

# Sociological Methods & Research

<http://smr.sagepub.com/>

---

## Doorstep Interactions and Interviewer Effects on the Process Leading to Cooperation or Refusal

Gabriele B. Durrant and Julia D'Arrigo

*Sociological Methods & Research* published online 25 March 2014

DOI: 10.1177/0049124114521148

The online version of this article can be found at:

<http://smr.sagepub.com/content/early/2014/03/24/0049124114521148>

---

Published by:



<http://www.sagepublications.com>

**Additional services and information for *Sociological Methods & Research* can be found at:**

**Email Alerts:** <http://smr.sagepub.com/cgi/alerts>

**Subscriptions:** <http://smr.sagepub.com/subscriptions>

**Reprints:** <http://www.sagepub.com/journalsReprints.nav>

**Permissions:** <http://www.sagepub.com/journalsPermissions.nav>

>> [OnlineFirst Version of Record](#) - Mar 25, 2014

[What is This?](#)

# Doorstep Interactions and Interviewer Effects on the Process Leading to Cooperation or Refusal

Sociological Methods & Research

1-29

© The Author(s) 2014

Reprints and permission:

[sagepub.com/journalsPermissions.nav](http://sagepub.com/journalsPermissions.nav)

DOI: 10.1177/0049124114521148

[smr.sagepub.com](http://smr.sagepub.com)



Gabriele B. Durrant<sup>1</sup> and Julia D'Arrigo<sup>1</sup>

## Abstract

This article presents an analysis of interviewer effects on the process leading to cooperation or refusal in face-to-face surveys. The focus is on the interaction between the householder and the interviewer on the doorstep, including initial reactions from the householder, and interviewer characteristics, behaviors, and skills. In contrast to most previous research on interviewer effects, which analyzed final response behavior, the focus here is on the analysis of the *process* that leads to cooperation or refusal. Multilevel multinomial discrete-time event history modeling is used to examine jointly the different outcomes at each call, taking account of the influence of interviewer characteristics, call histories, and sample member characteristics. The study benefits from a rich data set comprising call record data (paradata) from several face-to-face surveys linked to interviewer observations, detailed interviewer information, and census records. The models have implications for survey practice and may be used in responsive survey designs to inform effective interviewer calling strategies.

---

<sup>1</sup> Southampton Statistical Sciences Research Institute (S3RI), University of Southampton, Southampton, UK

## Corresponding Author:

Gabriele Durrant, Southampton Statistical Sciences Research Institute, University of Southampton, Southampton SO17 1BJ, UK.

Email: [g.durrant@southampton.ac.uk](mailto:g.durrant@southampton.ac.uk)

## Keywords

tailoring, interviewer effects, multilevel modeling, call record data, paradata

## Introduction

The falling response rates in many surveys and countries (Baruch and Holtom 2008; de Heer 1999; de Leeuw and de Heer 2002) have shifted attention to the determinants of the response process in sample surveys. For interviewer-administered surveys, it has been recognized that interviewers play an important role in gaining both contact (e.g., Purdon, Campanelli, and Sturgis 1999) and cooperation from sample survey members (Couper and Groves 1992; Groves and Couper 1998; Jaekle et al. 2009; O'Muircheartaigh and Campanelli 1999; Pickery and Loosveldt 2002, 2004). In face-to-face surveys, the interviewer needs to make one or several visits, referred to as *calls*, to each household to first establish contact and finally cooperation. Often hypothesized to be a key element in persuading sample members to take part in a survey is the *interaction process* between householders and interviewers at each contact (Bates, Dahlhamer, and Singer 2008; Couper and Groves 2002; Groves, Cialdini, and Couper 1992; Groves and Couper 1996; Groves and Heeringa 2006; Groves and McGonagle 2001; Sturgis and Campanelli 1998). This process may be best investigated by analyzing the actual *call process* and the interactions between the householder and the interviewer on the doorstep (Groves and Couper 1998).

This article aims to analyze the influence of the interviewer on the process leading to cooperation or refusal in face-to-face surveys. Of particular interest is the interaction process between the householder and the interviewer, including initial reactions from the householder to the survey request, and interaction effects between householder and interviewer characteristics. The article investigates the role of interviewer strategies, behaviors, and attitudes on the cooperation process. Rather than focusing on the *final* outcome of response or nonresponse as most previous research on interviewer effects (Durrant et al. 2010; Groves and Couper 1998; O'Muircheartaigh and Campanelli 1999; Pickery and Loosveldt 2000), the analysis here is carried out at the call level, focusing on the outcome *at each call*.

Although the analysis of call record data has in recent years found increasing attention (Bates et al. 2008; Groves and Heeringa 2006; Wagner 2013), the influence of the interviewer on a call-by-call basis on the response process has not yet been studied. One reason for this is that it is not obvious how

best to analyze such complex time-dependent call record data. Durrant, D'Arrigo, and Steele (2011, 2013) introduce methodologies to analyze such complex data and showcase this using an application. Durrant et al. (2011) focus on the correlates of time to first contact, in particular on best times to contact a household. Durrant et al. (2013) analyze the correlates of the process to cooperation and refusal using a basic model, including household-level variables and some standard call record variables only but without analysis of interviewer effects and household-interviewer interactions. Their focus is on best times to achieve cooperation. The article here extends this previous research by using the principles of the methods developed but applying them to the investigation of interviewer and sample member interaction effects. More specifically, this analysis asks questions such as "Does the initial householder reaction predict the probability of cooperation at a future call?" "Does the way the interviewer makes contact with a household at a particular call influence cooperation rates at the current or future calls?" "What is the influence of interviewer characteristics, behaviors, and skills, in particular those related to tailoring, on cooperation rates across calls?" and "How do influences of call characteristics and the initial reactions of the householder change depending on the experience and the tailoring ability of the interviewer?"

Most of the previous research on call record data has focused on ways to establish contact with a household (Durrant et al. 2011; Kulka and Weeks 1988; Wagner 2013; Weeks et al. 1980). Since the contact and the cooperation/refusal stages are two quite distinct processes (Groves and Couper 1998; Lynn and Clarke 2002; Nicoletti and Perachi 2005), analyzing time to first contact is therefore not of direct relevance here. To investigate the influence of tailoring and the interaction process on the doorstep, which is the key focus of this article, a call-by-call analysis with information on what happens on the doorstep is most promising. So far, however, only few researchers have started to look at the influence of the interviewer on nonresponse during the call process. Blom, de Leeuw, and Hox (2010) investigate multilevel logistic models using data from the European Social Survey, taking account of households within interviewers and countries, to study cross-country differences. Jaeckle et al. (2009) focus on the effect of interviewer personality traits. However, information on nonresponding units is in both cases limited. Bates et al. (2008) make use of doorstep concerns and call histories to predict survey response, however without controlling for the influence of the interviewer. Purdon et al. (1999) investigate the effects of calling times on cooperation, but only presenting descriptive statistics and not accounting for the clustering of calls and households within interviewers. If interviewer influences were analyzed, then information about the household was usually very

limited, meaning that interaction effects could not be investigated, or call information was not taken into account (e.g., Blom et al. 2010; Groves and Couper 1998; Jaeckle et al. 2009).

In this article, a multilevel multinomial model is used to examine jointly the different outcomes at each call, taking account of the influence of interviewer characteristics, call histories, and sample member characteristics. The multilevel modeling approach for the analysis of call record data, which can be complex with many different hierarchies, was first presented for the cooperation process in Durrant et al. (2013). This method is used and extended here. As in the previous work, the multilevel model accounts for clustering of calls within households and interviewers. Here, and this is new, the focus is on the influence of call variables that describe part of the interaction, interviewer characteristics, skills and behaviors, and cross-level interactions. By including interviewer-level variables, we aim to explain part of the significant interviewer variance.

The work is guided by the conceptual framework for response behavior, as developed in Groves and Couper (1998), and by sociological and psychological concepts (Goyder 1987; Groves et al. 1992). In addition to household characteristics, influences that have been hypothesized to play a key role in the response process include (a) call-level characteristics (Bates et al. 2008; Blom, Lynn, and Jaeckle 2008; Groves and Couper 1996), (b) the initial interviewer–householder interaction on the doorstep (Bates et al. 2008; Groves and McGonagle 2001), and (c) the influence of the interviewer (Blom et al. 2010; Groves and Couper 1998; Groves and McGonagle 2001; O’Muircheartaigh and Campanelli 1999; Pickery, Loosveldt, and Carton 2001). Our analysis aims to investigate the influence of all of these components. In the literature, primarily three factors have been identified that may drive interviewer effects on nonresponse outcomes: interviewer experience (Durbin and Stuart 1951; Groves and Couper 1998), attitudes and confidence of the interviewer (Groves and Couper 1998), and strategies and behaviors. In particular, the importance of being responsive toward the individual sample member, including tailoring of strategies toward respondents and awareness of respondent concerns (Groves and Couper 1998; Morton-Williams 1993) has been stressed. It may be hypothesized that fixed effects of interviewer strategies, behaviors, and tailoring approaches are more important at the call level than for final response analysis.

This article benefits from rich information on interviewers based on a survey of interviewers working for the U.K. Office for National Statistics (ONS). This information was linked to the survey outcome from six U.K. household surveys. An advantage of the study is that rich information from

interviewer observations and U.K. Census records is available for both responding and nonresponding sample members. In addition, comparatively detailed call record information, so-called paradata (Couper 1998), is available, including the time and day of the call, the outcome of the call, initial reactions from household members, and basic characteristics of the person the interviewer talked to on the doorstep.

The research is anticipated to influence field decisions, particularly in adaptive and responsive survey designs, where survey data collection outcomes are continuously monitored allowing early intervention and the alteration of the survey design (Groves and Heeringa 2006; Laflamme, Maydan, and Miller 2008; Kirgis and Lepkowski 2013). The results may guide strategies for interviewer training and performance management. The work may also contribute to further methodological development and to provide guidance to survey practitioners on how to use and analyze call record data and information on interviewers.

The remainder of the article is structured as follows. First, the different linked data sources are discussed. Then, the multilevel multinomial model and the results are presented. The final section provides a summary of the main findings and a discussion of the results.

## Data

### *Information on Interviewers*

The study makes use of the U.K. 2001 Census link study (Durrant et al. 2010), which combines the survey outcome of six face-to-face U.K. household surveys with detailed information on sample member from the U.K. 2001 Census, at the individual and household level, and rich paradata, including interviewer observation data about the households and areas, call record data, and information on interviewers. (Another U.K. Census has recently been conducted; however, linked data of this type have not yet been available for analysis.)

A comprehensive survey of interviewers working for the ONS in 2001 was conducted, designed to coincide with the 2001 Census (Freeth, Kane, and Cowie 2002). The survey collected information on sociodemographic characteristics, interviewer work history and qualification, interviewer attitudes (including attitudes to the persuasion of reluctant respondents and working at different times and days of the week), interviewing behaviors, strategies, and doorstep approaches, in particular indicators of tailoring abilities. Conceptually, this interviewer survey builds on previous work by

Groves and Couper (1998) and Hox and de Leeuw (2002). The response rate was relatively high (84 percent). The interviewer survey was designed to be conducted prior to the fieldwork in question. (A very small number of interviewers filled in the questionnaire slightly later than planned. However, the implications for data analysis are expected to be small.) Participation in the survey was voluntary and interviewers who participated in the survey were paid 1 hour for their time. The survey was not anonymous since identifying information was needed to link the interviewer information to all other data sources.

### *Call Record Data and Interviewer Observations*

The call data contain, in addition to the standard recordings, such as day, time, and outcome of the call, information about the initial reactions of the householder, such as whether the householder asked any questions and whether any positive or negative comments were made. The interviewer also recorded basic characteristics of the householder at each call, such as gender and approximate age, and information on how the contact was established (i.e., face-to-face, via an intercom system, or through a window or door). This type of information, often not collected in standard call record data, allows analysis of the householder–interviewer interaction and possible tailoring effects. The outcome of each call, the dependent variable in our analysis, distinguishes cooperation, refusal, making an appointment for another time, and “postponements.” Such “postponements” are defined as broken appointments and circumstances where the interviewer withdraws to come back later, if the interviewer is unable to make contact with a responsible resident or feels threatened. Cooperation is defined as at least one household member agrees to respond to the survey. The call data are recorded at every call and therefore are call dependent, that is, *time varying*. In addition, the interviewer collected so-called interviewer observation variables, that is, basic information about the household and immediate neighborhood, such as type of accommodation, indications of the presence of children, the condition of the house relative to others in the area, if the house is part of a council housing estate and an indication of how safe the interviewer feels walking in the area after dark. These variables are collected only once—if possible, at the first call. They are therefore *time invariant*.

### *Quality of the Paradata*

Following the increasing use of paradata, the quality of such data has been a topic of recent debate (Sinibaldi, Durrant, and Kreuter 2013; West 2013).

Since call record data are not part of the standard survey data, they are not undergoing the same editing checks and may therefore require further cleaning and editing before use. Paradata may be subject to missing data and measurement error. For this study, we were able to work closely together with the U.K. ONS, which ensured the linkage of the paradata to the survey and census records and reduced the likelihood of errors using additional editing checks. The interviewer observed paradata is subject to a comparatively small amount of missing values ranging from less than 2 percent for standard call variables, such as date and time of the call or type of house, to about 12 percent for more difficult to observe call record variables, such as presence of children. Furthermore, the unique features of this data set, in particular the linkage to the U.K. Census records, made it possible to assess the measurement error properties of some of the interviewer observed variables. Comparing interviewer observations with Census records, such as the type of house, adults in employment, ethnicity, and presence of children, showed a high correspondence between the variables analyzed (between 88 and 97 percent agreement), meaning that the interviewer-observed variables are of comparatively high quality (Sinibaldi et al. 2013).

### *Analysis Sample*

Our analysis sample contains 38,816 contact calls. Noncontact calls are not investigated in this analysis (for an analysis of contact/noncontact calls, see Durrant et al. 2011). Households that were never contacted, vacant, and non-residential addresses, reissues and unusable records were excluded from the analysis. The final analysis sample includes a total of 15,782 households, nested within 565 interviewers. Most of the guidelines to interviewers provided by the survey organization refer to contacting strategies. Some general guidelines are provided to interviewers on how to avoid or deal with a refusal on the doorstep, including calling back at least once after a refusal. It should be noted that the call record data are not from a randomized experiment but reflect observational data, which has both advantages and disadvantages (see also the discussions of this issue in Conrad et al. 2013; Durrant et al. 2013; Groves and Couper 1998; Purdon et al. 1999). The data allow the analysis at each call, in particular the interaction process between householder and interviewer, a key component in determining response behavior. However, causal effects are difficult to assess.

The six face-to-face surveys included in the study 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 Labor Force Survey (LFS). The final refusal rates across the six surveys range from about 14 percent for the LFS to about 30 percent for the EFS, which may be explained by the differences in survey topics, interview length, length of data collection period, and additional requirements such as a diary.

## Statistical Data Analysis

### *The Event History Model*

The analysis makes use of a multilevel multinomial logistic discrete-time event history regression model (Durrant et al. 2013; Steele, Diamond, and Wang 1996), which accounts for the clustering of calls within households and interviewers. The dependent variable in the model,  $y_{tij}$ , the response outcome at call  $t$ , made to household  $i$  by interviewer  $j$  conditional on contact being achieved at call  $t$ , contains four response outcomes: (1) refusal, (2) appointment made, (3) other form of postponement, and (4) cooperation. Modeling the log odds of outcome  $s$  ( $s = 1, 2, 3$ ) relative to outcome 4 (cooperation), the multilevel multinomial logistic model employed here can be expressed as follows:

$$\log\left(\frac{\pi_{tij}^{(s)}}{\pi_{tij}^{(4)}}\right) = \beta^{(s)'} x_{tij}^{(s)} + \delta^{(s)'} z_{ij}^{(s)} + \gamma^{(s)'} g_j^{(s)} + \lambda^{(s)} u_{ij} + \zeta^{(s)} v_j, \quad (1)$$

where  $x_{tij}^{(s)}$  is a vector of time-varying covariates, with coefficient vector  $\beta^{(s)}$ , which includes characteristics of the current call  $t$ , two-way interactions between call level variables, interaction effects between householder and interviewer, and information from calls prior to  $t$ . The vectors  $z_{ij}^{(s)}$  and  $g_j^{(s)}$  represent household and interviewer-level variables, with coefficient vectors  $\delta^{(s)}$  and  $\gamma^{(s)}$ , respectively. The household-level variables include information from both census and interviewer observation data and are added as controls. The household and interviewer random effects  $u_{ij}$  and  $v_j$ , assumed to be normally distributed, represent remaining variation including unobserved household and interviewer-level characteristics. Their variance parameters,  $\sigma_u^2$  and  $\sigma_v^2$ , are the residual between-interviewer and between-household variances in the log odds of refusal versus cooperation, the log odds of appointment versus cooperation and the log odds of postponement versus cooperation. Both random effects are specified with outcome-specific coefficients or “loadings,”  $\lambda^{(s)}$  and  $\zeta^{(s)}$ , respectively (with  $\lambda^{(1)}$  and  $\zeta^{(1)}$  fixed at one for identification; Steele, Sigle-Rushton, and Kravdal 2009). This implies that the unobserved household and interviewer effects can vary across the three

different outcomes (refusal, appointment, and postponement), allowing for a common set of unmeasured household and interviewer characteristics.

To estimate the models, maximum likelihood estimation is employed, using the aML software (Lillard and Panis 2003). To evaluate model fit, likelihood ratio tests are used (Goldstein 2011). This allows the comparison of nested models, for example to evaluate if a model including call-record and interviewer-level variables, leads to an improvement compared to a standard model only including basic household information. Predicted probabilities are derived using the method described in detail in Durrant et al. (2013; see also Rasbash et al. 2009) to help interpret the results, in particular of interaction effects, and to investigate effect sizes, given the large number of contact calls.

### *Details of Modeling and Choice of Explanatory Variables*

A multilevel modeling framework for the analysis of call record data on cooperation was first introduced by Durrant et al. (2013), providing a detailed justification for the method. That article presented a basic model, controlling only for standard household and call-level variables to illustrate the method. The model employed here extends this approach. In the current article, the focus is on the influence of interviewer characteristics, cross-level interaction effects, that is, statistical interaction effects between householder and interviewer characteristics, and the effects of time varying variables, in particular variables that describe the initial interaction with the household on the doorstep. It is of interest to what extent the observable interviewer characteristics explain part of the unexplained interviewer variance. Here, the model conditions on contact being made with the household, since the focus is on the interaction between the householder and interviewer.

The following modeling strategy is used here. First, random effect models without any covariates are fitted to explore the random structure. Then, various sets of explanatory variables are added into the model one at a time to investigate their relevance on the response process. After first controlling for household-level variables, from both census and interviewer observation data, call record variables and finally interviewer characteristics and cross-level interactions are included. The models aim to control for survey design differences. A model including call characteristics and interviewer observation variables but not census variables is also explored.

Motivated by the conceptual framework for response behavior (Groves and Couper 1998), the final model explores the influence of (a) call-level characteristics describing the initial interaction between householder and interviewer, (b) interviewer-level characteristics and strategies, and

(c) cross-level interaction effects, all of which had not been explored in the basic model by Durrant et al. (2013). The model controls for basic call record variables and household characteristics, as already included in the simpler model by Durrant et al. (2013).

The advantage of using a multinomial model, rather than fitting separate binary logistic models for each type of outcome, is that the effects of household and interviewer characteristics on the probability of refusal, appointment, and postponement may be evaluated simultaneously and tested for equivalence. The specification in (1) allows for a different set of covariates to be included in the three outcome equations. For covariates in the model, their effects may differ for the three outcome types and it may be of interest to test whether a given characteristic has the same effect on the three outcomes. The final model investigates differential effects of interviewer characteristics across the three different forms of nonparticipation (dependence of  $g_j^{(s)}$  on  $s$ ). This may have practical implications. It may be hypothesized that particular interviewer characteristics are associated with certain outcomes. For example, female interviewers may be more likely to make appointments or to withdraw if they feel uncomfortable on the doorstep.

Often, cross-level interactions between householder and interviewer characteristics cannot be investigated since either information (or even both) are not available. An advantage of our data set is that the exploration of cross-level interactions, in particular with interviewer doorstep approaches, is possible. We are interested if certain interviewers are more effective in handling more difficult cases. For example, more confident and more experienced interviewers may be expected to be more successful in achieving response from younger people, single households, from people that make negative comments, have questions or households that do not allow face-to-face contact. Further, certain doorstep approaches may have a positive influence on the response behavior of certain types of households. The influence of tailoring strategies may also have an impact on response as main effects. Interviewers who use certain tailoring strategies may perform better on the doorstep. Although sociodemographic interviewer characteristics as main effects have not been found to play a role (Groves and Couper 1998; O'Muircheartaigh and Campanelli 1999), they may have an influence in the interaction with householder characteristics, particularly with characteristics of the person on the doorstep. For example, older interviewers may be more successful in achieving cooperation with older householders. The modeling approach here allows testing for such effects.

## Results

Differential effects of observed and unobserved interviewer characteristics and the influence of time varying call record information on the three nonparticipation outcomes are investigated. First, the random effects structure is discussed and then characteristics of the call including initial reactions of the householder are explored. Finally, main and interaction effects of interviewer characteristics on the process leading to cooperation are presented.

### *Random Interviewer Effects*

We investigate the influence of unmeasured interviewer characteristics, represented by  $\zeta^{(s)}v_j$ , on the three forms of nonparticipation. Different specifications of model (1) were explored. Across all models, we found significant residual variation ( $\sigma_v$ ) in the log odds of a nonresponse outcome between interviewers, which holds also after controlling for household and basic area characteristics, call record variables, and interviewer characteristics. This implies that interviewers have a significant influence on the response outcome of a household, as would be expected in line with previous research on final response outcome (O'Muircheartaigh and Campanelli 1999; Pickery and Loosveldt 2002; Durrant et al. 2010). Slightly surprisingly, we did not find evidence for differential random interviewer effects on the three nonparticipation outcomes due to unobserved interviewer characteristics: The loadings  $\zeta^{(s)}$  ( $s = 1, 2, 3$ ) on the interviewer random effect are assumed to be equal. (The likelihood ratio test statistic for a test of the null hypothesis  $H_0: \zeta^{(1)} = \zeta^{(2)} = \zeta^{(3)} = 1$  is 2.80 on 2 *df*,  $p = .246$ .) This means that unmeasured interviewer characteristics have the same effect on the log odds of each of the three nonparticipation outcomes. The final models explored are therefore a simplification of model with a random effect for interviewers,  $v_j$ , but without loading  $\zeta^{(s)}$  (i.e., the interviewer random effect loadings  $\zeta^{(s)}$  are constrained to be equal to one).

Table 1 presents an overview of estimated household and interviewer random effect parameters for different specifications of the multilevel multinomial model, with the null model (model 0) only including random effects, model 1 with added household-level variables from both the census and interviewer observations, model 2 with added call record variables and model 3, the final model, with also interviewer level variables. We can see that, even after controlling for call-record and household characteristics, the common interviewer random variance remains significant, supporting the

**Table 1.** Estimated Household and Interviewer Random Effects from Different Specifications of the Multilevel Multinomial Logistic Regression Model.

Model	Log Likelihood	Parameter Estimates (Standard Error)						
		Household		Household Random Effect Loadings $\lambda^{(s)}$			Interviewer Common Standard Deviation $\sigma_v$	
		Common Standard Deviation $\sigma_u$	$\lambda^{(1)a}$ Refusal	$\lambda^{(2)}$ Appointment	$\lambda^{(3)}$ Other postponement			
Model 0 (random effects only)	-43,610.39	4.492 (0.133)***		0.040 (0.004)***	0.458 (0.013)***	0.224 (0.018)***		
Model 1 (added household-level variables)	-43,322.66	4.182 (0.115)***		0.040 (0.005)***	0.469 (0.012)***	0.203 (0.018)***		
Model 2 (added call record variables)	-29,956.25	1.900 (0.096)***		0.873 (0.041)***	1.280 (0.055)***	0.755 (0.043)***		
Model 3 (added interviewer level and interaction effects)	-29,899.60	1.900 (0.095)***		0.875 (0.040)***	1.286 (0.055)***	0.734 (0.043)***		

Note: Likelihood ratio test statistics: Between model 0 and 1: Likelihood ratio test statistic is  $2 \times 287.73$ , on 33 df,  $p = .000$ . Between model 1 and 2: Likelihood ratio test statistic is  $2 \times 13,366.41$ , on 84 df,  $p = .000$ . Between model 1 and 3: Likelihood ratio test statistic is  $2 \times 13,423.06$ , on 108 df,  $p = .000$ . Between model 2 and 3: Likelihood ratio test statistic is  $2 \times 56.65$ , on 24 df,  $p = .000$ .

<sup>a</sup>Constrained to equal 1.

\*\*\*Significantly different from zero at the 1 percent level.

hypothesis that interviewers indeed play an important role on the nonresponse outcome at a particular call. Adding in interviewer-level and interaction effects explains part of this variation as would be expected and significantly improves the fit of the model (see likelihood ratio test between models 2 and 3). However, only a relatively small reduction in the interviewer variance can be observed. At the household level, the results also show significant residual variation in the log odds of a nonresponse outcome between households across all models. Contrary to the common interviewer random effect, there is, in addition, evidence of differential effects of unmeasured household characteristics  $u_{ij}$  across the three outcomes (based on  $t$ -tests that the loadings for postponement and appointment are equal to one:  $t = 3.1$ ,  $p = .002$  for  $H_0: \lambda^{(2)} = 1$  and  $t = 5.1$ ,  $p = 0.000$  for  $H_0: \lambda^{(3)} = 1$ ). These differential effects indicate a stronger household effect for postponement and a weaker effect for appointments across all models. As one may expect, the household random effect reduces by about half when household and call record variables are entered (models 1 and 2), implying that a significant part of the household variation is explained by these characteristics. The household random effect remains stable when interviewer effects are included (model 3). The likelihood ratio test statistic between models 1 and 2 indicates that adding in call record variables significantly improves the fit of the model which supports findings in Bates et al. (2008) that information on the call “greatly improve” nonresponse models. Slightly surprising at first sight is the observed increase in the interviewer random effect when call record variables are entered. However, as discussed in Snijders and Bosker (1999:217, 228-29) in the case of a multilevel logistic model, entering highly significant level-1 variables (here the call record variables) tends to increase the random effects of higher levels (here interviewer random effects), which is explained on the basis of the threshold representation.

### *Household–interviewer Doorstep Interactions*

Table 2 presents parameter estimates of two multilevel multinomial models. For easier comparisons, we have included the estimates from the basic model (a), which only controls for time-invariant household-level variables and some basic call record variables but neither interviewer nor interaction effects (this model was presented in Durrant et al. 2013). Then, the parameter estimates for the extended model (b) are presented, including in addition call record variables describing the interaction at the doorstep, interviewer characteristics, and cross-level interaction effects, which is the focus of this article.

**Table 2.** Estimated Coefficients (and Standard Errors in Parentheses) of Two Multilevel Multinomial Logistic Models: (A) Basic Model Controlling Only for Household Characteristics and Basic Call Variables (but without interaction call variables, interviewer characteristics, and cross-level interactions; model presented in Durrant et al. 2013), and (B) Extended Model Including in Addition Call Variables from the Contact Process at the Doorstep, Interviewer Level Variables and Cross-level Interviewer Interactions (Model 3).

Variable (ref = Reference category)	Categories	(A) Basic Model			(B) Extended Model		
		$\hat{\beta}$ (SE( $\hat{\beta}$ )) Refusal	$\hat{\beta}$ (SE( $\hat{\beta}$ )) Appointment Made	$\hat{\beta}$ (SE( $\hat{\beta}$ )) Other Postponement	$\hat{\beta}$ (SE( $\hat{\beta}$ )) Refusal	$\hat{\beta}$ (SE( $\hat{\beta}$ )) Appointment Made	$\hat{\beta}$ (SE( $\hat{\beta}$ )) Other Postponement
Constant		-2.073 (0.314)	2.136 (0.129)***	0.954 (0.168)***	-2.737 (0.222)***	0.157 (0.155)	-0.715 (0.186)***
Basic Call Variables (time varying call record variables)							
Previous contact indicator (ref = First contact)	Contact previously made	-0.738 (0.111)***	-1.704 (0.073)***	-2.002 (0.081)***	-0.242 (0.109)**	-1.597 (0.076)***	-1.862 (0.089)***
Number of contact calls previously made	—	-1.282 (0.053)***	-1.012 (0.029)***	-0.838 (0.031)***	-1.376 (0.051)***	-1.193 (0.036)***	-1.162 (0.038)***
Number of noncontact calls made until first contact	—	0.082 (0.033)**	-0.093 (0.015)***	-0.152 (0.023)***	0.049 (0.021)**	-0.175 (0.015)***	-0.277 (0.020)***
Number of intermediate noncontact calls after first contact was made	—	0.692 (0.040)***	0.411 (0.023)***	0.390 (0.029)***	0.522 (0.034)***	0.444 (0.026)***	0.378 (0.032)***
Day and time of contact (ref = Sun-Mon-Tue eve)	Sun-Mon-Tue a.m.	0.408 (0.174)**	-0.067 (0.097)	0.258 (0.118)**	0.451 (0.180)**	0.153 (0.121)	-0.229 (0.150)
	Sun-Mon-Tue p.m.	-0.375 (0.112)***	-0.590 (0.061)***	-0.153 (0.075)**	-0.388 (0.105)***	-0.708 (0.071)***	-0.364 (0.087)***
	Wed-Sat a.m.	-0.003 (0.148)	-0.236 (0.083)***	-0.037 (0.100)	-0.102 (0.154)	-0.663 (0.108)***	-0.123 (0.131)
	Wed-Sat p.m.	-0.366 (0.107)***	-0.444 (0.059)***	-0.131 (0.072)**	-0.491 (0.102)***	-0.501 (0.070)***	-0.141 (0.085)**
	Wed-Sat eve	-0.103 (0.098)	-0.101 (0.054)**	0.077 (0.066)	0.145 (0.097)	-0.128 (0.066)**	0.018 (0.080)
Previous Appointment Indicator (ref = No prior appointment made)	Prior appointment made	-1.888 (0.109)***	-1.887 (0.059)***	-1.555 (0.068)***	-1.998 (0.201)***	-2.003 (0.116)***	-1.998 (0.143)***

(continued)

**Table 2. (continued)**

Variable (ref = Reference category)	(A) Basic Model			(B) Extended Model		
	$\hat{\beta}$ (SE( $\hat{\beta}$ )) Refusal	$\hat{\beta}$ (SE( $\hat{\beta}$ )) Appointment Made	$\hat{\beta}$ (SE( $\hat{\beta}$ )) Other Postponement	$\hat{\beta}$ (SE( $\hat{\beta}$ )) Refusal	$\hat{\beta}$ (SE( $\hat{\beta}$ )) Appointment Made	$\hat{\beta}$ (SE( $\hat{\beta}$ )) Other Postponement
Contact Process on the Doorstep (Time Varying Call						
Record Variables	Not included in basic model					
How contact was made at doorstep (ref = Face-to-face)						
Question made by householder during the interviewer introductory conversation (ref = No question made)						
Comment made by householder during the interviewer introductory conversation (ref = No comment made)						
Age of main person the interviewer talked to (ref = 60 and over)						
Gender of main person the interviewer talked to (ref = Male)						
Interviewer-level Variables (Time Invariant)	Not included in basic model					
Interviewer experience (ref = 3 years or more)						
Interviewer qualification (ref = Degree or postgraduate, other higher education)						
Can convince reluctant respondents (ref = Less confident)						
Should persuade most reluctant respondent (ref = Strongly agree/agree)						
Can use wide variety of approaches (ref = Strongly agree/agree)						

**Table 2. (continued)**

Variable (ref = Reference category)	(A) Basic Model			(B) Extended Model		
	$\hat{\beta}$ (SE( $\hat{\beta}$ )) Refusal	$\hat{\beta}$ (SE( $\hat{\beta}$ )) Appointment Made	$\hat{\beta}$ (SE( $\hat{\beta}$ )) Other Postponement	$\hat{\beta}$ (SE( $\hat{\beta}$ )) Refusal	$\hat{\beta}$ (SE( $\hat{\beta}$ )) Appointment Made	$\hat{\beta}$ (SE( $\hat{\beta}$ )) Other Postponement
Cross-level Interviewer Interactions	Not included in basic model					
Interviewer experience $\times$ comment made (ref = 3 years or more and No comment made)	Up to 2 years* Positive/neutral			-0.941 (0.280)***	-0.334 (0.100)***	-0.817 (0.124)***
	Up to 2 years*At least one negative comment			-0.729 (0.205)***	-0.210 (0.145)	-0.677 (0.155)***
Can use wide variety of approaches $\times$ question asked (ref = Strongly agree/agree and No question made)	Neither agree nor disagree* At least one question			0.414 (0.173)**	0.494 (0.116)***	0.608 (0.144)***
	Disagree/strongly disagree* At least one question			-0.418 (0.231)*	0.087 (0.146)	-0.039 (0.186)

Note: Time-invariant variables included in both model as controls (coefficients not shown here; for further information on their effects, see Durrant et al. 2013): Interviewer observations: type of accommodation, house in a better or worse condition than others in area; Household-level variables: household type, preschool children present, London indicator, urban/rural indicator, indicator if adults in employment, qualification of household reference person, survey indicator. The models are estimated using full information maximum likelihood. As a closed form solution to the maximum likelihood function does not exist, the residuals at each level are "integrated out" numerically using Gauss-Hermite quadrature. The number of quadrature points used is 16. Approximate standard errors are computed based on an approximation to the Hessian matrix. The missing value categories have been suppressed to save space. Coding of time of call: a.m. = .00–12.00, p.m. = 12.00–17.00, eve = 17.00–00.

\*Significant at the 10 percent level.  
 \*\*Significant at the 5 percent level.  
 \*\*\*Significant at the 1 percent level.

Comparing both models (models A and B), the size of the effects of the basic call record variables, change slightly, but there are no noticeable differences in the significance and the direction of these effects, as one would expect. The basic call record variables, such as the day and time of the call, if an appointment was previously made, number of calls until first contact, and number of intermediate noncontact calls are all found to be significant in predicting response at a call. For example, if the previous call was an appointment, this reduces the probability of a refusal at the next call. These basic call record effects were already explored in the example model in Durrant et al. (2013) to illustrate the method of analyzing time varying variables and are not discussed here further.

We now turn to the effects of the initial interaction between householder and interviewer on the doorstep and the effects of interviewer characteristics (model B), the main focus of this article. As may be expected, the way the contact between the interviewer and the householder was made on the doorstep—if directly face-to-face, or indirectly via a closed window or door, or through an intercom system—seems to make a difference: Direct contacts lead to lower rates of refusals, appointments, and postponements. Not opening the door for the interviewer may indicate a higher suspicion toward strangers (It should be noted that this effect remains after controlling for household and area characteristics.).

The direct response from the householder on the doorstep is found to be a good indicator of cooperation, in line with the findings in Bates et al. (2008). If the householder shows some interest by asking at least one question refusal, appointments and postponements are less likely to occur. Similarly, if the householder makes at least one positive or neutral comment, the likelihood for a refusal or a postponement is reduced in comparison to no comment made. Interestingly, the likelihood for an appointment increases in this case. As might be expected, if the householder makes at least one negative comment, refusal, appointments, and postponements are significantly more likely. The findings support the results in Bates et al. (2008) who also report significant effects of variables describing the initial reaction of the householder to the survey request. Such information improves models predicting response in comparison to models based only on time-invariant information or models only accounting for the basic call history, such as number of previous calls.

In addition, characteristics of the person on the doorstep, such as gender and approximate age, seem to be useful in predicting the outcome of a call. Women are significantly more likely to make appointments than men; postponements are also more likely to occur. This may reflect a greater

reluctance toward strangers or a fear of crime among women (Clemente and Kleiman 1977; Morton-Williams 1993). Other factors due to differences in lifestyles may also contribute to this effect. Women may be more likely to be looking after children when at home and may prefer the interviewer to call back at a more convenient time. No differences in the immediate refusal rates between men and women are observed. The older the householder on the doorstep, the less likely are refusal, appointments, and postponements, in particular for householders aged 60 years and older. If the person is less than 16, refusals, appointments, and postponements are highest.

Interestingly, with an increasing number of contact calls refusals, appointments and postponements seem to be less likely to occur, that is, the odds of cooperation increase with each additional contact made. This supports the findings of Groves and Heeringa (2006) and Sangster and Meekins (2004) who also report a significant positive effect of a prior contact with the household on the likelihood of a main interview. The number of contact calls may be interpreted as an indirect measure of an interaction between the householder and interviewer. The finding may imply that an ongoing interaction between the interviewer and the householder may be more likely to lead to a positive outcome, which would support the interaction hypothesis of Groves and Couper (1996, 1998) that an ongoing interaction may impact positively on the likelihood of response. The effect could also indicate that interviewers are persistent in returning to a household if they feel they have a chance of a positive outcome.

### *Influence of Interviewer Characteristics and Cross-level Interactions*

A range of interviewer level characteristics are explored that may constitute part of the significant interviewer variance. We first comment on potential interviewer main effects before describing cross-level interactions. First, the model controls for interviewer work experience and qualification, often found to be significant predictors in the analysis of interviewer effects (Groves and Couper 1992; Groves and Couper 1998; Hox and de Leeuw 2002). We find that interviewers with up to two years' experience have significantly higher immediate refusal rates than interviewers with three or more years. In fact, when we used the experience variable with a finer categorization of four groups we found that with increasing number of years of experience the immediate refusal rate decreased, and was particularly low for interviewers with nine or more years' experience. This supports findings in Groves and Couper (1998) who report a linear relationship between (final) response rates and length of experience. Interestingly, interviewers with

lower levels of experience also show higher appointment and postponement rates. One may expect higher refusal rates for interviewers with a lower qualification. Here, no noticeable differences in the immediate refusal rates for different levels of interviewer qualification have been found, at least not after controlling for interviewer experience. Interviewers with low or no qualifications are less likely to experience appointments or postponements.

In line with previous research, we find significant evidence for the importance of interviewer attitudes (Groves and Couper 1998; Hansen 2007; Hox and de Leeuw 2002). An indicator of the interviewer's self-confidence seems to be predictive of the outcome: Interviewers who are more confident in their ability to convince reluctant respondents have indeed significantly less refusals. Interestingly, these interviewers also experience significantly less appointments and postponements. Interviewers who agree that they should persuade most reluctant respondents are also significantly less likely to experience a refusal. No differences on making appointments are observed.

With regard to making appointments, we also find that all of the household-level variables included in the model reveal significant effects on making appointments, for example households with children, living in a house or in an urban area are significantly more likely to book appointments (estimates of coefficients not individually listed in Table 2). To summarize, the likelihood of making appointments therefore seems to be dependent on all three types of characteristics: interviewer, household, and call characteristics.

A number of variables indicating the tailoring ability of the interviewer are explored. However, we did not find much conclusive evidence. The variables were either only marginally significant (e.g., if interviewer finds it difficult to modify their approach depending on the respondent) or not significant (e.g., if interviewers think they can vary their approach). For one variable, the significant coefficient was in the opposite direction to that anticipated (interviewers who disagree with the statement that they can vary their approach from situation to situation are less likely to get a refusal). Once we included important controls, such as interviewer experience, these effects were not significant any more in the final model. Such tailoring indicators are explored further in interactions with the householder on the doorstep (see subsequently). Similarly, interviewing strategies, such as altering the introduction depending on the household they visit or how best to introduce themselves and the survey, are explored but again, once variables such as interviewer experience and qualifications are controlled for, these variables are not significant any more. A potential problem with most of these variables is that they report what the interviewer generally does and reflect

the interviewer's own perception of their behavior, but do not necessarily indicate what actually happens at each call (see also the argument in Groves and Couper 1998). One variable, however, is recorded at the call level (in addition to the self-report from the interviewer on what they generally do), which allows testing of call-specific influences. Leaving a card or message behind, recorded for each call, reduces the likelihood of a postponement at the next call but otherwise does not show a significant effect (and is therefore not included in the final model).

We have already noted that interactions with the householder and tailoring ability of the interviewer have been hypothesized to play an important role in successful interviewing strategies. A number of interaction effects are investigated and two are included in the final model (Table 3). We hypothesized that more experienced interviewers may be better at responding to comments and questions. We do not find any difference in the various levels of interviewer experience regarding their ability to handle negative or positive comments. However, if there are no comments, possibly indicating a lack of engagement with the survey request and a potentially more difficult respondent, more experienced interviewers seem to be better at achieving cooperation and have a lower probability to receive a refusal than less experienced interviewers. Similarly, we find that, if the householder does not ask any questions, potentially indicating a more difficult case, then less experienced interviewers have lower cooperation, higher refusal, and higher postponement rates (effect not included in final model). If the householder asks at least one question, then both experienced and less experienced interviewers seem to perform equally well. Experience did not significantly interact with the way the contact was made on the doorstep or with demographic characteristics of the person on the doorstep, for example with age and gender. Interactions between the level of confidence and questions asked or comments made were either not significant or not easily interpretable. If a householder does not ask any questions, indicating a potentially more difficult case, then interviewers that can use a wide variety of approaches have higher cooperation, slightly lower refusal, and lower postponement rates. This may support the hypothesis that interviewers that are more able to tailor their approach to the respondent may be more successful (Groves and Couper 1998). Householder-interviewer interactions on sociodemographic characteristics (such as age and gender of interviewer with age and gender of person on the doorstep) were all not found to be significant. This is in line with previous results that also did not find much support for sociodemographic interaction effects (Durrant et al. 2010; O'Muirheartaigh and Campanelli 1999).

**Table 3.** Predicted Probabilities of Each Outcome (in Percentage) for Cross-level Interviewer Interactions (Model 3).

		Interaction between Interviewer Experience and Comment Made					
		Comment	Cooperation	Refusal	Appointment Made	Other Postponement	
Interviewer experience	Up to 2 years	Negative	23.69	32.18	16.60	27.53	
		Positive/Neutral	67.01	0.46	26.74	5.79	
		No	65.58	1.39	16.13	16.90	
	3+ years	Negative	23.95	31.01	15.99	29.05	
		Positive/Neutral	65.66	0.52	27.32	6.50	
		No	71.18	0.83	15.74	12.25	
Interaction between Can Use Wide Variety of Approaches and Question Asked							
		Outcome at Current Call					
		Question Asked	Cooperation	Refusal	Appointment Made	Other Postponement	
Can use wide variety of approaches	Strongly agree/agree	Yes	69.43	0.87	22.18	7.52	
		No	57.91	2.35	23.56	16.18	
	Neither agree nor disagree	Yes	62.75	1.02	25.07	11.16	
		No	58.14	2.42	21.19	18.25	
	Strongly disagree/disagree	Yes	68.16	0.63	22.68	8.53	
		No	56.71	2.58	21.89	18.82	

## Summary and Implications for Survey Practice

This article presents analysis of interviewer effects in the nonresponse process using interviewer call record data. Of particular interest are interaction effects between householders and interviewers on the doorstep, influences of call characteristics, tailoring strategies of interviewers, and the effects of interviewer characteristics. The aim is to better understand the process leading to cooperation or refusal, analyzing response outcome at each call, rather than focusing on predicting final response. The main findings are as follows:

1. Call characteristics are important predictors when analyzing the response outcome of each call. Characteristics of the interaction process between the interviewer and the householder and information about the initial reaction of the householder on the doorstep measured at the call-level are of particular relevance, including how contact was established, sociodemographic interviewer observations of the person at the door, and whether this person asked questions or made comments.
2. We find that the more contact calls made, the higher the odds of cooperation. This may provide some evidence that keeping in contact with the household may increase the chances of a successful interview. The finding could support the hypothesis expressed in Groves and Couper (1996, 1998) that maintaining the interaction with the household is more likely to lead to cooperation. Rather than pressing for an immediate cooperation, the interviewer may be advised to keep the conversation and the contact with the household going, for example by making an appointment for another time.
3. Unmeasured interviewer characteristics have a significant effect on nonparticipation outcomes, in line with previous research that investigated final response outcome. Interestingly, no evidence for differential effects due to unmeasured interviewer characteristics on the three nonparticipation outcomes are found, that is, the influence of the interviewer random effect is the same across the three nonparticipation outcomes.
4. A number of interviewer characteristics are found to have significant effects on the process leading to cooperation or refusal. In line with previous research on final response outcome, the attitude of the interviewer toward refusal conversion and the interviewer's self-confidence play an important role. The length of interviewer experience is significantly negatively associated with refusal on the

doorstep (although, using observational data, we cannot disentangle whether interviewer experience is cause or outcome of refusal probability). Interestingly, interviewers with lower levels of experience also show higher appointment and postponement rates. We find some evidence that more experienced interviewers handle more difficult cases better. Analyzing main interviewer effects, we do not find conclusive evidence that interviewers who indicate to tailor their approach to be more effective. We do not find much support for differential interviewer effects on the three nonparticipation outcomes. For example, interviewer's experience, qualification, and confidence impact all three nonparticipation outcomes.

5. The likelihood of making an appointment seems to be dependent on all three types of influences: interviewer and household characteristics and the circumstances of the call. For example, householders who are female, younger than 60 years of age (in particular if younger than 16), live in a house or have preschool children are more likely to make an appointment. If the call is made in the evening, the probability of appointment is significantly higher than for a call during daytime.

A potential limitation of the data is that, although the majority of characteristics are recorded at the call level, some information on specific interviewing strategies only reflect what an interviewer does in general. More information at the call level may therefore be beneficial to identify general trends on interviewer tailoring abilities. Another potential limitation is that the data are not obtained via a controlled experiment but reflect observational data and statements about causal effects may be limited. Response patterns may change over time and although our models include a wide variety of variables to control for these effects, further research will be required to analyze trends over time, for example using call record data from a longitudinal study.

The findings exhibit various implications for survey practice. Such models and the variables identified here as important may be used in responsive survey designs (Groves and Heeringa 2006; Kirgis and Lepkowski 2013; Laflamme et al. 2008), 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. The overall aim is *not* nonresponse adjustment for postsurvey use and estimation, but to inform the survey process *during* data collection, for example to reduce costs and to improve the quality of the resulting survey data. Specific recommendations for survey practice and responsive survey designs include the following:

Time-varying call record information, such as features of the call history and of the current call, play a key role in predicting the outcome of each call. The routine collection of such variables would therefore be beneficial to survey agencies. An increased number of initial or intermediate noncontacts and certain comments and questions from a householder already indicate a reduced likelihood to respond at a future call. Such signs help interviewers and survey agencies to flag more difficult cases early on and to inform intervention schemes that survey agencies can employ before the end of the data collection period to reduce final nonresponse rates. The findings may help survey agencies to determine how best to approach a household at the next call. The survey agency can then respond to such early clues by changing the contacting strategy, by offering a higher incentive, and sending a more targeted invitation letter or a more experienced interviewer. Survey agencies may use the information to inform when best to call to achieve cooperation, in particular if no prior appointment has been made.

The models also inform improvements to interviewer calling strategies, interviewer training, interviewer selection, and evaluation of interviewer performance. Interviewers may be trained to pick up important clues about the characteristics of the household or the future response behavior early on and to feed these back to the field management via an automated system particularly useful for responsive survey designs. More recently, some survey agencies have started to ask the interviewer to evaluate a household's willingness to respond which has been shown to be a good indicator for future response behavior (Copas and Farewell 1998; Eckman 2011; Wagner and Guyer 2005). The research also identifies areas where interviewers may be better trained in responding to initial reactions of the householder on the doorstep. Interviewer's experience and confidence of the interviewer were found to play a key role with more experienced and more confident interviewers showing higher likelihoods to achieve cooperation. Less experienced and less confident interviewers showed significantly higher appointment and postponement rates.

Interviewer observations such as characteristics of the house and neighborhood have been shown in previous studies to be useful in predicting response (Durrant et al. 2011; West and Kreuter 2011). Here, we find that call characteristics, observations of the initial interaction and characteristics of the householder play a key role. Survey agencies may therefore collect such information routinely on their surveys.

## **Acknowledgment**

We thank two anonymous referees for their very helpful comments.

## Authors' Note

This work contains statistical data from Office for National Statistics (ONS) which is Crown copyright and reproduced with the permission of the controller of Her Majesty's Stationery Office (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 data sets that may not exactly reproduce National Statistics aggregates.

## Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research was funded by the U.K. Economic and Social Research Council (ESRC), "The Use of Paradata in Cross-Sectional and Longitudinal Research," grant number: RES-062-23-2997.

## References

- Baruch, Yehuda and Brooks C. Holtom. 2008. "Survey Response Rate Levels and Trends in Organizational Research." *Human Relations* 61:1139-60.
- Bates, Nancy, James Dahlhamer, and Eleanor Singer. 2008. "Privacy Concerns, Too Busy, or Just Not Interested: Using Doorstep Concerns to Predict Survey Nonresponse." *Journal of Official Statistics* 24:591-612.
- Blom, Annelies G., Edith D. de Leeuw, and Joop J. Hox. 2011. "Interviewer Effects on Nonresponse in the European Social Survey." *Journal of Official Statistics* 27: 359-77.
- Blom, Annelies G., Peter Lynn, and Annette Jaeckle. 2008. "Understanding Cross-national Differences in Unit Non-response: The Role of Contact Data." ISER Working Article Series, Working Article No. 2008-01, Institute for Social and Economic Research, University of Essex, Colchester, UK.
- Clemente, Frank and Michael B. Kleiman. 1977. "Fear of Crime in the United States: A Multivariate Analysis." *Social Forces* 56:519-31.
- Conrad, Frederick, B. Jessica, B. Jose, K. Frauke, M. G. Robert, V. David, and Colleen McClain. 2013. "Interviewer Speech and the Success of Survey Invitations." *Journal of the Royal Statistical Society: Series A* 176:191-210.
- Copas, Andrew J. and Vern T. Farewell. 1998. "Dealing with Non-ignorable Non-Response by Using an 'Enthusiasm-to-respond' Variable." *Journal of the Royal Statistical Society, Series A* 161:385-96.

- Couper, Mick P. 1998. "Measuring Survey Quality in a CASIC Environment." *Proceedings of the Survey Research Methods Section, American Statistical Association* 48:743-72.
- Couper, Mick P. and Robert M. Groves. 1992. "The Role of the Interviewer in Survey Participation." *Survey Methodology* 18:263-78.
- Couper, Mick P. and Robert M. Groves. 2002. "Introductory Interactions in Telephone Surveys and Nonresponse." Pp. 161-78 in *Standardization and Tacit Knowledge: Interaction and Practice in the Survey Interview*, edited by Douglas W. Maynard, Hanneke Houtkoop-Steenstra, Nora C. Schaeffer, and Johannes van der Zouwen. New York: Wiley.
- de Heer, Wim. 1999. "International Response Trends: Results of an International Survey." *Journal of Official Statistics* 15:129-42.
- de Leeuw, Edith and Wim de Heer. 2002. "Trends in Household Survey Nonresponse: A Longitudinal and International Comparison." Pp. 41-54 in *Survey Nonresponse*, edited by R. M. Groves, D. A. Dillman, J. L. Eltinge, and R. J. A. Little. New York: Wiley.
- Durbin, J. and A. Stuart. 1951. "Differences in Response Rates of Experienced and Inexperienced Interviewers." *Journal of the Royal Statistical Society Series A* 114:163-206.
- Durrant, Gabriele B., Julia D'Arrigo, and Fiona Steele. 2011. "Using Field Process Data to Predict Best Times of Contact Conditioning on Household and Interviewer Influences." *Journal of the Royal Statistical Society Series A* 174:1029-49.
- Durrant, Gabriele B., Julia D'Arrigo, and Fiona Steele. 2013. "Analysing Interviewer Call Record Data by Using a Multilevel Discrete-Time Event History Modelling Approach." *Journal of the Royal Statistical Society Series A* 176: 251-69.
- Durrant, Gabriele B., Robert M. Groves, Laura Staetsky, and Fiona Steele. 2010. "Effects of Interviewer Attitudes and Behaviours on Refusal in Household Surveys." *Public Opinion Quarterly* 74:1-36.
- Eckman, Stephanie. 2011. "Options for Paradata in CATI Surveys: Exploring Willingness Rating." Paper presented at the International Workshop on Moving Responsive Design Forward, 3-4 November, Nuremberg, Germany.
- 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, UK, 1-18.
- Goldstein, Harvey. 2011. *Multilevel Statistical Models 4th ed.* Chichester, England: Wiley.
- Goyder, John. 1987. *The Silent Minority: Nonrespondents on Sample Surveys.* Boulder, CO: Westview Press.

- Groves, Robert M., Robert B. Cialdini, and Mick P. Couper. 1992. "Understanding the Decision to Participate in a Survey." *Public Opinion Quarterly* 56: 475-95.
- Groves, Robert M. and Mick P. Couper. 1996. "Contact-level Influences on Cooperation in Face-to-face Surveys." *Journal of Official Statistics* 12:63-83.
- Groves, Robert M. and Mick P. Couper. 1998. *Nonresponse in Household Interview Surveys*. New York: Wiley.
- Groves, Robert M. and Steven G. Heeringa. 2006. "Responsive Design for Household Surveys: Tools for Actively Controlling Survey Errors and Costs." *Journal of the Royal Statistical Society Series A* 169:439-59.
- Groves, Robert M. and Katherine A. McGonagle. 2001. "A Theory-guided Interviewer Training Protocol Regarding Survey Participation." *Journal of Official Statistics* 17:249-66.
- 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* 19:112-21.
- Hox, Joop and Edith de Leeuw. 2002. "The Influence of Interviewers' Attitude and Behavior on Household Survey Nonresponse: An International Comparison." Pp. 103-19 in *Survey Nonresponse*, edited by R. M. Groves, D. A. Dillman, J. L. Eltinge, and R. J. A. Little. New York: Wiley.
- Jaecle, Annette, Peter Lynn, Jennifer Sinibaldi, and Sarah Tipping. 2009. "Interviewer Characteristics, their Doorstep Behaviour, and Survey co-operation." ISER Working Paper Series, Institute for Social and Economic Research, University of Essex, Colchester, UK.
- Kirgis, N. G. and J. M. Lepkowski. 2013. "Design and Management Strategies for Paradata-driven Responsive Design: Illustrations from the 2006-2010 National Survey of Family Growth." Pp. 123-144 in *Improving Surveys with Process and Paradata*, edited by F. Kreuter. New York: Wiley.
- Kulka, Richard A. and Michael F. Weeks. 1988. "Towards the Development of Optimal Calling Protocols for Telephone Surveys: A Conditional Probabilities Approach." *Journal of Official Statistics* 4:319-58.
- Laflamme, Francois, Mike Maydan, and Andrew Miller. 2008. "Using Paradata to Actively Manage Data Collection Survey Process." *Proceedings of the Survey Research Methods Section, American Statistical Association* 630-37. Accessed February 7, 2014. <https://www.amstat.org/sections/srms/Proceedings/y2008/Files/300608.pdf>.
- Lillard, Lee A. and Constantijn W. A. Panis. 2003. *aML Multilevel Multiprocess Statistical Software Version 2.0*. Los Angeles, CA: EconWare.
- Lynn, Peter and Paul Clarke. 2002. "Separating Refusal Bias and Non-contact Bias: Evidence from UK National Surveys." *Statistician* 51:319-33.

- Morton-Williams, Jean. 1993. *Interviewer Approaches*. Aldershot, England: Dartmouth.
- Nicoletti, Cheti and Franco Peracchi. 2005. "Survey Nonresponse and Survey Characteristics: Microlevel Evidence from the European Community Household Panel." *Journal of the Royal Statistical Society A (Statistics in Society)* 168:763-81.
- 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* 162:437-46.
- Pickery, Jan and Geert Loosveldt. 2000. "Modeling Interviewer Effects in Panel Surveys: An Application." *Survey Methodology* 26:189-98.
- Pickery, Jan and Geert Loosveldt. 2002. "A Multilevel Multinomial Analysis of Interviewer Effects on Various Components of Unit Nonresponse." *Quality and Quantity* 36:427-37.
- 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* 20:77-89.
- Pickery, Jan, Geert Loosveldt, and Ann Carton. 2001. "The Effects of Interviewers and Respondents Characteristics on Response Behavior in Panel Surveys." *Sociological Methods and Research* 29:509-23.
- Purdon, Susan, Pamela Campanelli, and Patrick Sturgis. 1999. "Interviewer's Calling Strategies on Face-to-face Interview Surveys." *Journal of Official Statistics* 15: 199-216.
- Rasbash, Jon, Fiona Steele, William J. Browne, and Harvey Goldstein. 2009. *A User's Guide to MLwiN v2.10*. Bristol, UK: Centre for Multilevel Modelling, University of Bristol.
- Sangster, Roberta L. and Brian J. Meekins. 2004. "Modeling the Likelihood of Interviews and Refusals: Using Call History Data to Improve Efficiency of Effort in a National RDD Survey." *Proceedings of the American Statistical Association, Section of Survey Research Methods*, 4311-17. Accessed February 7, 2014. <http://www.amstat.org/sections/srms/Proceedings/y2004/files/Jsm2004-000520.pdf>.
- Sinibaldi, Jennifer, Gabriele B. Durrant, and Frauke Kreuter. 2013. "Evaluating the Measurement Error of Interviewer Observed Paradata." *Public Opinion Quarterly* 77:173-93. Special issue: Topics in Survey Measurement and Public Opinion.
- Snijders, Tom and Roel Bosker. 1999. *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modelling*. London, UK: Sage publications.
- Steele, Fiona, Ian Diamond, and Duolao Wang. 1996. "The Determinants of the Duration of Contraceptive Use in China: A Multilevel Multinomial Discrete-hazards Modeling Approach." *Demography* 33:12-23.
- Steele, Fiona, Wendy Sigle-Rushton, and Oystein Kravdal. 2009. "Consequences of Family Disruption on Children's Educational Outcomes in Norway." *Demography* 46:553-74.

- 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* 40:121-39.
- Wagner, James. 2013. "Using Paradata-driven Models to Improve Contact Rates in Telephone and Face-to-face Surveys." Pp. 145-170 in *Improving Surveys with Process and Paradata*, edited by F. Kreuter. New York: Wiley.
- Wagner, James and H. Guyer. 2005. "Statistical Propensity Models to Predict Likelihood of Survey Response Compared to Interviewer Judgments of Likelihood of Response." Paper presented at the *Annual Conference of the Midwestern Association for Public Opinion Research*, Chicago, IL.
- Weeks, Michael F., Bruce L. Jones, Ralph E. Folsom, and Charles H. Benrud. 1980. "Optimal Times to Contact Sample Households." *Public Opinion Quarterly* 44: 101-14.
- West, Brady T. 2013. "An Examination of the Quality and Utility of Interviewer Observations in the National Survey of Family Growth." *Journal of the Royal Statistical Society* 176:211-25. Special issue on "The use of paradata in social survey research."
- West, Brady T. and Frauke Kreuter. 2011. "Observational Strategies Associated with Increased Accuracy of Interviewer Observations: Evidence from the National Survey of Family Growth." In *Proceedings of the Survey Research Methods Section, American Statistical Association*. Accessed February 7, 2014. [http://www-personal.umich.edu/~bwest/just\\_jsm\\_proc\\_2011.pdf](http://www-personal.umich.edu/~bwest/just_jsm_proc_2011.pdf).

## Author Biographies

**Gabriele Durrant** is Associate Professor in Social Statistics, Department of Social Statistics and Southampton Statistical Sciences Research Institute at the University of Southampton, United Kingdom. She is also director of CASS (Courses in Applied Social Surveys) at the University of Southampton. She is the principal investigator of the ESRC-funded research project "The Use of Paradata in Cross-Sectional and Longitudinal Research," grant number: RES-062-23-2997.

**Julia D'Arrigo** was a research fellow at the Southampton Statistical Sciences Research Institute at the University of Southampton. She is now working in the insurance industry.