To be published in the Journal of the Royal Statistical Society, Series A, 2011

Using field process data to predict best times of contact conditioning on household and interviewer influences

Gabriele B. Durrant[1], Julia D'Arrigo[1] and Fiona Steele[2]
[1]Southampton Statistical Sciences Research Institute (S3RI)
University of Southampton, UK
(g.durrant@southampton.ac.uk)

[2]Centre for Multilevel Modelling
University of Bristol, UK

Summary

Establishing contact is an important part of the response process and effective interviewer calling behaviours are critical in achieving contact and subsequent cooperation. This paper investigates best times of contact for different types of households and the influence of the interviewer on establishing contact. Recent developments in the survey data collection process have led to the collection of so-called field process or paradata, which greatly extend the basic information on interviewer calls. This paper develops a multilevel discrete time event history model based on interviewer call record data to predict the likelihood of contact at each call. The results have implications for survey practice and inform the design of effective interviewer calling times, including responsive survey designs.

Key Words: paradata, interviewer call-record data, responsive survey design, multilevel discrete-time event history models.

Address for correspondence: Gabriele Durrant, Southampton Statistical Sciences Research Institute, University of Southampton, SO17 1BJ, Southampton, g.durrant@southampton.ac.uk

1. Introduction

In recent years many surveys have seen a decline in response rates (De Heer, 1999). Survey agencies have to undertake great efforts to increase response rates and, at the same time, to reduce the costs of survey data collection. Establishing contact is an important part of the response process, which is often costly and time-consuming (Weeks et al., 1980; Groves and Couper, 1998; Cunningham et al., 2003). Effective interviewer calling behaviours are therefore critical in achieving contact and subsequent cooperation. Although survey agencies have become increasingly interested in understanding and improving the process of data collection, research so far has analysed primarily the final outcome of contact/non-contact rather than the process leading to contact (Weeks et al, 1980; O'Muircheartaigh and Campanelli, 1999; Durrant and Steele, 2009).

The increasing interest in the data collection process has led more recently to the development of so-called field process data or *paradata* (Couper, 1998). The term is used to describe empirical measurements about the process generating the survey data, such as time and day of the call and, for face-to-face surveys, interviewer observations about the physical and social characteristics of the selected housing unit and the neighbourhood. An increasingly important area for the use of paradata in survey organisations is responsive survey design (Groves and Heeringa, 2006; Laflamme et al., 2008; Laflamme, 2009), where the continuous measurement and monitoring of the process and survey data offers the opportunity to alter the design during the course of the data collection to reduce costs and to increase the quality of the survey data. So far, however, few studies have used paradata for updates on the progress of data collection, as decision-making tools during data collection or for adjustment at the data analysis stage.

To date, analyses of paradata and interviewer calling behaviour, in particular for face-to-face surveys, have been limited. Much of this research has focused on the average best times of day and days of the week to establish contact, without controlling for household

characteristics and prior call information (e.g. Weeks et al., 1980). Greenberg and Stokes (1990) and Kulka and Weeks (1988) conditioned on previous call times but did not have household-level information available. Some studies controlled for basic information about the household or area, but without deriving household-specific estimates of the probability of contact (Purdon et al., 1999; Groves and Couper, 1998; Brick et al. 1996; O'Muircheartaigh and Campanelli, 1999). Most research on optimal calling scheduling has been carried out in the context of telephone surveys (e.g. Weeks et al., 1987; Greenberg and Stokes, 1990; Brick et al. 1996) rather than face-to-face surveys, although the latter offer a much wider range of observational information available for each household and call (Groves and Couper, 1998; Greenberg and Stokes, 1990). Techniques to analyse such data have often been limited to descriptive statistics and simple logistic regression modeling, and usually only one survey was considered (e.g. Weeks et al., 1987; Purdon et al., 1999; Groves and Couper, 1996; Wood et al., 2006; Elliott et al., 2000). Although often acknowledged as important for securing cooperation, few studies have considered the role of the interviewer on the contact process (for examples see Purdon et al., 1999; Groves and Couper, 1998; Blom and Blohm, 2007), and those that have, used basic analysis techniques or had only limited information about interviewers.

A major advantage of this study is that we have access to rich paradata including information recorded by the interviewer at each call to the household (even if contact was not made), interviewer observations about the household and neighbourhood, and detailed information about the interviewers themselves. The dataset combines call-record data from six major UK face-to-face surveys, which allows more general inferences to be made than in previous work. A key strength of these data is that individual and household characteristics from the UK 2001 Census are linked to the paradata for both contacted and non-contacted households. The resulting data have a multilevel structure with households nested within a cross-classification of interviewers and areas. As identified by Groves and Heeringa (2006, p.

455), research is needed to establish how best to use such paradata to inform nonresponse processes, as well as further methodological development in the specification of models based on such data.

This paper aims to build and improve response propensity models based on paradata to predict the likelihood of contact at each call, conditioning on household and interviewer characteristics. We use multilevel discrete-time event history analysis (Steele et al., 2004) to model the propensity of contact, allowing for household, interviewer and area effects in a cross-classified model. The model conditions on information available for each household, such as from administrative data and prior calls, and includes call record data as time-varying covariates. The key research questions are:

1. What are the best times of the day and days of the week to establish contact?
2. What are the best times to establish contact with certain types of households, in particular households that are generally more difficult to contact?
3. To what extent does establishing contact and the success of the timing of the call depend on interviewer characteristics?

The paper aims to provide guidance to academic researchers and survey practitioners on how to model and use interviewer call record data for the design of effective and efficient interviewer calling strategies. It is anticipated that this research will inform the improvements of responsive survey designs and the design of call-backs and follow-ups of nonrespondents, with implications for survey agencies for the allocation of time and staff resources. Although survey organisations may not have access to information such as the census variables considered in this study, the analysis provides useful information about the type of data that could be beneficial for predicting contact and survey organisations could explore proxies for such variables from available data sources. It would also be possible to train interviewers to collect relevant observation data on earlier calls. If some attributes of the households are

observable, survey designs might be altered to improve efficiency, reduce costs and increase contact rates (Groves and Couper, 1998). The paper is organised as follows. Section 2 describes the data available. The methodology for the analysis is presented in Section 3. Section 4 outlines the rationale for the modelling, the choice of variables and the modelling strategy. Section 5 discusses the results. A summary of the findings with implications for survey practice is provided in Section 6.

2. Data

2.1 Field process data (paradata)

This study takes advantage of comparatively rich field process data (paradata) captured during the data collection period of six face-to-face UK government surveys in 2001. In each survey interviewers recorded information on each call to a household via an interviewer observation questionnaire. The key advantage of these data is that they have been linked to individual and household information from the UK 2001 Census, interviewer information from a survey of all face-to-face interviewers working for the UK Office for National Statistics (ONS) in 2001, and area information from registers and aggregated census information. The timing of the study was chosen to coincide with the last UK Census in 2001.

The available paradata include records of calls and interviewer observations about the household and neighbourhood captured by the interviewer during data collection. The call record data include the time and day of call, brief information on the contact strategy used at the call, and the outcome of the call. The interviewer also recorded (usually at the first visit) their observations about the household and neighbourhood, such as if there were any physical barriers to the house, type of accommodation, quality of housing and information about the household composition, such as any signs of the presence of children. The interviewer observation data are, in principle, available even if no contact was made with the household.

The interviewer is said to have made *contact* with a household at a given call - the dependent variable in our analysis - if he/she was able to talk to at least one responsible resident at the sampled household, either face-to-face or through an entry phone. The guidelines provided to interviewers by the survey organisation state that the final response outcome for an address cannot be coded as 'non-contact' until at least four calls have been made. At least two of these calls should be in the evening or on a Saturday. In our dataset the maximum number of calls made to a household is 15. The study includes households selected for interview in one of the six surveys during May-June 2001, the months immediately following the 2001 Census. The call record data are available for 16,799 households (after excluding vacant and non-residential addresses, re-issues and unusable records, as described in Durrant and Steele, 2009), of which 1,017 households were never contacted. This results in an overall final non-contact rate of about 6%. Although the non-contact rate may not appear very large in comparison to the refusal rate (for the surveys considered here around 15-30%), establishing contact is a costly and time-consuming process. Our dataset contains a total of 69,619 calls to households of which more than half (37,879 calls) were made to establish first contact or until the household was coded as a non-contact.

The six face-to-face household surveys for which the interviewer call record data were collected 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 non-contact rates for the six surveys range from 3% to about 10% which may be explained by differences in the survey design, length of data collection period, minimum number of calls to be made, interviewer workload, interviewer qualifications and interviewer training. Further details about these surveys can be found in Durrant and Steele (2009) and at www.statistics.gov.uk.

It should be noted that the ideal dataset for such an analysis would be based on fully randomized calling times for all sample units. Such a design would, however, be practically

impossible, at least for face-to-face surveys; it could be achieved to some extent for telephone surveys (Groves and Couper, 1998). The dataset here, similar to previous work, provides information on observed calling times, i.e. the times that the interviewer chose to call on a household. If an interviewer's decision to call at a particular time can be regarded as independent of the characteristics of the sample unit, a departure from fully randomised calls should not be important. It seems reasonable to assume that interviewers choose when to make their first call with little, if any, prior knowledge about the sampling units. However, the timing of subsequent calls may depend on additional knowledge that the interviewer obtained at an earlier call. We therefore control in our models for characteristics of the households that are related to differential interviewer calling strategies, in particular household and area characteristics from both the census and the interviewer observation data. This issue has been discussed further in Purdon et al. (1999, p. 201), Groves and Couper (1998, p. 82) and Kulka and Weeks (1988).

2.2 Linked data

The field process data were linked to demographic and socio-economic individual and household level information from the UK 2001 Census, available for both contacted and non-contacted sample households. It should be noted that some of the information from the interviewer observation data coincides with information recorded via the census (e.g. type of accommodation) and wherever possible we used the interviewer observation variables. Detailed information about the interviewer was linked to the household level information. These data were obtained via a separate face-to-face survey (Interviewer Attitude Survey) of ONS interviewers during June 2001, at around the time of the survey and census data collection period. The information on interviewers includes socio-demographic characteristics, and employment background, such as pay grade and experience, workload and planning,

attitudes, strategies and behaviours for dealing with non-contacts as well as information about doorstep approaches.

Area information is available from aggregated census data, where area is defined as the local authority district. The dataset contains a total of 565 interviewers and 392 areas. It should be noted that in clustered survey designs an interviewer is normally assigned to one primary sampling unit (PSU) and their workload consists of all sampled households in that PSU. Interpenetrated sampling designs may be used to avoid confounding of area and interviewer effects, where interviewers are allocated at random to households. Such designs enable, at least to some extent, a separation of interviewer and PSU effects. However, due to the high costs of implementing interpenetrated designs, only very few studies of this kind exist (O'Muircheartaigh and Campanelli, 1999; Schnell and Kreuter, 2001). Usually, if no such design has been employed, area effects are ignored in the analysis or area information is not available (e.g. Pickery and Loosveldt, 2004). Although the surveys included in this study did not employ randomised interpenetrated sampling designs, a complete confounding of area and interviewer effects was avoided because most interviewers work on a number of surveys and some interviewers work across PSUs. In particular, we allow for area effects in our models where areas are defined as local authority districts, with the PSUs not strictly nested within the local authority districts but crossing boundaries. 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. As described in Section 3 we use a multilevel cross-classified model to analyse this type of data. For other examples of the use of multilevel cross-classified models and a detailed discussion of different forms of (partial) interpenetrated sampling designs see Durrant et al. (2009) and von Sanden (2004), respectively.

Deterministic matching (Herzog et al, 2007; Ch. 8.3) was used to link the various datasources based on key identifying variables, including UK address id number (the Ordnance Survey Address Point Reference which uniquely defines and locates a postal address

based on postcode, house and flat number etc.), gender, age or date of birth and if necessary further identifying information, as well as information routinely collected as part of the survey administration, such as interviewer id. The linkage was carried out separately for every survey. For about 95% of all households a match to the census records was found. For the analysis sample used here, any potential error due to incorrect matching is assumed to be small for the following reasons: a.) due to the uniquely defined postal system in the UK exact matching at the address level is likely to be achieved; also, the address id number is used across all surveys and censuses in the UK; b.) the analysis sample used here only requires linkage at the household level but not on the individual level which would be more difficult; c.) for the case of a multi-occupied address further identifying information was used. If an exact match was not found a match was selected at random, which was, however, carried out in less than 1% of matched cases; d.) cases causing higher linkage errors such as households that moved during the short period between the census and the survey, non-residential, vacant and second homes were excluded from the analysis (see above). A number of quality checks and a significant amount of clerical review were carried out to identify and minimise any potential linkage errors. No potential effects due to the loss of unmatched cases were found, for example comparing the distribution of key variables before and after the linkage (Beerten and Freeth, 2004).

The effects of linkage errors on data analysis are analogous to those of measurement error and may therefore lead to an attenuation of regression coefficients and increased variability (Fuller, 1987; Lahiri and Larsen, 2005). However, effects on the data analysis here, due to the potential small record linkage errors, are assumed to be small. Effects of linkage and measurement errors on regression analysis are further discussed in Scheuren and Winkler (1993) and specifically for multilevel models in Goldstein (2010; Ch. 14).

3. Multilevel discrete time hazard model for the probability of contact

Multilevel event history analysis (see e.g. Steele, Diamond and Amin, 1996) was used to model the probability of contact at a particular call, given that no contact was made prior to that call (i.e. the number of calls to first contact). Households that were not contacted by the end of the data collection period have right-censored contact histories. Denote by $y_{i(jk)t}$ the binary indicator of contact, coded 1 if contact is made with household i of interviewer j in area k at call t and 0 if the contact attempt fails. The grouping of the j and k indices in parentheses, (jk), indicates a cross-classification of interviewers and areas. The conditional probability of contact at call t given no contact before t – commonly referred to as the discrete-time hazard function – is defined as $\pi_{i(jk)t} = \Pr(y_{i(jk)t} = 1 \mid y_{i(jk)t-1} = 0)$. The multilevel cross-classified discrete-time hazard model, allowing for a clustering of households within a cross-classification of interviewers and areas, may be written

$$\log\left(\frac{\pi_{i(jk)t}}{1-\pi_{i(jk)t}}\right) = \alpha_t + \boldsymbol{\beta}' \, \boldsymbol{x}_{i(jk)t} + u_j + v_k \,. \qquad (1)$$

$\boldsymbol{x}_{i(jk)t}$ is a vector of covariates, with coefficients $\boldsymbol{\beta}$, including time-varying attributes of calls (e.g. time and day of contact attempt), time-invariant characteristics of households, interviewers and areas, and two-way interactions between call and household-level variables. α_t is a function of the call number t ('time') which allows the probability of contact to vary across calls; here α_t was initially fitted as a step function, i.e. $\alpha_t = \alpha_1 D_1 + \alpha_2 D_2 + ... + \alpha_T D_T$ where $D_1, D_2, ..., D_T$ are dummy variables for calls $t = 1,...,T$ with T the maximum number of calls, but simpler monotonic functions were also explored. Unobserved interviewer and area characteristics are represented respectively by independent random effects u_j and v_k, assumed to follow normal distributions: $u_j \sim N(0, \sigma_u^2)$ and $v_k \sim N(0, \sigma_v^2)$.

After restructuring the data so that, for each household, there is a record for every contact attempt, the multilevel discrete-time event history model (1) can be estimated as a cross-classified model for the binary responses $y_{i(jk)t}$. Estimation is carried out using Markov chain Monte Carlo (MCMC) methods as implemented in the MLwiN software (Browne, 2009; Rasbash et al., 2009a). To aid interpretation of the fitted model, predicted probabilities of contact are calculated for each value of the categorical covariates, holding constant the values of all other covariates in the model. To obtain mean probabilities, we average across interviewer and area-specific unobservables by taking random draws from the interviewer and area random effect distributions. The simulation approach involves generating a large number of pairs of random effect values from independent normal distributions with variances $\hat{\sigma}_u^2$ and $\hat{\sigma}_v^2$, calculating a predicted probability based on each pair of generated values and the estimated coefficients, and taking the mean across the simulated values. This procedure is implemented in MLwiN and described in Rasbash et al. (2009b).

4. Choice of explanatory variables and modelling strategy

The conceptual framework of Groves and Couper (1998) for household survey nonresponse identifies a number of important influences on the process of contact, including the timing and frequency of the calls, social environmental and socio-demographic characteristics, at home patterns of the householders and the presence of any physical impediments to gaining access to the household. Such attributes may be separated into factors that are under the control of the interviewer or survey organisation and factors outside their control (Purdon et al., 1999). Our analysis aims to control for all of these effects. Examples of variables under the direct control of the interviewer or survey organisation are the time of day and day of the week of the call and the time between calls. Most previous research has analysed the overall best times to contact and found that evening and weekend calls are optimal (Weeks et al., 1980;

Swires-Hennessy and Drake, 1992). Survey agencies, however, need to decide how best to allocate limited staff and time resources and not all calls can be made in the evenings and at weekends. A logical question to ask is which households have the highest chance of contact during the day, so that evening and weekend times can be reserved for more difficult cases. The survey organisation may then refine the calling strategy in light of information available about a household, for example as part of a responsive survey design. We therefore explore interactions between call times and household characteristics to determine best times of contact for particular households.

Of particular interest is the influence of interviewer observation variables as survey agencies can collect this information for all households, including non-contacts. Such data are especially useful when no information from administrative data or census is available. Interviewer data include (time-invariant) information about physical barriers to accessing the household (e.g. a locked common entrance, locked gate or entry phone), the presence of security devices (e.g. security staff, CCTV cameras or burglar alarm), indications about boarded-up or uninhabitable buildings in the area, household composition, quality of the housing and how safe the interviewer would feel walking in the area after dark.

Contact strategies and interviewer behaviours, such as attempting to establish contact by telephone or leaving a card or message at a call, are further examples of variables under the control of the interviewer or survey organisation. Such call-specific variables are included in the models as time-varying covariates. Some further time-varying variables, such as the time between calls, were derived from the call record history. An interesting question for survey agencies is whether changing the timing of the call increases the likelihood of contact and we therefore investigate the influence of the call history (see Purdon et al., 1999; Groves and Couper, 1998 and Kulka and Weeks, 1988). A separate indicator for the first call was included in the model and variables relating to earlier calls, such as the time of the previous call, were

coded zero for the first call. This coding allows the coefficients of these call history variables to be interpreted as effects for second and subsequent calls.

Factors that are outside the direct control of the interviewer include characteristics of the household or area that indicate at home patterns and lifestyle of household members, attributes of the social environment, socio-demographic characteristics and indicators of physical impediments to accessing the household. We investigate the influence of variables that may be regarded as proxies for the time spent at home and lifestyle, such as indicators of a single-person household, presence of dependent children, pensioners, carers or a person with a limiting long-term illness and adults in employment, as well as social, environmental and socio-demographic attributes; at the area level we considered a range of socio-demographic variables, mostly aggregated information from the 2001 Census.

Although not always possible, there are various ways for survey organisations to obtain information about a household, or at least about the area, prior to (or sometimes after) the start of fieldwork. This information could come from the sampling frame, census, registers or administrative data -possibly only at an aggregated level- or in case of a longitudinal survey from a previous wave. The availability of such data may depend on the country: Scandinavian countries and the Netherlands have access to rich administrative data (for an example see Cobben and Schouten, 2007). Any such prior information may be used to direct interviewer calling efforts at the start of data collection. Furthermore, interviewers may already have some prior knowledge about the areas and the type of households they have to contact. After the first call the interviewer should be able to gather more information about each household, e.g. based on visual observations or by talking to neighbours. Subsequent calling strategies may then depend on this information.

Previous research on the influence of the interviewer on the nonresponse process has focused on the cooperation/refusal stage (Durrant et al., 2009; O'Muircheartaigh and Campanelli, 1999; Groves and Couper, 1998). Few studies have considered the role of

interviewers in establishing contact (Purdon et al., 1999; Groves and Couper, 1998; O'Muircheartaigh and Campanelli, 1999). Purdon et al. (1999) and Groves and Couper (1998) argue that the impact of interviewer characteristics should operate through the time, day and frequency of calling, as these are the only parts of the contact process that interviewers can control. After adjusting for the timing of the call the interviewer should not play a significant role. Groves and Couper (1998) nevertheless investigate if there are any further net effects of interviewer characteristics and explore simple relationships between interviewer attributes and the probability of contact. Purdon et al. (1999) find a significant influence of interviewer pay grade (although only based on a single-level model which does not allow for clustering due to unmeasured interviewer characteristics). In this paper, we hypothesise that characteristics such as the qualification, pay grade and experience of the interviewer may play a role in establishing contact. Such variables may be indicators of an interviewer's ability to judge best times of contact for different households. Another mechanism through which attributes of the interviewer may impact could be through knowledge of the area and types of households. Moreoever some interviewers may be better at organising their workload and prioritising their cases, leading to higher contact rates. In addition we explore the extent to which interviewer strategies may influence the probability of contact. The survey organisation may also have limited influence over certain interviewer characteristics, in the sense that interviewers may be assigned to households based on available information about interviewers, areas and households. For example, more experienced interviewers may be allocated to more difficult cases or areas. Interactions between interviewer and household characteristics were investigated to see which interviewers may be better at establishing contact with generally harder to reach households. To analyse differences in effectiveness of interviewers at certain times of the day, interactions between interviewer and call characteristics were explored.

Due to the large number of variables available, testing of main effects and interactions was primarily guided by theories of contact and interpretation (Durrant and Steele, 2009;

Groves and Couper, 1998). Variables that were not significant at the 10% level, and did not interact significantly with other variables, were excluded from the final model. Some variables in the dataset are subject to a small amount of item nonresponse. To maximise the size of the analysis sample we allowed for a missing category for those variables subject to item nonresponse. The coefficients of dummy variables for these categories were not significant in the final model and are not shown in the tables of results. We investigated a series of models starting with a simple specification including only dummy variables for survey, the previous call indicator and the number of previous calls. We then added interviewer and area random effects in a cross-classified multilevel model. Next, we entered time-invariant household and time-varying call-level variables and two-way interactions between household and call characteristics. Finally, we include interviewer-level variables to examine the extent to which these may explain between-interviewer variance in the contact rate.

5. Results

5.1 The hazard rate and average best times of contact

Analysing the hazard of contact at each call, based on a simplified version of model (1) with only dummy variables for call number (results not shown), we found, in line with previous studies (e.g. Purdon et al., 1999; Groves and Couper, 1998), that the probability of contact declines almost linearly with the number of calls, from about 50% for the first call to just under 30% for the 15th call. We therefore simplified the specification of the baseline logit hazard, α_t in (1), by including the number of previous calls as a linear term.

Table 1 shows the probability of contact at the first call by time of day and day of the week. By far the most popular times to call are weekday afternoons, followed by weekday evenings and weekday mornings, with a clear decline in the frequency of calls from the beginning to the end of the week for all times of the day. Few calls are made at the weekend, in particular on a Sunday due to interviewer working practices. Calling on weekday evenings

yields the highest probability of contact, with a particularly high probability towards the beginning of the week and decreasing thereafter. Weekend calls also lead to a higher probability of response, with Sunday evenings showing a similar pattern to early weekday (Mon-Wed) evenings. The next most productive times to call are weekday afternoons. Weekday mornings are generally the worst times to establish contact. During the week, afternoons are better than mornings but it is the other way round at the weekend.

[Table 1 about here]

These indicative findings largely support the conclusions of previous research, that evenings and weekends are optimal times to call (Weeks et al., 1980; Swires-Hennessy and Drake, 1992; Purdon et al. 1999; Groves and Couper, 1998). These results and some initial modelling informed the categorisation of the calling time variable used in the final model (Table 2) distinguishing early week (Mon-Wed), late week (Thu-Fri) and weekend (Sat-Sun) and morning, afternoon and evening. Since the contact probability is similar for Sunday evening and early weekday evenings (and there are few Sunday evening calls) we merged the Sunday evening and early week (Mon-Wed) evening categories in the calling time variable.

5.2 Best times of contact for different types of households

The chance of making contact at a given time of day will depend on the characteristics of the household that indicate the householder's at-home patterns. We now investigate the best times to establish contact with certain types of households, in particular those households that are generally more difficult to contact. Table 2 presents parameter estimates of two multilevel discrete-time hazard models which take account of household, area and interviewer characteristics and interactions between time-varying variables and household and interviewer characteristics. Model A excludes census variables since these would not normally be available to a survey agency. Model B represents the final model, including census information. The inclusion of census variables reduces the DIC (Deviance Information Criterion, Spiegelhalter

et al., 2002) by only a small amount (i.e. by 163 from 46936 for Model A to 46773 for Model B), indicating that a model based only on interviewer observation variables does not have much less predictive power than the full model. Furthermore, there are no differences in the direction of effects between the two models, implying that similar results can be obtained also in the absence of additional administrative data, i.e. when the survey agency can only rely on recordings by the interviewers to obtain information about nonresponding households.

From Table 2 we see that the probability of contact is highest for the first call. The highly significant negative coefficient for number of previous calls indicates a decrease in the odds of contact by 10% for each additional call net of all other factors in the model ([1-exp(-0.110)]*100=10%). We tested for non-proportional effects of covariates by interacting each with number of previous calls, but there was no evidence to suggest that the effect of any variable differed across calls. In the following we distinguish between interviewer observation and census variables, although in practice, at least some of the census variables could be substituted by variables based on interviewer observations. It is well known that single-person households, households without children or with primarily young people, and households in urban areas and in flats are the most difficult to contact (Durrant and Steele, 2009; Groves and Couper, 1998), and our results confirm these findings. To aid interpretation of the interaction terms, predicted probabilities are provided in Table 3. (These have been calculated for call 1 but the pattern in probabilities is exactly the same for subsequent calls because the lack of interactions with the number of previous calls implies that all effects are constant across calls.)

[Tables 2 and 3 about here]

Household and neighbourhood characteristics based on interviewer observations

We first considered the effects of a range of interviewer observations. All of these variables were predictive of contact in initial modelling (i.e. before controlling for a range of household and interviewer effects), which suggests such variables are useful for guiding the

process of establishing contact in the field. As may be expected, houses with no security device visible - such as a security gate, burglar alarm, CCTV cameras or security staff - were easier to contact (Table 2). An observation that can be relatively easily recorded by the interviewer is whether the household lives in a house or a flat. For almost all call times, it is easier to establish contact with householders living in a house rather than a flat, and this is true even after controlling for household characteristics such as location, number of people in the household and presence of children. We also explored interactions between interviewer observation variables and time of call. The interaction with type of accommodation (Table 3) reveals that on afternoons, for any day of the week, it is easier to make contact with residents of houses than of flats. Householders living in flats are most likely to be contacted in the evenings and on weekend mornings. Contact was found on average to be more difficult when the interviewer recorded that houses in the area were in a fair or bad state of repair and that the house was in a worse condition than others in the area. The interaction term between timing of the call and state of repair of houses in the area provides some indication that the contact rate is better for houses in a fair or bad state of repair for Thur-Sun mornings. From Table 3, we can also see that for almost all call times the probability of contact is higher for households with children, with particularly high probabilities on weekday evenings, all afternoons and Mon-Wed mornings. The fact that weekday afternoons are good times may be related to children being back home from school. For households without children, calls made on weekdays during the day are the least likely to result in contact, whereas weekday evenings are the most promising. (Information about the presence of children is available from both the interviewer observation questionnaire and the census information. We decided to use the census variable in the final model due to the lower level of item nonresponse and potentially higher data quality of this variable.)

Two other call-specific variables that are under the control of the survey organisation, and that may determine best times of contact, are the timing of the previous call and the

length of time since the last call. Considering the main effect of time of previous call only (without the interaction term in the model) we found that if the previous call was already a weekday evening call then establishing contact at the next call becomes increasingly less likely, indicating a potentially difficult to contact household. We found some indications for a significant interaction between time of current call and time of previous call (Tables 2 and 3). If the previous call was a weekend call, it seems advisable to call early during the week either in the morning or evening, or on a weekend morning. If the previous call was on a weekday afternoon, promising times to call are evening and weekend and Mon-Wed mornings. If the previous call was made during the evening, calling again during the evening is most likely to lead to contact. Overall evenings and weekends are reliably good times to call. These findings suggest that interviewers may have some (although limited) opportunity for increasing contact rates by changing the time of the call, especially if it is to an evening or weekend. Similar conclusions were drawn by Weeks and Kulka (1988), although they present only descriptive statistics for the timing of the first three calls. Purdon et al. (1999) did not find a significant interaction between time of current and time of previous call, and Groves and Couper (1998) did not find interpretable conditional effects of the timing of previous calls.

The effect of the number of days between calls suggests that leaving a few days between calls, ideally about one or two weeks, increases the probability of contact compared to returning on the same day. The increased probability of contact for call-backs after one or two weeks may reflect effects of additional knowledge about the household gathered by the interviewer at the earlier call which led them to adopt such a calling schedule. For example, interviewers may have found out from neighbours that the household was on holiday.

Household characteristics from the census

The contact rate for weekday mornings (Mon-Wed) or afternoons (Mon-Fri) is higher for households without any adults in employment than for households with at least one employed

resident (Table 3), as would be expected. The reverse effect is found for evenings. For households with adults in employment the probability of contact for both weekday and weekend evenings is higher than for households in unemployment. There is a lower chance of contact for households with adults in employment on weekend mornings than for households in unemployment but weekend afternoons perform very similarly. (An indicator of whether any adults are in employment is also available from the interviewer observation questionnaire. Again due to the higher data quality of census data we included the census measure in the final model. For an example where information on employment status and unemployment benefits is available from administrative sources see Cobben and Schouten, 2007.)

The interviewer has a good chance of finding someone at home during the week if there is at least one pensioner present. We see particularly high probabilities of contact during the day in the early part of the week for pensioner households. Weekday evenings are also good times to establish contact with pensioners. Compared to other types of households, the contact rate for households with pensioners is relatively low at the weekend, particularly mornings. This may be partially explained by older people being more likely to have religious commitments on a Sunday for example. For households without a pensioner weekday evenings and weekend mornings are the best times to call. Households with at least one person with a limiting long term illness (LLTI) have high probabilities of contact throughout the week, as would be expected since such persons may be more likely to be at home due to their restricted daily activities and some may have a carer present. The probabilty of contacting these households is particularly high during the week (Mon-Wed), which is almost as good a time to call as evenings and weekends. Information on the presence of persons with a LLTI may be available in register or administrative databases (for an example see Cobben and Schouten, 2007). Alternatively, some crude indicators may be captured by the interviewer, for example via observations regarding wheelchair access to the house or a disabled parking permit visible in the car. From Table 2 we see that the number of people in the household has

a significant effect on the probability of contact, with larger households being easier to contact than single-person households. This may be expected since it will be more likely to find at least one person at home for larger households.

Geographical location and type of area are usually regarded as important predictors of non-contact (Groves and Couper, 1998). However, after controlling for household characteristics and random area effects the London and urban-rural indicators were no longer significant. Area-level variables (e.g. unemployment rate, percentage of older people and children etc.) were all significant before controlling for household and call-level information, but not in the final model. This implies that area variables may be regarded as weak proxies for household characteristics, in line with the findings of O'Muircheartaigh and Campanelli (1999). In the absence of other information, knowledge about the area would therefore be advantageous and predictive of contact.

The above findings are based on a pooled analysis of six UK surveys which are expected to differ in their contact rates, for example because of differences in their design, such as length of data collection period. We find that the LFS has a significantly higher probability of contact than the other surveys considered. This may be due to a number of factors, such as LFS interviewers working only on that survey. They also have a comparatively lower workload, and receive more intensive interviewer training, although it should be noted that the LFS also has a shorter data collection period than the other surveys.

5.3 Influences of the interviewer on the process of contact

There is significant variation between interviewers in their contact rates in all models. The inclusion of the interviewer characteristics reduced the between-interviewer variance from 0.11 to 0.08, explaining about 27% of the interviewer variance. Interpreting $\exp(\hat{\sigma}_u)$ $(= \exp(\sqrt{0.08}) = 1.33)$ as the effect of a one standard deviation increase in the unobserved characteristics represented by the random interviewer effect we find that, after adjusting for

covariates in the model, an interviewer whose unobserved characteristics place them at one standard deviation above the average has a 33% higher odds of making contact than an 'average' interviewer. The between-area variance was found to be substantially smaller than the between-interviewer variance, and controlling for household-level and call-level variables halved the between-area variance; in the final model area effects are only marginally significant at the 10% level (Table 2).

The effects of a number of interviewer characteristics were investigated to explain the between-interviewer variance in contact rates, including socio-demographic characteristics, experience and work background and interviewer strategies. It may be argued that more experienced and higher qualified interviewers may be better at establishing contact (see Groves and Couper, 1998, p. 95). We found pay grade of interviewers to be an important factor in explaining part of the differences between interviewers, with interviewers in higher pay grades being better at establishing contact. A similar effect was found in Purdon et al. (1999), which was counter to their *a priori* hypothesis of no interviewer effects after controlling for the timing of the call. We also found that interviewers with a higher qualification such as a university degree or postgraduate education have higher contact rates. This may indicate that certain types of interviewers may be better at judging best times to call, for example through gathering information about the household from observation and talking to neighbours, and using such information to tailor their calling strategy to maximise the chance of contact.

We also find that older interviewers (50 years and over) are more successful at establishing contact which may reflect their greater experience or the fact that they may appear more trustworthy. Another possible explanation is that older interviewers may have fewer time-constraining commitments outside their job, such as looking after young children, allowing greater flexibility on calling times. We also explored the interaction between age of the interviewer and timing of the call (Table 3), and found some evidence that older

interviewers may be better in judging the best times to call: older interviewers are more likely than younger interviewers to achieve contact on weekday evenings and on weekend mornings.

Somewhat surprisingly, we did not find any significant effects of the number of years of interviewer experience after controlling for the timing of the call as well as household and area characteristics. This is in line with Groves and Couper (1998) who also did not find an effect of interviewer experience. The expected positive association between experience and the probability of contact might be more adequately captured by pay grade and qualification and, to some extent, age which were all found to be significant. It may be argued that the pay grade of the interviewer captures a combination of length of experience and interviewer performance, with better performing interviewers expected to be on higher pay grades. This combination of characteristics may therefore be more important in explaining differences between interviewers rather than simply the length of time an interviewer has been in the job.

Since survey agencies are particularly interested in behavioural differences between interviewers, we also explored the extent to which interviewer strategies influence the probability of contact. We found that interviewers who always or frequently use the phone to establish contact, rather than visiting the household in person, perform worse than interviewers who rarely or never use the phone. This may be an indicator of interviewer effort, with interviewers putting in more effort and dedicating more time to each sample unit being more successful. Somewhat surprisingly some interviewer strategies, such as how often they check with neighbours, were not found to explain differences amongst interviewers, although it should be noted that these measures of interviewer practice are self-reported rather than from direct observation. As suggested by Groves and Couper (1998) it may be preferable to ask interviewers to record their strategy for each call or household. We find some support for their recommendation: the variable indicating whether it is the interviewer's general practice to leave a card or message behind had no significant effect on contact, while the time-varying

covariate capturing the same information for each call was found to be significant, showing an increase in the probability of contact at the next call if a card or message was left (Table 2).

It may be argued that more experienced interviewers and interviewers on higher pay grades are better at establishing contact with harder-to-reach households. Effects of this type could help to inform the allocation of certain interviewers to potentially more difficult households. However, we did not find any significant interaction effects of this type.

6. Summary and Discussion

This paper uses multilevel discrete-time event history analysis to model the process of establishing contact with sample members in face-to-face surveys. Our unique data allow exploration of the best times to contact different types of households, controlling for interviewer effects. Our findings can be summarised as follows:

1. The results support earlier findings that weekday evenings and weekend daytimes are, on average, the best times to call. Furthermore, we find that the best times to call depend on household characteristics, especially markers for at home patterns. Differences in optimal calling times have been found e.g. by type of accommodation and the presence of children, pensioners or unemployed persons.
2. There is substantial evidence that interviewer observations about a household and neighbourhood are useful for predicting best times of contact. Interviewer observation variables were predictive of contact before and after controlling for additional information about a household (from the census in the present study).
3. We find that area-level variables are predictive of contact before controlling for other household and calling variables, but they were not significant in the final model. Therefore, in the absence of additional information, area characteristics are useful for predicting contact.

4. We have found significant effects of interviewer characteristics on contact. Important in explaining interviewer differences are pay grade, qualifications and age. Interviewer experience was not found to be important after controlling for these factors. There is evidence that some interviewers may be more effective in establishing contact at certain times, which may indicate better judgement of when best to call. There is little empirical support for the hypothesis that some interviewers are more successful in establishing contact with more difficult households, such as single households.
5. It is of interest to know whether certain interviewer strategies are helpful in establishing contact. Our model showed some significant effects of such strategies, for example the probability of contact was higher at the next call if the interviewer left a card or message. Our results also suggest that interviewer strategies measured at the call or household level have greater predictive power than measurements at the interviewer level. We also found some indication that changing the time of the call may lead to higher contact rates, in particular when changing to evening and weekend calls.

The results have wide ranging implications for survey practice. They may inform the design of efficient and effective calling behaviours and follow-ups as well as responsive survey designs to increase response rates and to potentially reduce nonresponse bias. The type of model presented may be used to predict the likelihood of contact at the next call, conditioning on information known to the survey organisation or interviewer at each point in time - even in the absence of information like here from the census. Furthermore, probabilities of contact for different types of households can be derived conditioning on household characteristics that may be known to the survey organisation prior to or during data collection. Due to limited time and staff resources, not all calls can be made in the evenings and at weekends and survey organisations need to make informed decisions which households to call upon during the day. By identifying the types of household that have a high chance of being contacted during the

day, survey agencies can allocate staff and time resources more efficiently. The focus was on face-to-face surveys but some findings may also apply to telephone surveys.

The study highlights the benefits of prior information about sample units for improving prediction of contact, and survey agencies should exploit possibilities of data linkage to boost information available about each household or area. Such additional information may come from the sampling frame, registers or administrative data, as well as previous waves in the case of a longitudinal study - available prior to data collection. Information may also come from interviewer observations obtained during data collection. The availability of such additional data may depend on the country and some restrictions on data linkage may apply due to confidentiality and data disclosure concerns. The analysis highlights the usefulness of field process data (paradata) to inform interviewing calling strategies. This also has implications for interviewer training and interviewers will need to receive guidance on the type of data to be collected. In particular, careful consideration should be given to what kind of data should be recorded for each call, such as interviewer observations about the household and information obtained from neighbours.

The significant interviewer effects imply that survey agencies may have a greater choice than previously thought regarding how best to contact a household, rather than, as was hypothesised in Purdon et al. (1999), simply decisions on the timing of calls. For example, certain interviewers may be allocated to more difficult times or cases – at least within fieldwork constraints such as travelling times and costs. It may also be advantageous for the survey organisation to be aware of other time commitments of interviewers; for example interviewers who have only a limited capacity to make evening and weekend calls may need additional support or may be allocated certain cases or areas.

The paper also provides guidance to academic researchers and survey practitioners on how best to use paradata collected in the field and contributes to the methodological developments in the specification of response propensity models based on such data. The

paper aims to contribute to the development of a theoretical framework for the analysis and definition of interviewer calling behaviours and strategies to establish contact. The estimated response propensities obtained from the event history models may ultimately be used for adjustment and estimation at the data analysis stage.

7. Acknowledgement

The research is 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. This work contains statistical data from ONS which is Crown copyright and reproduced with the permission of the controller of HMSO and Queen's Printer for Scotland. The use of the ONS statistical data in this work does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data. This work uses research datasets which may not exactly reproduce National Statistics aggregates.

References

Beerten, R. and Freeth, S. (2004) Exploring Survey Nonresponse in the UK: The Census-Survey Nonresponse Link Study. 59th *Annual Meeting of the American Association for Public Opinion Research (AAPOR)*, 1-16.

Blom, A.G. and Blohm, M. (2007) The Effects of First Contact by Phone: Evidence for the European Social Survey. Paper presented at the *18th International Workshop on Survey Household Nonresponse*, Southampton, UK.

Brick, J.M. and Allen, B. and Cunningham, P. (1996) Outcomes of a Calling Protocol in a Telephone Survey. *Proceedings of the Survey Research Methods Section, American Statistical Association,* 142-149.

Browne, W.J. (2009) *MCMC Estimation in MLwiN v2.1*. Centre for Multilevel Modelling, University of Bristol.

Cobben, F. and Schouten, B. (2007) A follow-up with basic questions of nonrespondents to the Dutch Labour Force Survey. *Discussion paper 07011, Statistics Netherlands*, available from:http://www.cbs.nl/NR/rdonlyres/42B0DC4A-D071-46F8-A249-408D7786E2F3/0/200711x10pub.pdf

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

Cunningham, P., Martin, D. and Brick, M. (2003) An Experiment in Call Scheduling. *Proceedings of the Section on Survey Research Methods, American Association for Public Opinion Research*, 2003, 59-66.

De Heer, W. (1999) International Response Trends: Results of an International Survey. *Journal of Official Statistics*, 15, 2, 129-142.

Durrant, G.B. and Steele, F (2009) Multilevel Modelling of Refusal and Noncontact Nonresponse in Household Surveys: Evidence from Six UK Government Surveys. *J. R. Statist. Soc. A,* 172, 2, 361-381.

Durrant, G.B., Groves, R., Staetsky, L. and Steele, F. (2009) Effects of Interviewer Attitudes and Behaviours on Refusal in Household Surveys, *Public Opinion Quarterly*, 74,1, 1-36.

Elliott, M.R., Little, R.J.A. and Lewitzky, S. (2000) Subsampling Callbacks to Improve Survey Efficiency. *Journal of the American Statistical Association*, *Applications and Case Studies*, 95, 451, 730-738.

Fuller, W. A. (1987) *Measurement Error Models*, Wiley, New York.

Goldstein, H. (2010) *Multilevel Statistical Models*, 4th ed., Chichester: Wiley.

Greenberg, B.S. and Stokes S.L. (1990) Developing an Optimal Call Scheduling Strategy for a Telephone Survey. *Journal of Official Statistics,* 6**,** 4, 421-435.

Groves, R.M. and Couper, M.P. (1996) Contact-Level Influences on Cooperation in Face-to-Face Surveys. *Journal of Official Statistics*, 12, 1, 63-83.

Groves, R.M. and Couper, M.P. (1998) *Nonresponse in Household Interview Surveys*, New York.

Groves, R.M. and Heeringa, S.G. (2006) Responsive Design for Household Surveys: Tools for Actively Controlling Survey Errors and Costs, *J. R. Statist. Soc. A,* 169, 3, 439-459.

Herzog, T. N., Scheuren, F. J. and Winkler, W. E. (2007) *Data Quality and Record Linkage Techniques*. New York: Springer.

Kulka, R.A. and Weeks, M.F. (1988) Towards the Development of Optimal Calling Protocolls for Telephone Surveys: A Conditional Probabilities Approach. *Journal of Official Statistics*, 4,4, 319-358.

Laflamme, F. (2009) Experiences in Assessing, Monitoring and Controlling Survey Productivity and Costs at Statistics Canada, paper presented at the 57th Session of the International Statistical Institute Conference, held on 16th - 22nd August 2009, Durban, South Africa.

Laflamme, F., Maydan, M. and Miller, A. (2008) Using Paradata to Actively Manage Data Collection Survey Process, *Proceedings of the Survey Research Methods Section, American Statistical Association*, forthcoming.

Lahiri, P. and Larsen, M. (2005), Regression analysis with linked data, *Journal of the American*

O'Muircheartaigh, C. and Campanelli, P. (1999) A Multilevel Exploration of the Role of Interviewers in Survey Nonresponse. *J. R. Statist. Soc. A,*162, 3, 437-446.

Pickery, J. and Loosveldt, G. (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.

Purdon, S., Campanelli, P. and Sturgis, P. (1999) Interviewer's Calling Strategies on Face-to-Face Interview Surveys. *Journal of Official Statistics,* 15, 2, 199-216.

Rasbash, J., Charlton, C., Jones, K. and Pillinger, R. (2009b) *Manual Supplment to MLwiN v2.1*. Centre for Multilevel Modelling, University of Bristol.

Rasbash, J., Steele, F., Browne, W.J. and Goldstein, H. (2009a) *A User's Guide to MLwiN v2.1*. Centre for Multilevel Modelling, University of Bristol.

Scheuren, F., and Winkler, W. E. (1993) Regression Analysis of Data Files that are Computer. Matched, *Survey Methodology*, 19, 1, 39-58.

Schnell, R. and Kreuter, F. (2001) Separating Interviewer and Sampling Point Effects. *Journal of Official Statistics*, 21, 3, 389-410.

Spiegelhalter, D. J., Best, N.G., Carlin, B. P. and van der Linde, A. (2002) Bayesian Measures of Model Complexity and Fit (with Discussion). *J. R. Statist. Soc.* B, 64, 583-639.

Statistical Association, 100, 222-230.

Steele, F., Diamond, I. and Amin, S. (1996) Immunization Uptake in Rural Bangladesh: a Multilevel Analysis. *J. R. Statist. Soc. A,* 159, 289-299.

Steele, F., Goldstein, H. and Browne, W. (2004). A General Multistate Competing Risks Model for Event History Data, with an Application to a Study of Contraceptive Use Dynamics. *Statistical Modelling*, 4, 2, 145-159.

Swires-Hennessy, E. and Drake, M. (1992). The Optimum Time at Which to conduct Interviews, *Journal of the Market Research Society*, 34, 1, 61-72.

Von Sanden, N. D. (2004) *Interviewer Effects in Household Surveys: Estimation and Design*, PhD Thesis, University of Wollongong. Available from: http://www.library.uow.edu.au/adt-NWU/public/adt-NWU20051202.102014/index.html

Weeks, M.F., Jones, B.L., Folsom, R.E. and Benrud, C.H. (1980) Optimal Times to Contact Sample Households. *Public Opinion Quarterly*, 44, 1, 101-114.

Weeks, M.F., Kulka, R.A. and Pierson, S.A. (1987) Optimal Call Scheduling for a Telephone Survey. *Public Opinion Quarterly,* 51, 4, 540-549.

Wood, A.M. and White, I.R. (2006). Using Number of Failed Contact Attempts to Adjust for Non-ignorable Non-response. *J. R. Statist. Soc. A,* 169, 3, 525-542.

Table 1: Probability of contact at first call, by day and time of call.

		Contact probability	Total number of first calls made	% of all first calls
Monday	Morning	0.46	682	4.1
	Afternoon	0.49	3310	19.8
	Evening	0.67	947	5.7
Tuesday	Morning	0.39	505	3.0
	Afternoon	0.48	2796	16.7
	Evening	0.63	810	4.8
Wednesday	Morning	0.36	327	2.0
	Afternoon	0.47	2176	13.0
	Evening	0.61	683	4.1
Thursday	Morning	0.44	290	1.7
	Afternoon	0.46	1864	11.1
	Evening	0.59	492	2.9
Friday	Morning	0.39	221	1.3
	Afternoon	0.42	1014	6.1
	Evening	0.57	286	1.7
Saturday	Morning	0.50	60	0.4
	Afternoon	0.53	202	1.2
	Evening	0.43	51	0.3
Sunday	Morning	0.50	10†	<1.0
	Afternoon	0.50	16†	<1.0
	Evening	0.67	9†	<1.0
Total		--	16799	100

Morning: 0.00-12.00, Afternoon: 12.00-17.00, Evening: 17.00-0.00
† indicates cells with a sample size of less than 30

Table 2: Estimated coefficients (and standard errors) for two multilevel cross-classified logistic models for contact: Model A without census variables and Model B with census variables.

Variable (ref= Reference category)	**Categories**	**Model A** $\hat{\beta}$ $(ste(\hat{\beta}))$	**Model B** $\hat{\beta}$ $(ste(\hat{\beta}))$
Constant		0.011 (0.086)	-0.870 (0.111)***
Survey indicator (ref = EFS)	FRS	0.076 (0.054)	0.077 (0.050)
	GHS	0.052 (0.047)	0.022 (0.044)
	OMN	0.171 (0.049)***	0.064 (0.045)
	NTS	-0.026 (0.049)	-0.008 (0.046)
	LFS	0.682 (0.053)***	0.280 (0.057)***
Call Record Data (time variant variables)			
Previous call indicator (ref= First call)	Call previously made	-0.645 (0.061)***	-0.550 (0.060)***
Number of calls previously made		-0.083 (0.009)***	-0.111 (0.009)***
Day and time of call (ref = Sun-Wed eve)	Mo-Wed am	-0.536 (0.144)***	-0.305 (0.196)
	Mo-Wed pm	-0.541 (0.084)***	-0.457 (0.115)***
	Thu-Fri am	-0.727 (0.208)***	-1.110 (0.284)***
	Thur-Fri pm	-0.792 (0.111)***	-0.625 (0.146)***
	Thu-Fri eve	-0.087 (0.113)	-0.118 (0.152)
	Sat-Sun am	-0.600 (0.379)	-0.282 (0.493)
	Sat-Sun pm	-0.281 (0.234)	-0.346 (0.306)
	Sat eve	0.053 (0.644)	-2.472 (1.651)
Time of previous call (ref= Weekday evening)	Weekend	0.704 (0.147)***	0.615 (0.141)***
	Weekday morning	-0.008 (0.104)	-0.018 (0.104)
	Weekday afternoon	0.175 (0.052)***	0.172 (0.052)***
Number of days between calls (ref= Same day)	1-3 days	0.095 (0.043)**	0.089 (0.042)**
	4-8 days	0.257 (0.046)***	0.245 (0.045)***
	9-14 days	0.332 (0.080)***	0.311 (0.080)***
	15+ days	0.428 (0.154)***	0.290 (0.155)*
Card/message left (ref= No card/message left)	Card/message left	0.104 (0.035)***	0.095 (0.035)***
Interviewer Observations (time invariant)			
Security device (ref= security device visible)	No security device visible	0.210 (0.030)***	0.192 (0.031)***
Type of accommodation (ref= Not house, i.e. flat, mobile home, other)	House	0.467 (0.058)***	0.350 (0.057)***
Houses in area in good or bad state of repair (ref= Good)	Fair-Bad	-0.238 (0.052)***	-0.186 (0.050)***
House in a better or worse condition than others in area (ref= Better)	About the same	-0.127 (0.039)***	-0.068 (0.040)
	Worse	-0.308 (0.056)***	-0.272 (0.056)***
Dependent children present (ref= Not present)	Present	0.323 (0.059)***	----
Household-level variables from the Census (time invariant)			
Age (household reference person) (ref= 16 - 34)	35 - 49	----	0.165 (0.033)***
	50 - 64	----	0.389 (0.038)***
	65 - 79	----	0.444 (0.069)***
	80 and older	----	0.535 (0.080)***
Household type (ref= Single household)	Couple household	----	0.425 (0.027)***
	Multiple household	----	0.402 (0.075)***
Pensioner in household (ref= No pensioner in household)	Pensioner in household	----	0.113 (0.082)
Person with a limiting long term illness present (LLTI) (ref= Not present)	Household with one or more people with LLTI	----	0.085 (0.055)
Dependent children present (ref= Not present)	Present	----	0.557 (0.054)***
Adults in employment (ref= No)	Yes	----	0.120 (0.064)**

Interviewer-level Variables (time invariant)			
Pay grade (ref= Merit 1 and 2)	Interviewer and advanced interviewer Merit 3 and field manager	0.144 (0.038)*** 0.128 (0.043)***	0.079 (0.047)* 0.129 (0.057)**
Interviewer qualification (ref= Degree or postgraduate, other higher education)	A levels GCSE, qualifications below this level, no qualification	-0.110 (0.047)** -0.022 (0.035)	-0.148 (0.059)** -0.032 (0.043)
Interviewer Age (ref= 50 years or more)	Under 50 years	-0.122 (0.056)**	-0.142 (0.062)**
Use phone to make appointment (ref= Always, frequently, sometimes)	Rarely, never	0.097 (0.033)***	0.103 (0.041)**
Interactions between interviewer observations and household characteristics			
Day and time of call * Dependent children present (ref=Sun-Wed eve and No dependent children)	Mo-Wed am * Children Mo-Wed pm * Children Thu-Fri am * Children Thu-Fri pm * Children Thu-Fri eve * Children Sat-Sun am * Children Sat-Sun pm * Children Sat eve * Children	-0.416 (0.131)*** -0.256 (0.074)*** -0.260 (0.190) -0.191 (0.093)** -0.043 (0.110) 0.187 (0.404) -0.152 (0.230) 0.063 (0.578)	-0.090 (0.126) 0.146 (0.069)** -0.093 (0.187) 0.061 (0.090) -0.155 (0.098) -0.613 (0.358)* -0.116 (0.207) -0.267 (0.524)
Day and time of call * Adults in employment (ref= Sun-Wed eve and No adults in employment)	Mo-Wed am * Yes Mo-Wed pm * Yes Thu-Fri am * Yes Thu-Fri pm * Yes Thu-Fri eve * Yes Sat-Sun am * Yes Sat-Sun pm * Yes Sat eve * Yes	---- ---- ---- ---- ---- ---- ---- ----	-0.552 (0.143)*** -0.590 (0.080)*** -0.083 (0.202) -0.591 (0.103)*** 0.034 (0.118) -0.381 (0.364) -0.028 (0.243) 2.669 (1.518)*
Day and time of call * Household with a person with limiting long term illness (LLTI) (ref= Sun-Wed eve and No person with LLTI)	Mo-Wed am * LLTI Mo-Wed pm * LLTI Thu-Fri am * LLTI Thu-Fri pm * LLTI Thu-Fri eve * LLTI Sat-Sun am * LLTI Sat-Sun pm * LLTI Sat eve * LLTI	---- ---- ---- ---- ---- ---- ---- ----	0.152 (0.118) 0.315 (0.069)*** 0.193 (0.166) 0.131 (0.087) -0.045 (0.104) 0.369 (0.297) 0.274 (0.199) 0.435 (0.536)
Day and time of call * Pensioner in household (ref= Sun-Wed eve and No pensioner)	Mo-Wed am * Pensioner Mo-Wed pm * Pensioner Thu-Fri am * Pensioner Thu-Fri pm * Pensioner Thu-Fri eve * Pensioner Sat-Sun am * Pensioner Sat-Sun pm * Pensioner Sat eve * Pensioner	---- ---- ---- ---- ---- ---- ---- ----	0.342 (0.153)** 0.318 (0.088)*** 0.629 (0.213)*** 0.246 (0.113)** 0.034 (0.128) -0.717 (0.385)*** 0.069 (0.265) 1.600 (1.551)
Day and time of call * Indicator if house (ref= Sun-Wed eve and and Not house)	Mo-Wed am * House Mo-Wed pm * House Thu-Fri am * House Thu-Fri pm * House Thu-Fri eve * House Sat-Sun am * House Sat-Sun pm * House Sat eve * House	-0.531 (0.139)*** -0.258 (0.078)*** -0.338 (0.199)* -0.035 (0.104) -0.040 (0.105) 0.106 (0.347) -0.065 (0.214) -0.371 (0.567)	-0.519 (0.145)*** -0.191 (0.078)** -0.158 (0.201) 0.065 (0.104) 0.048 (0.100) 0.311 (0.357) -0.090 (0.214) -0.110 (0.564)
Day and time of call * Indicator if house in a good or bad state of repair (ref= Sun-Wed eve and Good)	Mo-Wed am * Fair/Bad Mo-Wed pm * Fair/Bad Thu-Fri am * Fair/Bad Thu-Fri pm * Fair/Bad Thu-Fri eve * Fair/Bad Sat-Sun am * Fair/Bad Sat-Sun pm * Fair/Bad Sat eve * Fair/Bad	0.012 (0.117) 0.198 (0.066)*** 0.536 (0.163)*** 0.243 (0.085)*** 0.157 (0.092)* 0.509 (0.327) -0.200 (0.202) 0.031 (0.496)	0.036 (0.120) 0.150 (0.065)** 0.631 (0.169)*** 0.199 (0.085) 0.120 (0.090) 0.485 (0.327) -0.144 (0.197) -0.168 (0.483)

Day and time of call * Time of previous call (ref= Sun-Wed eve and Weekday eve)	Mo-Wed am * Weekend	0.078 (0.408)	-0.007 (0.417)
	Mo-Wed pm * Weekend	-0.714 (0.223)***	-0.567 (0.224)**
	Thu-Fri am * Weekend	-0.552 (0.785)	-0.211 (0.766)
	Thu-Fri pm * Weekend	-0.189 (0.460)	0.003 (0.465)
	Thu-Fri eve * Weekend	-0.682 (0.459)	-0.675 (0.443)
	Sat-Sun am * Weekend	-0.240 (0.681)	0.065 (0.667)
	Sat-Sun pm * Weekend	-0.833 (0.306)***	-0.761 (0.297)**
	Sat eve * Weekend	-1.319 (0.587)**	-1.203 (0.580)**
	Mo-Wed am * Weekday am	0.090 (0.245)	0.098 (0.246)
	Mo-Wed pm * Weekday am	0.086 (0.135)	0.156 (0.137)
	Thu-Fri am * Weekday am	0.447 (0.298)	0.492 (0.301)
	Thu-Fri pm * Weekday am	-0.102 (0.168)	0.043 (0.170)
	Thu-Fri eve * Weekday am	0.379 (0.190)**	0.359 (0.185)**
	Sat-Sun am * Weekday am	0.574 (0.524)	0.438 (0.521)
	Sat-Sun pm * Weekday am	0.149 (0.521)	0.214 (0.508)
	Sat eve * Weekday am	0.014 (1.690)	-0.581 (1.628)
	Mo-Wed am * Weekday pm	0.163 (0.143)	0.211 (0.146)
	Mo-Wed pm * Weekday pm	-0.039 (0.067)	-0.009 (0.067)
	Thu-Fri am * Weekday pm	-0.063 (0.179)	-0.074 (0.183)
	Thu-Fri pm * Weekday pm	-0.034 (0.086)	0.014 (0.086)
	Thu-Fri eve * Weekday pm	0.025 (0.087)	-0.021 (0.083)
	Sat-Sun am * Weekday pm	0.772 (0.313)**	0.853 (0.313)***
	Sat-Sun pm * Weekday pm	-0.444 (0.205)**	-0.458 (0.201)**
	Sat eve * Weekday pm	0.108 (0.584)	-0.048 (0.607)
Interactions between interviewer observations and interviewer characteristics			
Day and time of call * Interviewer Age (ref= Sun-Wed eve and 50 years or more)	Mo-Wed am * under 50 yrs	0.096 (0.118)	0.108 (0.123)
	Mo-Wed pm * under 50 yrs	0.017 (0.066)	0.035 (0.067)
	Thu-Fri am * under 50 yrs	0.044 (0.171)	0.130 (0.171)
	Thu-Fri pm * under 50 yrs	-0.023 (0.087)	-0.012 (0.087)
	Thu-Fri eve * under 50 yrs	-0.194 (0.093)**	-0.204 (0.092)**
	Sat-Sun am * under 50 yrs	-0.776 (0.339)**	-0.716 (0.337)**
	Sat-Sun pm * under 50 yrs	0.061 (0.200)	0.029 (0.193)
	Sat eve * under 50 yrs	0.026 (0.443)	-0.142 (0.440)
Interviewer variance	--	0.089 (0.013)***	0.078 (0.011)***
Area variance	--	0.006 (0.005)	0.009 (0.005)*

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
*** significant at the 1% level

Coding of time of call: am = 0.00-12.00, pm=12.00-17.00, evening (eve)= 17.00-0.00

Table 3: Predicted probabilities of contact (in %) for two-way interactions with variable 'day and time of call' (for model B including census variables).†

		Type of accommodation		State of repair of houses in area		Dependent children present		Time of previous call			
		House	Flats, other	Good	Fair-Bad	Present	Not present	Weekend	Weekday am	Weekday pm	Weekday eve
Day and time of call	Mo-Wed am	38.2	42.2	47.9	44.3	54.6	43.2	60.7	48.0	55.4	46.1
	Mo-Wed pm	42.3	38.6	44.2	43.4	56.6	39.6	43.5	45.7	46.3	42.4
	Sun-Wed eve	58.1	49.6	55.4	50.8	63.9	50.6	67.7	53.0	57.6	53.5
	Thu-Fri am	28.6	24.9	29.5	39.3	35.3	25.7	36.6	38.2	29.9	27.9
	Thu-Fri pm	44.5	34.8	40.2	40.5	50.5	35.7	53.3	39.0	42.8	38.4
	Thu-Fri eve	56.4	46.7	52.5	50.9	57.5	47.7	49.1	58.8	54.3	50.6
	Sat-Sun am	58.8	42.7	48.5	55.8	42.4	43.8	62.9	56.8	70.4	46.6
	Sat-Sun pm	47.5	41.2	46.9	39.0	53.0	42.2	41.5	49.8	38.3	45.1
	Sat eve	9.9	7.9	9.8	7.1	10.7	8.2	5.3	5.3	10.3	9.2

		Adults in employment		Pensioner in household		Person with LLTI		Interviewer age	
		No adult	1+	Present	Not present	Present	Not present	Under 50 years	50 years or more
Day and time of call	Mo-Wed am	50.8	40.4	56.3	45.2	51.4	45.6	49.6	50.4
	Mo-Wed pm	47.1	36.0	52.0	41.5	51.7	42.0	44.1	46.7
	Sun-Wed eve	58.6	61.0	55.4	52.6	55.1	53.1	54.4	57.8
	Thu-Fri am	31.9	32.7	43.7	27.2	39.9	27.6	31.4	31.6
	Thu-Fri pm	43.0	32.2	46.1	37.6	43.1	38.0	39.0	42.6
	Thu-Fri eve	55.3	59.0	53.3	49.7	51.1	50.2	46.5	55.0
	Sat-Sun am	51.4	45.0	31.8	45.7	57.2	46.2	31.0	51.0
	Sat-Sun pm	49.8	52.0	48.6	44.2	53.4	44.6	46.7	49.4
	Sat eve	10.9	65.5	34.6	8.9	14.2	9.0	8.3	10.8

† 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.

The call indicator variable has been fixed for call 1 to obtain these predicted probabilities but the trend in predicted probabilities would be the same for subsequent calls since interactions with the call-variable were not included.

Coding of time of call: am = 0.00-12.00, pm=12.00-17.00, evening (eve)= 17.00-0.00