitudes and behaviors, are unlikely to be correlated with area characteristics such as population density. (We tested this assumption by examining pairwise correlations between key area and interviewer level variables and found no significant relationships.) Under this assumption, a significant effect of an interviewer-level variable may be interpreted as an interviewer rather than an area effect.

Analysis Methods

A multilevel cross-classified logistic model is used. We are interested in the effects of interviewer characteristics and interviewer-household interactions on the probability of refusal. A multilevel model allows for the correlation in nonresponse probabilities for households allocated to the same interviewer. Using standard regression analysis, which does not account for such clustering, results in underestimation of standard errors, which in turn leads to overstatement of the statistical significance of effects. The downward bias in standard errors is especially severe for coefficients of higher-level variables, interviewer characteristics in the present case. A cross-classified model

accommodates the effect of more than one type of nesting which occurs at the *same* level (here interviewers and areas), allowing for the cross-classification of areas and interviewers (see Goldstein, 2003). Other authors have used similar multilevel models for the analysis of interviewer effects on nonresponse (Pickery et al., 2001; Hox, 1994; O'Muircheartaigh and Campanelli, 1999).

We denote by $y_{i(jk)}$ the outcome for household i contacted by interviewer j in area k, where the cross-classification of interviewers and areas is indicated by placing their indices in parentheses. The dependent variable is coded as

$$y_{i(jk)} = \begin{cases} 0 & \text{cooperation} \\ 1 & \text{refusal.} \end{cases}$$

Denoting the probability of refusal by $\pi_{i(jk)} = \Pr(y_{i(jk)} = 1)$, and taking cooperation as the reference category, the multilevel cross-classified logistic model for refusal can be written as

$$\log\left(\frac{\pi_{i(jk)}}{1-\pi_{i(jk)}}\right) = \boldsymbol{\beta}^T \boldsymbol{x}_{i(jk)} + u_j + v_k, \qquad (1)$$

where $\boldsymbol{x}_{i(jk)}$ is a vector of household, interviewer and area-level covariates and interactions, $\boldsymbol{\beta}$ is a vector of coefficients and u_j and v_k are random effects, representing unobserved interviewer and area effects respectively. The random effects are assumed to follow normal distributions, i.e. $u_j \sim N(0, \sigma_u^2)$ and $v_k \sim N(0, \sigma_v^2)$. The variance parameters σ_u^2 and σ_v^2 are respectively the residual between-interviewer and between-area variances in the log-odds of refusal versus cooperation. Variables that were not statistically significant at least at the 10% level were removed from the model. Due to the large number of available predictors and possible interaction terms we restricted our analysis to terms of scientific interest as informed by the theoretical substantive framework. The models were estimated using Markov chain Monte Carlo (MCMC) methods in MLwiN (Rasbash et al. 2008b; Browne, 2008) with non-informative (also known as flat or diffuse) priors. The parameter estimates and standard errors are the

means and standard deviations of 80,000 chains, after a burn-in of 5000 (Browne, 2008). To aid interpretation we calculated predicted probabilities, using a simulation approach to average over the interviewer and area random effect distributions (Rasbash et al., 2008a).

Results

Between-Interviewer Variance in Survey Participation

Table 2 summarizes estimates of the interviewer and area random effect variances from alternative specifications of the multilevel models for refusal. All models include dummy variables for survey to account for design differences among the six surveys. The table also shows the DIC (deviance information criterion) diagnostic which can be used for model comparison, with a smaller DIC indicating a better fit (Spiegelhalter et al. 2002).

[Table 2]

Comparing Model 0, including only survey effects and no interviewer or area random effects, with the same model including either interviewer (Model 1a) or area (Model 1b) random effects, we see that the DIC is reduced by 121 and 67 respectively, suggesting between-interviewer and to a lesser extent between-area variation in cooperation rates. For both models the interviewer and area variances are significant with the area variance just over half the size of the interviewer variance. We then fitted a cross-classified model that accounts for interviewer and area effects simultaneously (Model 2), reducing the DIC by 7 in comparison to the model with only interviewer effects (Model 1a). The interviewer variance is still highly significant and about three times larger than the area variance, which is only marginally significant. The standard deviation (std) of the interviewer effect is estimated as $\sqrt{0.077} = 0.28$. We can say that

an increase of 1 std in the unobserved interviewer characteristics is associated with a 32% increase in the odds of a refusal.[1]

Including household-level characteristics (Model 3) reduces the DIC by a further 415 (=18338-18753) and halves the area variation, now no longer significant, suggesting that household characteristics explain most of the area variation, as would be expected (O'Muircheartaigh and Campanelli, 1999; Schnell and Kreuter, 2005). The interviewer variation appears almost unaffected by the addition of household-level variables. Including interviewer-level variables, their interactions with household characteristics and the survey indicators (Model 4) explains about half of the interviewer variation. Nevertheless, there remains a small amount of unexplained significant interviewer variation. This compares to findings in O'Muircheartaigh and Campanelli (1999) who found interviewer effects were no longer significant once the effects of interviewer-level variables were controlled. The area variance is unaffected by the introduction of the interviewer-level variables, suggesting that we are indeed explaining interviewer variation rather than area variation with the interviewer-level characteristics. The final model (Model 5) also accounts for area variables, which have virtually no effect on either the interviewer-level variance or the DIC diagnostic and reduce the area variability by only a small amount. The interpretation of the final model is now discussed.

What are the Attributes of Interviewers that Affect Cooperation Rates?

There are three streams of research that have focused on the mechanisms underlying interviewer effects on cooperation rates. The earliest (Durbin and Stuart, 1951) observed that more experienced interviewers, probably through skill acquisition, achieved higher cooperation rates. The second, compatible with the first, observed that independent of

[1] $\exp(0.28) = 1.32$, so a 1-unit (standard deviation) increase in u_j is associated with a 32% increase in the odds of a refusal.

experience, perceived confidence in one's abilities allowed interviewers to perform better (Groves and Couper, 1998). Finally, several works suggest that high-cooperation rate interviewers are focused on respondent concerns (Morton-Williams, 1993) and tailor their introductory behaviour to individual respondents (Groves and Couper, 1998). The weakness of the past research, however, is that it generally failed to statistically control for attributes of the sample households, risking confounding between interviewer and householder effects. Our data allow us to explore simultaneously each of these three streams of research and to control for household and area characteristics. We describe the findings of the final model (Model 5), presented in Tables 3-4.

[Tables 3-4]

The influence of the household-level variables in Model 5, while focused on the interviewer, can be summarized as follows (Durrant and Steele, 2009): refusal is higher for a household in London, urban areas, which did not move during the last year, has no car, is occupied by a single person, had a household reference person of lower education, is self-employed and regards his/her health as not good. On the other hand, refusals are lower in households with children and caregivers. We also found interviewer observations to be significant and two were included in the final model: the refusal probability is higher if the house is in a worse condition than others in the area and if the householder at first contact is male. After controlling for household characteristics, only one area-level variable remained significant - the percentage of the population 0-4 years old, showing a positive effect on refusal. We now turn to the interpretation of the three sets of interviewer influences: experience, confidence in abilities and behavior.

Interviewer experience and pay grade

A common finding in the literature is that refusal rates decrease with increasing length of experience (Groves and Couper, 1998; Pickery and Loosveldt, 2002; Hox and DeLeeuw, 2002). Some of these studies used interviewer corps with relatively few long-tenured interviewers. A potential limitation of research on interviewers is the problem of self-selection. We may expect better interviewers to stay in their job for longer and worse performing interviewers to change to a different job sooner. We are therefore limited in the interpretation of causal effects of e.g. length of experience; only an experimental design may be able to disentangle such effects.

When experience is the only interviewer-level variable in our model, we also found it to predict lower refusal rates for more experienced interviewers. However, after controlling for the effect of pay grade, a different relationship between experience and the refusal rate emerges. With that statistical control, interviewers who have been in the job for 9 years or more seem to perform significantly less well than those with less experience. There is also an indication that the probability of refusal declines after 1-2 years experience (although the effect is not significant). This curvilinear relationship between performance and length of experience has been hypothesized in the literature but has not before been supported by empirical evidence (Groves and Couper, 1998). We interpret this as a potential disentangling of two underlying processes: most interviewers receive jumps in pay grade as they gain experience. Measuring only the effects of experience confounds these two processes. Skill level, reflected in pay grade, appears to be the real underlying mechanism driving cooperation rates, not the simple length of time employed.

We also find that, after controlling for pay grade and experience, interviewers who have worked 5 or more hours per day in the previous year have lower refusal rates than those who did not work or worked fewer hours, supporting a potential role of

interviewer practice. It is common for survey organizations to provide larger workloads to higher-performing interviewers, and thus this finding may be endogenous to refusal rates.

Interviewer confidence and attitudes

The second set of prior research examined whether interviewer confidence, attitudes and expectations affect response propensities. Prior work on interviewer attitudes has been limited to bivariate or interviewer-level analyses, usually with the interviewer-level cooperation rate as the dependent variable (Groves and Couper, 1998; Hox and DeLeeuw, 2002). In our analysis of household-level response, we find important effects of interviewer attitudes on cooperation rates. Particularly, we find a strongly significant effect of interviewer confidence and attitude towards persuasion of reluctant respondents, both measured independently of the survey in question. Interviewers who report more confidence in their ability to persuade reluctant respondents, who believe they can persuade when others cannot and who disagree with the statement 'no matter what I do some respondents will never agree to participate' show a lower probability of refusal. Interviewers who agree they should persuade reluctant respondents also have a lower refusal rate than interviewers that disagree. We found some indication that if interviewers believe refusal is due to something they did, refusal rates are higher. This finding may also indicate that interviewers who are less confident about their behavior may be less successful - although not significant in the final model. These findings on confidence show the important role of positive expectation (Groves and Couper, 1998). It may indicate interviewers who believe in themselves and their techniques may be able to persuade. The results also stress the importance of being willing and able to persuade reluctant respondents, with interviewers who feel they can and should persuade being more successful.

Interviewers who report that a refusal affects their behavior have lower refusal rates. Rather than interpreting this as an effect of confidence, it may be the case that these interviewers take the experience of a refusal as an opportunity to change their behavior and to adapt to new demands, leading to an improved performance. This may indicate an ability to tailor and to adjust the interviewer's technique to the sample unit. Interviewers were asked whether they supported 'sending a different interviewer if the respondent refused.' Those who support that policy tended to have lower refusal rates. This may indicate that interviewers who do not wish interference from other interviewers and who are possibly less open-minded towards another person's skills are more likely to experience higher refusal rates.

The attitude of the interviewer to travelling longer distances, staying overnight and working evenings and weekends were significant in initial modeling but not in the final model. We may conclude that more persistent interviewers, interviewers who believe in themselves, who are confident and feel able to persuade reluctant respondent are successful at gaining cooperation.

Interviewer behaviors, strategies and approaches at the doorstep

Our measures of interviewer behaviors are self-reports from the questionnaire, administered independent of their working on the survey in question, asking interviewers about what they *generally* do, rather than being recorded at the contact level. It is clear from the literature that such measures may not fully reflect the tendency for successful interviewers to vary their behaviors across respondents (Morton-Williams, 1993). Thus, we focused particularly on variables that may indicate the ability and willingness to tailor the approach to each household and on variables representing doorstep approaches. The following were used as indicators of tailoring ability: ability to deal with everybody in the same manner, use of the same/different introduction, complimenting respondents about

their house, the belief that respondents need unique approach, belief in changing approach, use of wide variety of approaches, finding it difficult to modify approach from situation to situation. Variables related to doorstep techniques include: waiting to explain survey, saying that topic should interest and that not a salesperson, indicating most people enjoy interview, the survey is a chance to express views, and explaining methods how sample members were selected. Although we found significant effects of some of these variables, most lost significance when other interviewer-level variables were added or had non-interpretable main effects. For example, interviewers who alter their introduction to fit each household they visit, who compliment the householder or who do not think they can deal with everybody in the same manner seem to have lower refusal rates, when variables were entered each on their own. However, all of these variables were non-significant in the final model.

Groves and Couper (1998) argue that main effects of interviewer behavior may be unlikely because it is not whether certain strategies are adopted *in general* that is important, but whether strategies are *tailored* towards sample units. We believe that progress on identifying the specific behavioral mechanisms requires measurement at the conversational level (Sturgis and Campanelli, 1998; van der Vaart et al. 2005; Couper and Groves, 2002; Maynard and Schaeffer, 2002). Rather than focusing on main effects of variables that may be interpreted as indications of tailoring it may be more important to investigate statistical interactions between interviewer and household characteristics. Given these results, we note that the measured effects of interviewer experience and attitudes may be confounded with real unmeasured behavioral differences. In contrast to past work, however, they do control for household-level attributes.

Are different interviewer attributes important for the participation of different sample persons?

An important value of the data is that they permit examining a host of as yet unanswered questions about whether some interviewer behaviors and attributes stimulate cooperation among some respondents but not others. In short, do some sample units react favorably to interviewer approaches that others would reject? Most of the past literature on this stems from one premise: interviewers and respondents sharing attributes might tend to produce higher cooperation rates. Due to data limitations, there has been little exploration of the statistical interactions between interviewers and householders in previous research (Groves and Couper, 1998; O'Muircheartaigh and Campanelli, 1999; Pickery and Loosveldt, 2002, 2004). One of the advantages of this study is that its rich information on interviewers, linked to household characteristics, permits such an analysis.

Socio-Demographic Interviewer Characteristics

We first test whether homogeneity between sample members and interviewers may result in higher probability of cooperation. The presence of such interactions may suggest ways of tailoring interviewing strategies for particular types of respondents. Information on such interactions may be used to match interviewers to sample units.

We find an interaction effect (significant at the 10% level) between the gender of the person at first contact and that of the interviewer. Female householders are more likely to respond than men if the interviewer is also female, while interviewer gender has no effect among male sample units (Table 4). We also found an effect (significant at the 5% level) in the same direction between gender of the HRP and that of the interviewer (effect was not additionally included in model). These findings may be explained by a potential fear of crime of a woman towards a male stranger. It could also be explained by

the *theory of liking* (Groves et al., 1992), which hypothesizes that people are favorably inclined towards those who they like or have something in common with, such as similar characteristics or attitudes.

We found no significant interaction effect between age of the interviewer and householder at first contact (using estimated age group of the householder obtained from interviewer observations). It should be noted that it is more difficult to estimate age group for the person at first contact than it is to record gender, resulting in a higher proportion of 'age not known'. This measurement problem may be one reason for not finding a significant age effect.

We also investigated interaction effects based on educational level. Since we do not have this information for the person at the doorstep we use the variable education of the HRP who is representative for the household as a whole. We can see that the refusal rate is significantly lower in the case when both the householder and the interviewer have no or only a low educational attainment as well as when both have a college degree (Table 4). The refusal rate is higher for the case when the interviewer has a low educational level but the householder has a college degree. The probability is even higher if the interviewer has a college degree but the householder does not, and is highest for the case where the interviewer has only a low or no educational attainment and the householder has a professional degree of some form. This effect indicates that sample members may be favorably inclined towards those who they have something in common with, which may support the theory of liking. In short, the interaction effects found provide evidence for the notion of similarity of interviewers and respondents generating higher cooperation. This may support the practical decision of survey managers to match interviewers with sample households if possible.

Influence of interviewer characteristics and self-awareness on persuading difficult households

The survey community has found (and our model supports), that certain types of households are more difficult to persuade, such as single-person households, households without children, households with unemployed persons. It is of interest to investigate the characteristics of interviewers that are more successful in persuading such difficult cases. In particular, we tested if level of confidence and experience of the interviewer had different effects on securing cooperation of more difficult households. The interviewer questionnaire also asked if the interviewer believed certain types of persons were more difficult to persuade to participate in surveys. We thus could also investigate the effect of self-awareness of the interviewer.

We found some support for these hypotheses. Interviewers who report they are more confident in persuading reluctant respondents are more likely to be successful in persuading households without children, a group of households that generally is more difficult to persuade (Table 4; significant at 10% level). We also found some effects of self-awareness of interviewers; for example if the interviewer reported finding it difficult to persuade households with children and the household has children the refusal rate was indeed higher (effect not included in the final model). In this case the interaction effect confirmed the self-awareness of the interviewer's difficulty in persuading a certain type of household. The findings may have implications for survey practice. It may be possible to allocate certain interviewers to more difficult cases, for example when re-issuing the case to a different interviewer, as part of responsive designs or a follow-up study.

Do Interviewer-Level Influences on Cooperation Depend on Survey Designs?

There is some evidence from studies of interviewer variance in responses, that interviewer effects arise most prominently when the challenge of the job becomes difficult (Mangione et al. 1992). The data set is uniquely well-suited to examining

whether the strong correlates of high interviewer cooperation rates are active for all six surveys. The hypothesis that would arise directly from the research literature is that survey designs that make less burdensome requests of respondents do not require the skills of the best interviewers. Indeed, it is common for interviewer managers to start off a new hire on "easy" cases or surveys when possible.

We considered a number of interactions of interviewer characteristics with the survey indicators. In particular we tested for survey-specific effects of a) socio-demographic characteristics of interviewers, such as age and gender; b) interviewer strategies and techniques, including style of doorstep approach such as use of the same or different introduction for each household, the ability to deal with everybody in the same manner, and what the interviewer says or does at the doorstep; and c) interviewer experience, interviewer education, experience of working for another survey organization and having another employment. In particular, effects c) were considered to explore the hypothesis that the level of experience and interviewer education are of greater importance in more complex surveys, such as those with a long questionnaire and a diary.

We found a survey-specific effect of how confident interviewers feel they are in persuading reluctant respondents (Table 4). For surveys that are more complex, i.e. have a more demanding survey topic, have a longer interview or require a diary, the level of confidence of the interviewer is important. Interviewers who feel more confident do significantly better than less confident interviewers in more complex surveys, such as the EFS with a long questionnaire, questions about income and expenditure and a two-week diary. For a less demanding survey, such as the LFS with only a short interview, the level of confidence does not lead to significant differences in the response propensities between more and less confident interviewers.

There are potentially important practical implications of this finding. It supports assertions that some survey protocols do not demand as much training on respondent recruitment issues as others and implies that allocation of experienced staff can be made more effectively.

Summary and Implications for Survey Practice

This paper has found empirical support for several key mechanisms through which interviewers affect survey cooperation rates:

a) We identified significant effects of pay grade and interviewer experience. While more senior interviewers tend to achieve higher cooperation rates, their skills recognized through promotion to higher pay grades seem to be more critical in understanding the effect; controlling on pay grade we actually find a decline in performance after 9 years in the job.

b) We found pervasive effects of interviewer confidence and attitudes; if interviewers express confidence in their abilities, they achieve higher cooperation rates; interviewers with a positive attitude towards persuasion or those who agree they should persuade reluctant respondents tend to have lower refusal rates. In short, confident, engaged interviewers seem to perform better.

c) Like other studies, we find little predictive power from reports *at the interviewer-level* of interviewer behavior and interviewer ability to adapt to different situations; we view this as support for the notion that behavioral impacts on cooperation rates depend on specific features of interactions with respondents (Maynard and Schaeffer, 1997).

d) Our unique data allowed the investigation of interaction effects between interviewers and householders. We found some evidence for the hypothesis that interviewers and respondents sharing attributes may tend to produce higher cooperation rates. In particular, we found that female householders seem to be more cooperative with

female interviewers than with male interviewers and that when interviewer and householder share educational backgrounds, higher response propensities exist.

e) We also investigated the influence of certain interviewer attributes on persuading more difficult cases. We found that confident interviewers did better than those with less confidence among households without children, a group commonly producing lower cooperation rates; similarly the more confident interviewers did better on the more difficult surveys. Both of these findings are consistent with the notion that, when the interviewing job gets difficult, the more confidence the interviewer has, the better the performance.

Although the illumination of these findings required complicated statistical modeling, there are practical implications for the field, for example regarding interviewer training and allocation of interviewers to sample units. Simply using the most experienced interviewers as a tool to maximize cooperation rates may be myopic; it would be better to examine the actual performance of interviewers including historical performance (in these data measured by pay grade), since some long-term interviewers may perform less well than those with less experience.

More confident interviewers seem to be most valuable on the more difficult surveys with simpler surveys showing a smaller effect of interviewer differences. The importance of confidence and a positive attitude towards persuasion also has implications for interviewer training. How interviewers assess their own abilities seems to make a difference, confidence and related attitudinal states measured independently of the survey performance seems to predict that performance. Although it may not be possible to completely disentangle whether interviewer confidence and positive expectation are cause or outcome of refusal probability, we control for two indicators of historical performance and previous success experience: interviewer pay grade, as a

measure of the skills and previous performance, and interviewer experience. The significant effects of both interviewer confidence and attitude can then be interpreted as net effects, indicating that, after controlling for pay grade and experience, these variables have indeed a positive influence on cooperation rates. The findings therefore support the idea of enhancing the self-image and a positive expectation during interviewer training. These two characteristics may also be already measurable at the recruitment stage, which may help the selection of interviewers. The effects of confidence and positive attitude may nevertheless be partially the result of previous successful experience of interviewers. Facilitating such a positive experience early on, e.g. during training, may therefore be beneficial in increasing interviewer confidence and positive expectation and may help to increase cooperation rates longer term.

Our findings on interviewer-respondent interaction effects, such as on gender and education, suggests the matching of interviewer characteristics to different subgroups of the population. In practice, it may be possible to match also on other characteristics such as age and ethnicity to achieve cooperation (see also Durrant et al., 2009). This may be of particular relevance for the design of interviewer call-backs, re-issues and follow-ups, such as in responsive survey designs, which are currently explored in some UK surveys. For example, in the case of an initial non-compliance by a sample person another more suitable interviewer may approach the household to persuade them, e.g. a female householder is approached by a female interviewer, which seems to be of particular importance amongst households from certain ethnic backgrounds, such as Asians (Durrant et al., 2009). Knowledge about the household may also inform the way the interviewer approaches the household and the interviewer behavior at the next call. Interviewer supervisors may have known or suspected some of the above, but our findings suggest that they may have more widespread applicability.

This research has identified gaps in current knowledge. The influences of interviewer behaviour as well as interviewer personality traits are not yet well understood. It seems advisable to measure interviewer behavior at the interaction level rather than the interviewer level. To understand the process of establishing cooperation better, interviewer call records need to be investigated, which only more recently have become available. It also seems advisable to control for previous interviewer performance which requires survey agencies to record and use these data. A largely unexplored area is interviewer effects in longitudinal surveys. Further research is currently under way to address these shortcomings.

Appendix:

Table A1: Wording of original question(s) from which explanatory variables included in the final multilevel model were derived.

Abbreviations as used in tables	Wording of questions in questionnaires
Variables from the 2001 UK Census form	
Educational attainment	Which of these qualifications do you have? 1+ O Levels/CSEs/GCSEs (any grades). 5+ O Levels, 5+ CSEs (grade 1), 5+ GCSEs (grades A-C), School Certificate. 1+ A levels/AS levels. 2+ A levels, 4+ AS levels, Higher School Certificate. First Degree (eg BA, BSc). Higher Degree (eg MA, PhD, PGCE, post-graduate certificates/diplomas). NVQ Level 1, Foundation GNVQ. NVQ Level 2, Intermediate GNVQ. NVQ Level 3, Advanced GNVQ. NVQ Levels 4-5, HNC, HND. Other Qualifications (eg City and Guilds, RSA/OCR, BTEC/Edexcel. No qualifications.
Household has no dependent children	This variable was derived from information provided on who lives in the household and their date of birth. Relevant information collected includes: 'List all members of your household who usually live at this address including yourself.' The form then asks for the household members and their relationships within the household (e.g. name of person 2 and relationship to person 1). The form then asks personal information for everyone listed, including 'What is your date of birth?' and 'Are you a school-child or student in full-time education?'.
London indicator	Geographical variables were derived from postcode information.
Rural indicator	Geographical variables were derived from postcode information.
Economic Activity	Last week, were you doing any work: as an employee, or as a trainee on a government sponsored training scheme, as a self-employed/freelance, or in your own/family business? If YES: Do (did) you work as an employee or are (were) you self-employed? Employee, Self-employed with employees, Self-employed/freelance without employees. If NO: Last week, were you any of the following? Retired, Student, Looking after home/family, Permanently sick/disabled, None of the above.
Perception of health	Over the last twelve months would you say your health has on the whole been: Good? Fairly good? Not good?
Caregiver in household	Do you look after, or give any help or support to family members, friends, neighbors or others because of long-term physical or mental ill-health or disability, or problems related to old age? No, Yes (1-19 hours a week), Yes (20-49 hours a week), Yes (50+hours a week).
Household type	This variable was derived based on the information given about the people in the household. The information collected included: 'List all members of your household who usually live at this address including yourself.' The form then asks for the household members and their relationships within the household (e.g. name of person 2 and relationship to person 1, where possible relationships include husband or wife, partner, son or daughter, step-child, brother or sister). The form then asks personal information for everyone listed, including 'What is your date of birth?'

Car Ownership	How many cars or vans are owned, or available for use, by one or more members of your household? None, One, Two, Three, Four or more
Household moved during last year	What was your usual address one year ago? The address shown on the front of the form. No usual address one year ago. Elsewhere.
Variables from the Interviewer Observation (IO) Questionnaire	
Gender of householder at first contact	Was the main person you talked to: A man/boy, a woman/girl, don't know/not sure (This information is recorded for every call made to the household, provided contact was established.)
House in a better or worse condition than others in area	Is the sampled house/flat/building in a better or worse condition than the others in the area? Better, Worse, About the same, Unable to code
How safe would you feel walking alone in this area after dark?	How safe would you feel walking alone in this area after dark? Very safe, Fairly safe, A bit unsafe, Very unsafe
Variables from the Interviewer Attitude Survey (IAS)	
Interviewer gender	Are you: Male, Female
Interviewer educational attainment	What is the highest educational qualification you have obtained? Higher degree and postgraduate qualifications Degree, or degree level equivalent Other Higher Education below degree level A Levels, vocational level 3 and equivalent O Levels or GCSE grade A-C & equivalents Qualifications below the above Trade Apprenticeships/Secretarial and Commercial Qualifications Other qualifications-level unknown No Qualifications
Pay grade	What is your current SSD (Social Survey Division) pay grade? Interviewer, Advanced Interviewer, Merit 1, Merit 2, Merit 3, Field Manager
Years of experience	How many years have you worked as an interviewer for SSD? (If under 1 year please enter 0. If one or more years please round to the nearest year)
Daily hours previous year weekdays	In the year from 1 April 2000 to 31 March 2001, what was the average number of hours a day you spent on interviewing on a weekday (Monday to Friday)? (By this we mean time spent in the field) (0…24) Please round to the nearest hour.
Should persuade most reluctant respondents	Interviewers should always try to persuade even the most reluctant respondents to participate. Strongly agree, agree, neither agree nor disagree, disagree, strongly disagree
Can persuade when others can't	I can persuade the people whom most other interviewers cannot persuade. Strongly agree, agree, neither agree nor disagree, disagree, strongly disagree
Can convince reluctant respondents	With enough effort I can convince even the most reluctant respondents to participate. Strongly agree, agree, neither agree nor disagree, disagree, strongly disagree
Refusal affects how behave	If you have experienced a refusal, how often would you say it affects how you behave at the next household in your assignment? Always, frequently, sometimes, rarely, never
No matter what I do, some respondents will never agree to participate	No matter what I do, there are some respondents who will never agree to participate. Strongly agree, agree, neither agree nor disagree, disagree, strongly disagree
If respondent refused because too busy it is better to send a different interviewer	When a respondent refused because (s)he is too busy, it is better to send a different interviewer than have the same interviewer return. Strongly agree, agree, neither agree nor disagree, disagree, strongly disagree

Table A2: Percentage distribution within type of response status for explanatory variables included in the final multilevel model (household, area and interviewer level). †

Variable	Categories	Cooperation (%)	Refusal (%)	Total (%)
Household-Level Variables		**n=13604**	**n=4097**	**n=17701**
Survey indicator	EFS	18.1	27.4	20.3
	FRS	11.6	13.1	11.9
	GHS	19.3	16.0	18.5
	OMN	16.6	17.8	16.9
	NTS	14.6	14.6	14.6
	LFS	19.8	11.1	17.8
Educational attainment (HRP)	No educational attainment	27.5	32.6	28.7
	O/A levels, GCSEs (UK school degrees)	38.8	33.4	37.6
	First/Higher/University degree	16.6	13.0	15.8
	Other educational attainment	5.6	5.9	5.6
	Missing	11.5	15.1	12.3
Household has no dependent children	Household has dependent children	31.8	25.6	30.4
	No dependent children	68.2	74.4	69.6
London indicator	Not London	90.2	86.6	89.4
	London	9.8	13.4	10.6
Rural indicator	Urban	88.3	90.7	88.8
	Rural	11.0	8.9	10.5
	Missing	0.7	0.3	0.7
Economic Activity (HRP)	Employee	51.3	45.6	49.9
	Self employed	8.8	10.4	9.2
	Unemployed	2.2	2.6	2.3
	Retired	16.9	16.5	16.8
	Looking after family	2.8	2.3	2.7
	Other (incl. student, permanently sick etc.)	6.5	7.5	6.8
	Missing	11.5	15.1	12.3
Perception of health (HRP)	Good	60.0	54.5	58.7
	Fairly good	28.2	31.7	29.0
	Not good	11.8	13.8	12.3
Caregiver in household	No	80.9	82.7	81.3
	Yes	19.1	17.3	18.7
Household type	Single	33.9	35.8	34.3
	Couple	55.6	50.8	54.6
	Multiple	10.5	13.4	11.1
Car Ownership	One car or more	75.2	70.3	74.0
	No car	24.8	29.7	26.0
Household moved during last year	No	92.0	94.0	92.5
	Yes	8.0	6.0	7.5
Interviewer Observation Variables (on household level)		**n=13604**	**n=4097**	**n=17701**
Gender of householder at first contact	Male	40.2	42.0	40.6
	Female	58.4	56.7	58.0
	Don't know/ not sure	1.3	1.3	1.3
House in a better or worse condition than others in area	Better	10.8	9.4	10.5
	Worse	6.4	8.5	6.8
	About the same	82.2	79.0	81.5
	Unable to code	0.6	3.1	1.2

† HRP= information based on household reference person

Variable	Categories	Total (%)
Area-Level Variable		**n=392**
Percentage of population between 0 and 4 years	Continuous variable (mean)	5.8
Interviewer-Level Variables		**n=564**
Interviewer gender	Male	60.1
	Female	39.9
Interviewer educational attainment	Degree or postgraduate, University degree	44.6
	Academic attainment below University degree (O/A levels, GCSEs)	51.2
	Lower, no, other educational attainment	4.8
Pay grade	Interviewer	46.8
	Advanced interviewer and merit 1 and 2	34.3
	Merit 3 and field manager	18.2
	Missing	0.5
Years of experience	Less than 1 year	20.6
	1 to 2 years	27.3
	3 to 8 years	27.7
	9 years or more	25.0
Daily hours previous year weekdays	0-4 hours	9.4
	5 and more hours	83.8
	Missing	6.7
Should persuade most reluctant respondents	Strongly agree, Agree	77.3
	Neither agree nor disagree	8.5
	Disagree, strongly disagree	14.2
Can persuade when other can't	Strongly agree, Agree	9.3
	Neither agree nor disagree	58.5
	Disagree, strongly disagree	29.2
	Missing	3.0
Can convince reluctant respondents	Less confident	83.3
	More confident	16.8
Refusal affects how behave	Rarely, never	77.7
	Always, frequently, sometimes	22.3
No matter what I do, some respondents will never agree to participate	Strongly agree, Agree	95.7
	Strongly disagree, disagree, neither agree nor disagree	4.3
If respondent refused because too busy it is better to send a different interviewer	Strongly agree, Agree	9.6
	Strongly disagree, disagree, neither agree nor disagree	90.4

References

Bailey, Leroy, Thomas F. Moore and Barbara A. Bailar. 1978. "An Interviewer Variance Study for the Eight Impact Cities of the National Crime Survey Cities Sample." *Journal of the American Statistical Association.* 73, 361: 16-23.

Browne, William J. 2008. *MCMC Estimation in MLwiN* v2.10. Centre for Multilevel Modelling. University of Bristol: Bristol.

Couper, Mick P. 1998. "Measuring Survey Quality in a CASIC Environment." *Proceedings of the Survey Research Methods Section. American Statistical Association.* 48:743-772.

Couper, Mick P. and Robert M. Groves. 2002. "Introductory Interactions in Telephone Surveys and Nonresponse", in: Douglas W. Maynard, Hanneke Houtkoop-Steenstra, Nora C. Schaeffer, Johannes van der Zouwen (eds). 2002. "*Standardization and Tacit Knowledge: Interaction and Practice in the Survey Interview*", New York, 2002, Wiley, 161-178.

Durbin, J. and Stuart, A. 1951. "Differences in Response Rates of Experienced and Inexperienced Interviewers." *Journal of the Royal Statistical Society, Series A*, 114, 1: 163-206.

Durrant, Gabriele B. and Fiona Steele. 2009. "Multilevel Modelling of Refusal and Noncontact Nonresponse in Household Surveys: Evidence from Six UK Government Surveys." *Journal of the Royal Statistical Society. Series A.* 172, 2, 361-381.

Durrant, Gabriele B., Harvey Goldstein, Laura Staetsky and Shaun Scholes. 2009. "Analysis of Attrition in the Family and Children Survey (FACS)" (unpublished working paper).

Feldman, J. John, Herbert Hyman and Catherine W. Hart. 1951-52. "A Field Study of Interviewer Effects on the Quality of Survey Data." *Public Opinion Quarterly.* 15, 4: 734-761.

Flores-Macias, Francisco and Chappell Lawson. 2008. "Effects of Interviewer Gender on Survey Responses: Findings from a Household Survey in Mexico." *International Journal of Public Opinion Research*. 20, 1: 100-110.

Fowler, Floyd J. and Thomas W. Mangione. 1990. *Standardized Survey Interviewing; Minimizing Interviewer-Related Error*. Newbury Park: Sage.

Freeman, John and Edgar W. Butler. 1976. "Some Sources of Interviewer Variance in Surveys." *The Public Opinion Quarterly*. 40, 1: 79-91.

Freeth, Stephanie, Catherine Kane and Allison Cowie. 2002. "Survey Interviewer Attitudes and Demographic Profile, Preliminary Results from the 2001 ONS Interviewer Attitudes Survey." *Office for National Statistics Working Paper*. London: 1-18.

Goldstein, Harvey. 2003. *Multilevel Statistical Models*. third edition. London: Edward Arnold.

Groves, Robert M. and Emilia Peytcheva. 2008. "The Impact of Nonresponse Rates on Nonresponse Bias: A Meta-Analysis." *Public Opinion Quarterly*. 72, 2: 167-189.

Groves, Robert M. and Mick P. Couper. 1998. *Nonresponse in Household Interview Surveys*. New York: John Wiley.

Groves, Robert M. and Nancy Fultz. 1985. "Gender Effects among Telephone Interviewers in a Survey of Economic Attitudes." *Sociological Methods & Research*. 14, 1: 31-52.

Groves, Robert M., Robert B. Cialdini and Mick P. Couper. 1992. "Understanding the Decision to Participate in a Survey." *Public Opinion Quarterly*. 56: 475-495.

Hansen, Kasper 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, 1: 112-121.

Hanson, Robert H. and Eli S. Marks. 1958. "Influence of the Interviewer on the Accuracy of Survey Results." *Journal of the American Statistical Association*. 53, 283: 635-655.

Hox, J. Joop. 1994. "Hierarchical Regression Models for Interviewer and Respondent Effects." *Sociological Methods & Research*. 22, 3: 300-318.

Hox, J. Joop. and Edith De Leeuw. 2002. "The Influence of Interviewers' Attitude and Behavior on Household Survey Nonresponse: An International Comparison." in: Robert M. Groves, Don A. Dillman, John L. Eltinge and Roderick J.A. Little. 2002. *Survey Nonresponse*, New York: 103-119.

Kane, Emily W. and Laura J. Macauley. 1993. "Interviewer Gender and Gender Attitudes." *Public Opinion Quarterly*. 57, 1: 1-28.

Keeter, Scott, Carolyn Miller, Andrew Kohut, Robert M. Groves and Stanley Presser. 2000. "Consequences of Reducing Nonresponse in a National Telephone Survey." *Public Opinion Quarterly*, 64: 125-48.

Killias, Martin and Christian Clerici. 2000. "Different Measures of Vulnerability in Their Relation to Different Dimensions of Fear of Crime." *The British Journal of Criminology*. 40: 437-450.

Mangione, Thomas, Floyd J. Fowler and Thomas A. Louis. 1992. "Question Characteristics and Interviewer Effects." *Journal of Official Statistics*. 8, 3: 293-307.

Maynard, Douglas W. and Nora C. Schaeffer. 1997. "Keeping the Gate: Declinations to the Request to Participate in a Telephone Survey Interview." *Sociological Methods and Research*. 26:34-79.

Maynard, Douglas, W. and Nora C. Schaeffer. 2002. "Refusal Conversation and Tailoring", in: Douglas W. Maynard, Hanneke Houtkoop-Steenstra, Nora C. Schaeffer, Johannes van der Zouwen (eds). 2002. "*Standardization and Tacit*

Knowledge: Interaction and Practice in the Survey Interview", New York, 2002, Wiley, 219-242.

Moorman, Patricia G., Beth Newman, Robert C. Millikan, Chiu-Kit J. Tse and Dale P. Sandler. 1999. "Participation Rates in a Case-Control Study: The Impact of Age, Race, and Race of Interviewer." *Annals of Epidemiology*. 9, 3: 188 -195.

Morton-Williams, Jean. 1993. *Interviewer Approaches*. Aldershot: Dartmouth.

O'Muircheartaigh, Colm and Pamela Campanelli. 1999. "A Multilevel Exploration of the Role of Interviewers in Survey Nonresponse." *Journal of the Royal Statistical Society. Series A*. 162, 3: 437-446.

Pickery, Jan and Geert Loosveldt. 2002. "A Multilevel Multinomial Analysis of Interviewer Effects on Various Components of Unit Nonresponse." *Quality and Quantity*. 36: 427-437.

Pickery, Jan and Geert Loosveldt. 2004. "A Simultaneous Analysis of Interviewer Effects on Various Data Quality Indicators with Identification of Exceptional Interviewers." *Journal of Official Statistics*. 20, 1: 77-89.

Pickery, Jan, Geert Loosveldt and Ann Carton. 2001. "The Effects of Interviewer and Respondent Characteristics on Response Behavior in Panel Surveys." *Sociological Methods and Research*. 29, 4: 509-523.

Rasbash, Jon, Chris Charlton and William J. Browne. 2008b. *MLwiN v2.10*. Centre for Multilevel Modelling, University of Bristol.

Rasbash, Jon, Fiona Steele, William J. Browne and Harvey Goldstein. 2008a. *A User's Guide to MLwiN version 2.10*. Centre for Multilevel Modelling. University of Bristol: Bristol.

Schnell, Rainer and Frauke Kreuter. 2001. "Separating Interviewer and Sampling Point Effects." *Journal of Official Statistics*. 21, 3: 389-410.

Spiegelhalter, David J., Nicky G. Best, Bradley P. Carlin and Angelika van der Linde. 2002. "Bayesian Measures of Model Complexity and Fit (with Discussion)", *Journal of the Royal Statistical Society. Series B.* 64, 4: 583-640.

Sturgis, Patrick and Pamela Campanelli. 1998. "The Scope for Reducing Refusals in Household Surveys: An Investigation based on Transcripts of Tape-recorded Doorstep Interactions." *Journal of the Market Research Society.* 40, 2: 121-139.

Van der Vaart, Wander, Yfke Ongena, Adriaan Hoogendoorn and William Dijkstra. 2005. "Do Interviewers' Voice Characteristics Influence Cooperation Rates In Telephone Surveys?." *International Journal of Public Opinion Research.* 18, 4: 488-499.

White, Amanda, Stephanie Freeth and Jean Martin. 2006. "Evaluation of Survey Data Quality Using Matched Census Survey Records." *Survey Methods Bulletin, Office for National Statistics.* London. 49, 7, 01: 39-46.

Table 1: Summary of main survey characteristics for the six surveys. (adapted from Durrant and Steele, 2009).

Survey Design Characteristic	**EFS**	**FRS**	**GHS**	**OMN**	**NTS**	**LFS**
Length of data collection period	1 month +1 week	1 month	1 month	3 weeks	2.5 to 6.5 weeks	7+7+2 days (spread over 13 week period)
Interviewer workload in number of addresses	18	24	23	30	23	20
Type of additional interviewer training given (in addition to standard training)	1 day	1 day	briefing	postal	1.5 days	4 days (interviewers work only on this survey)
Purpose leaflet available	Yes: in the field	Yes: in the field	Yes: in the field	Yes	Yes: postal (London only)	Yes: postal
Respondent incentives	Stamps; £10/£5 for diary	Stamps	None	Stamps	Pen and fridge magnet	None
Respondent rules	All house-holders aged 16+	All house-holders aged 16+	All house-holders aged 18+	One house-holder aged 16+	All house-holders aged 16+	All house-holders aged 16+
Proxy response allowed	Yes	Yes	Yes	No	Yes	Yes
Average length of interview (in mins)	70	80	70	26	60	30 (for wave 1)
Diary required (in addition to questionnaire)	Yes: 2 weeks	No	No	No	Yes: 1 week	No
Refusal rate	30.4%	24.4%	19.4%	21.9%	22.6%	13.9%

The surveys collect information based on the household as a whole and on the individuals within the households.

Information collected by survey:

EFS: Core topics include: household expenditure, rent and mortgage payments, taxes, benefits, detailed information about the income of each household member, and trends in nutrition.

FRS: Aims to provide information on living standards, people's relationship and interaction with the social security system. The questionnaire seeks information on income and benefits, tenure and housing costs, assets and savings, occupation and employment, health and ability to work, pensions and insurance, childcare and caregivers.

GHS: Core topics include: accommodation, consumer durables, housing tenure, migration, employment, pensions, education, health, smoking, drinking, family formation, and income.

NTS: Aims to provide a comprehensive picture of personal travel behaviour. Questions include ethnic group, place of work, reliability and frequency of local services such as buses and trains, use of vehicles, long distance journeys and travel outside of Great Britain.

OMN: Multi-purpose survey which aims to obtain information about the general population or about particular groups. The questionnaire is in two parts, including first a set of core classificatory questions and then a series of unrelated modules on varying topics at the request of customers. Core questions include information on demographic details, economic status, job details, employment status, full- or part-time working, tenure, and ethnic origin.

LFS: Aims to provide information about the UK labour market and unemployment. The survey seeks information on respondent's personal circumstances, their labour market status and income.

Table 2: Estimates (with 95% credible intervals) of the interviewer and area random effect variances from alternative specifications of the multilevel cross-classified logistic models for refusal. †

	Interviewer variance	SE	Area variance	SE	DIC
Model 0 (variable survey; hh level only)	--	--	--	--	18863
Model 1a (Model 0 with interv. Var)	0.094 (0.064; 0.129)	(0.017)	--	--	18742
Model 1b (Model 0 with area effects)	--	--	0.055 (0.032; 0.083)	(0.013)	18796
Model 2 (Model 1a with area var; cross-classified)	0.077 (0.047; 0.112)	(0.017)	0.026 (0.007; 0.052)	(0.011)	18735
Model 3 (Model 2 + household variables)	0.081 (0.049; 0.118)	(0.018)	0.013 (0.001; 0.037)	(0.010)	18338
Model 4 (Model 3 +interviewer variables)	0.040 (0.012; 0.070)	(0.015)	0.012 (0.001; 0.033)	(0.009)	18321
Model 5 (Model 4+area variables)	0.039 (0.015;0.069)	(0.014)	0.010 (0.001; 0.032)	(0.008)	18319

† The values in each cell are the point estimate (the means of 80,000 MCMC samples, with burn-in of 5,000) and the corresponding 95% interval estimate (the 2.5% and 97.5% points of the distribution). Standard errors (SE) are calculated as the standard deviations of the estimates from the MCMC samples.

Table 3: Estimated coefficients (and standard errors in parentheses) of multilevel cross-classified logistic model predicting probabilities of refusal (Model 5).

Variable (0 = Reference category)	**Categories**	$\hat{\beta}$ $(ste(\hat{\beta}))$
Constant		0.06916 (0.180)
Household-Level Variables		
Survey indicator † (0 EFS)	1 FRS 2 GHS 3 OMN 4 NTS 5 LFS	-0.385 (0.074)** -0.636 (0.069)** -0.445 (0.069)** -0.470 (0.071)** -1.256 (0.083)**
Educational attainment (HRP) † (0 No educational attainment)	1 O/A levels, GCSEs (UK school degrees) 2 First/Higher/College degree 3 Other professional degree/ educational attainment	-0.229 (0.070)** -0.550 (0.089)** -0.283 (0.126)**
Household has no dependent children (0 household has dependent children)	1 no dependent children	0.242 (0.050)**
London indicator (0 not London)	1 London	0.185 (0.077)**
Rural indicator (0 Urban)	1 Rural	-0.188 (0.069)**
Economic Activity (HRP) (0 Employed)	1 Self employed 2 Unemployed 3 Retired 4 Looking after family 5 Other (incl. student, permanently sick etc)	0.279 (0.066)** 0.135 (0.124) -0.162 (0.061)** -0.097 (0.129) -0.014 (0.086)
Perception on health (HRP) (0 Good)	1 Fairly good 2 Not good	0.130 (0.045)** 0.126 (0.064)**
Caregiver in household (0 No)	1 Yes	-0.152 (0.051)**
Household type (0 Single household)	1 Couple household 2 Multiple household	0.069 (0.048) 0.234 (0.069)**
Car Ownership (0 One car or more)	1 No car	0.117 (0.051)**
Household moved during last year (0 No)	1 Yes	-0.155 (0.077)**
Interviewer Observations		
Gender of householder at first contact † (0 Male)	1 Female	-0.021 (0.050)
House in a better or worse condition than others in area (0 Better)	1 Worse 2 About the same	0.433 (0.090)** 0.101 (0.063)
Area-Level Variables		
Percentage of population between 0 and 4 years (centred)	(continuous variable)	0.076 (0.037)*

Interviewer-level variables (IAS)		
Interviewer gender † (0 Male)	1 Female	-0.037 (0.066)
Interviewer educational attainment † (0 Degree or postgraduate, College degree)	1 Academic attainment below College degree (O/A levels, GCSEs) 2 Lower, no, other educational attainment	-0.057 (0.072) -0.488 (0.206)**
Pay grade (0 Interviewer)	1 advanced interviewer and merit 1 and 2 2 merit 3 and field manager	-0.117 (0.070) -0.382 (0.094)**
Years of experience (0 Less than 1 year)	1 1 to 2 years 2 3 to 8 years 3 9 years or more	-0.021 (0.073) 0.060 (0.090) 0.267 (0.114)**
Daily hours previous year weekdays (1 5 and more hours)	1 0-4 hours	0.118 (0.062)*
Should persuade most reluctant respondent (0 strongly agree, agree)	1 neither agree nor disagree 2 disagree, strongly disagree	-0.155 (0.082)* 0.106 (0.065)*
Can persuade when others can't (0 disagree, strongly disagree)	1 neither agree nor disagree 2 strongly agree, agree	-0.105 (0.049)** -0.300 (0.096)**
Can convince reluctant respondents † (0 Less confident)	1 More confident	-0.648 (0.204)**
Refusal affects how behave (0 Rarely, never)	1 always, frequently, sometimes	-0.135 (0.054)**
No matter what I do, some respondents will never agree to participate (0 strongly agree, agree)	1 strongly disagree, disagree, neither agree nor disagree,	-0.212 (0.109)**
If respondent refused because too busy it is better to send a different interviewer † (0 strongly agree, agree)	1 strongly disagree, disagree, neither agree nor disagree	0.154 (0.078)**

Continued…

Interactions between interviewer and household characteristics		
Interviewer gender * Gender of householder at first contact (0 Male and 0 Male)	1*1 Female – Female	-0.125 (0.076)*
Educational attainment (HRP)* Interviewer educational attainment (0 No educational attainment and 0 Degree or postgraduate, College degree)	1*1 O/A levels, GCSEs – Academic attainment below College degree 2*1 First/Higher/College degree – Academic attainment below College degree 3*1 Other professional degree/ educational attainment – Academic attainment below College degree 1*2 O/A levels, GCSEs – Lower or no educational attainment 2*2 First/Higher/College degree – Lower or no educational attainment 3*2 Other professional degree/ educational attainment – Lower or no educational attainment	0.051 (0.093) 0.084 (0.121) 0.045 (0.169) 0.121 (0.266) 0.627 (0.317)** 1.213 (0.451)**
Can convince reluctant respondents * household has no dependent children (0 Less confident and 0 household has dependent children)	1 More confident - no dependent children	0.198 (0.112)*
Survey-Specific effects		
Survey indicator * Interviewer can convince reluctant respondents (0 EFS and 0 Less confident)	1*1 FRS –more confident 2*1 GHS-more confident 3*1 OMN-more confident 4*1 NTS-more confident 5*1 LFS-more confident	0.290 (0.243) 0.196 (0.228) 0.295 (0.224) 0.405 (0.233)* 0.549 (0.222)**

The estimated coefficients and their standard errors are the means and standard deviations of parameter values across 80,000 Markov chain Monte Carlo samples, after the burn-in of 5000 and starting values from second order PQL estimation. The missing value categories have been suppressed to save space.

* significant at the 10% level
** significant at the 5% level
† interaction between interviewer characteristic and either survey or household characteristics
HRP information based on household reference person

Table 4: Predicted probabilities of refusal (in %) based on selected two-way interactions.†

Interaction between survey and interviewer attitude							
		EFS	**FRS**	**GHS**	**OMN**	**NTS**	**LFS**
Can convince reluctant respondent	Less confident	42.2	25.8	20.5	22.0	20.9	12.1
	More confident	30.7	22.0	16.4	19.3	19.6	11.7

Interaction between gender of the interviewer and householder at first contact			
		Interviewer Gender	
		Male	Female
Gender of householder at first contact	Male	23.9	23.0
	Female	23.5	20.8

Interaction between education of interviewer and HRP				
		Education of Interviewer		
		Degree or postgraduate (University degree)	Academic attainment below University degree, O/A levels, GCSEs (school degrees in UK system)	Lower educational attainment, no attainment, other educational attainment
Education of HRP	No educational attainment	26.8	25.7	17.9
	Academic attainment below College degree, O/A levels, GCSEs (school degrees in UK system)	22.7	22.5	16.4
	First/Higher/ College degree	17.5	18.1	20.4
	Other professional degree, other educational attainment (e.g. City and Guilds)	21.6	21.5	36.2

Interaction between interviewer can convince reluctant respondents and household without dependent children				
		Interviewer can convince reluctant respondents		
		Less confident	More confident	
Dependent Children	Household without dependent children	19.1	13.1	
	Household with dependent children	15.6	8.9	

† Predicted probabilities are calculated by varying the values of the two interacting variables, holding all other covariates at their sample mean value. In the case of a categorical variable, the dummy variable associated with a particular category takes on the value of the sample proportion in that category instead of the usual 0 or 1 value.