ESRC Project: Hierarchical Analysis of Unit-Nonresponse in Sample Surveys
Aims and Objectives and Progress during First Year

Overview	
Project title:	**Hierarchical Analysis of Unit Nonresponse in Sample Surveys**
Award No:	**RES-062-23-0458**
Award Holders:	**Dr Gabriele Durrant (PI), Prof John Micklewright, Prof Chris Skinner, Prof Fiona Steele**
Institutions:	University of Southampton and University of Bristol
Award start date:	01/09/2007
Award end date:	31/08/2010
Research fellows:	Julia D'Arrigo, Dr Sylke Schnepf (part-time), Laura Staetsky (part-time)
External collaborators:	Prof Jelke Bethlehem, Fannie Cobben, Prof Robert Groves, Jean Martin

Aims and Objectives

The broad aims of the project will be to improve understanding of nonresponse patterns and interviewer effects, hence informing efforts of survey practitioners to increase response rates and reduce nonresponse bias, and to advance methodology for adjusting for nonresponse. The specific aims of the project are:

1. to develop models for nonresponse in multistage surveys, allowing for correlation in response behaviour among survey units (e.g. sample members allocated to the same interviewer).
2. to analyse interviewer effects on the probabilities of noncontact and refusal.
3. to develop response propensity models on process data to predict the likelihood of contact at each call.
4. to develop post-survey adjustment methods for nonresponse with hierarchical survey data.
5. to provide guidance for social scientists and survey practitioners on the use of methods to reduce nonresponse bias and to increase response rates.
6. to disseminate the results of this work through publications, conference presentations, a short course, and an international research symposium.

The proposed research project is organised into four subprojects (1) – (4). A unifying theme of the proposed research is a focus on response within a hierarchical context, for example individual response within households or the responses of employees within firms. One key topic is the response behaviour of sampled units within interviewer assignments and the dependence of this behaviour on the interaction between units and interviewers. Another emphasis is multistage surveys, where the nonresponse process may itself be hierarchical; for example, in surveys of individuals within households or pupils within schools, nonresponse may occur at any of the nested levels. This work makes use of the UK Census Link Study, which comprises six major UK household surveys. This is a unique data source on household survey nonresponse in a European

context. We will also analyse the England component of the PISA study of schools and pupils, which provides an example of another type of multistage survey with a significant problem of nonresponse.

1. Modelling response patterns in multistage surveys

This subproject will research the specification of models which will provide a basis for the three other subprojects. An important aim is the investigation of the influence of the social context on individual action. We shall develop models to investigate the contextual nonresponse behaviour of both individuals within households and pupils within schools. The focus will be on surveys with multistage response patterns. We will also allow for different breakdowns of types of nonresponse, including proxy response behaviour, which is currently not well understood. The models will be built on current conceptual frameworks for survey participation (e.g. Groves et al 2000). To provide a general framework for representing the context-dependent behaviour of interest, we will use multilevel models for binary and multinomial responses.

2. Analysis of interviewer effects and interactions with sample units

Despite new modes of data collection, face-to-face interviewing is still one of the most important ways of collecting survey data. Interviewers have a vital role in contacting sample members and achieving their cooperation, and can be seen as a hierarchical component in the response process. A better understanding of interviewer effects is critical for the improvement of interviewer training and for informing strategies to maximise response. In particular, little is yet known about how best to match interviewers to sample members and how best to tailor interviewing techniques to certain individuals or subgroups to achieve higher response rates and to reduce nonresponse bias. This subproject analyses interviewer effects and their interactions with individual and household characteristics, including the investigation of survey specific and survey independent effects.

3. Hierarchical response propensity models based on process data

Establishing contact is an important part of the response process and effective interviewer calling patterns are critical in achieving contact and subsequent possible cooperation. As identified in Groves and Heeringa (2006), research is needed to establish how best to use such data to inform nonresponse processes, as well as further methodological development in the specification of response propensity models. So far, this area is not well researched and little is known about the benefits of process data for data collection, adjustment and analysis. The Census Link data offer unique possibilities on this score.

This subproject has two principal aims: a) to build and improve response propensity models based on process data to predict the likelihood of contact at each call, addressing the hierarchical structure of the data available, and b) to inform the design of effective and efficient interviewer calling strategies, call-backs and follow-ups to reduce nonresponse bias and to increase response rates. We will use multilevel discrete-time event history analysis (Steele et al 2004) to model the propensity of contact or interview at each call. The models will condition on all information available to the interviewer, such as from administrative data or prior calls, and will include call record data as time-varying covariates. It is anticipated that this work will inform improvements in the design of call-backs and follow-ups of nonrespondents. The work will also inform responsive survey designs, as defined by Groves and Heeringa (2006).

4. Weighting methods to adjust for nonresponse in multistage surveys

This subproject will explore the potential applications of hierarchical models of the nonresponse process to post-survey weighting adjustment methods. In particular, there appears to be little research on the potential to incorporate information about interviewers in weight construction and any potential problems that might arise. This subproject will develop a set of weighting methods for nonresponse adjustment in the kinds of surveys considered in subproject 1. It will evaluate these methods both theoretically and numerically, employing both the PISA and Census Link datasets. Particular consideration will be given to the potential role of interviewer information and the possible dependence of the properties of estimators on the choice and assignment of interviewers.

Progress during first year

During the first year of the project, we have focused on aims (1), (2) and (6) and have made good progress on all three. We have also recently started work on preparing the data and carried out some initial analysis for (3). Work on (5) has started with the delivery of two short courses, and two more in planning.

Publications

Durrant, G.B. and Steele, F. (2009) Multilevel modelling of refusal and noncontact in household surveys: evidence from six UK government surveys. *Journal of the Royal Statistical Society*, *Series A,* 172, 2, pp. 1-21 (to appear).

Durrant, G.B., Staetsky L., Groves, B. and Steele, F. (2008) Effects of interviewer attitudes and behaviours on refusal in household survey. In preparation; to be submitted to *Public Opinion Quarterly*.

Conference and workshop presentations

September 2007. Two presentations given at the 18th International Workshop on Household Survey Nonresponse in Southampton: 'Using multilevel multinomial models to analyse household and interviewer effects in survey nonresponse' (Durrant) and 'Analysing survey data where measurement and response behaviour is selective' (Skinner).

October 2007. 'Estimation of a distribution function under nonresponse using propensity score weighting' (Skinner), presentation at a workshop on Calibration and Estimation in Surveys, Statistics Canada, Ottawa.

December 2007. 'Estimation of a distribution function from survey data with nonresponse' (Skinner), seminar at the Institute of Statistics, University of Neuchatel, Switzerland.

July 2008. 'An exploration of the role of interviewers on survey nonresponse using a multilevel modelling approach' (Durrant), invited paper in a session on Survey Methods, NCRM Methods Festival, Oxford.

July 2008. 'How big was response bias in England to PISA 2003' (Mickewright), paper in a session on Administrative Data, NCRM Research Methods Festival, Oxford.

September 2008. 'An investigation of interviewer effects on household nonresponse in six UK government surveys' (Durrant), 7th International Conference on Social Science Methodology, Naples.

September 2008. Three presentations on research conducted under the project at the 19th International Workshop on Household Survey Nonresponse, in Slovenia: 'Using Paradata to Develop Hierarchical Response Propensity Models' (D'Arrigo and Durrant), 'The Effects of Interviewer Characteristics and Attitudes on Refusal in Face-to-Face Surveys' (Durrant) and 'Response to income questions in single question surveys' (Micklewright).

Papers in progress

Steele, F. and Durrant, G. (2008/2009): Alternative Approaches to Nonresponse Modelling (work in progress).

Durrant, G.B. and D'Arrigo, J. (2008): Using Paradata to Develop Hierarchical Response Propensity Models (work in progress). Draft paper available.

Micklewright, J., Schnepf, S. and Skinner, C. (2008) Analysis of Nonresponse Bias in Pisa Study Data (draft paper available).

Dissemination (with Users and the Academic Community)

Although in the early stages of the project, the research has been presented at a number of conferences and workshops, including invited sessions. In addition, we have discussed research ideas and findings with our external advisors and collaborators (including Robert Groves from the University of Michigan, Jean Martin, Ray Chambers from the University of Wollongong, Jelke Bethlehem and Fannie Cobben from Statistics Netherlands and members of staff from the Office for National Statistics). We also had meetings with colleagues working on other ESRC initiatives and related projects, notably the Survey Design and Measurement Initiative (including Peter Lynn, Gerry Nicolaas, John Bynner and Harvey Goldstein).

Short courses

'Nonresponse rates and nonresponse bias', Robert Groves from the University of Michigan, June 2008. (2-day course)

'Handling nonresponse in sample surveys', with collaborators from Statistics Netherlands (CBS) in March 2008. (2-day course)

'Missing data methods: imputation and weighting', main presenter James Carpenter (LSHTM), in Nov 2008. (3 day course)

'Handling nonresponse in sample surveys', Jelke Bethlehem, Fannie Cobben and Gabriele Durrant, in June 2009. (3-day course)

Outlook:

- Possibility of a research workshop/seminar at the RSS, organised by the Social Statistics Section.