

# Analysis of Household Survey Non-Response using a Multilevel Modelling Approach

Gabriele B. Durrant

Southampton Statistical Sciences Research Institute

25 April 2006 – Q2006



# Outline

- Introduction
- Data
- Methodology
- Refusal and Noncontact
- Interviewer Effects
- Summary
- Further Research



# Aims of the Study

- Gain deeper understanding of nature and causes of non-response
- Identification of factors influencing **household unit non-response** in government surveys – survey dependent and survey independent factors
- Identification of **interviewer** and area **effects**
- Understanding of interaction between household and interviewer
- ...



# The Data

- Project carried out in close collaboration with ONS
- **ONS Survey Nonresponse Census Link Study**
- Allows for comparisons of surveys with different designs and subject matters
- **The surveys are:**
  - General Household Survey (GHS)
  - Expenditure and Food Survey (EFS)
  - Family Resources Survey (FRS)
  - Omnibus Survey (OMN)
  - Labour Force Survey (LFS)
  - National Travel Survey (NTS)



# Main Advantages of the Data

- several major government surveys included in 1 study
- information available for respondents and non-respondents based on census data
- rich set of auxiliary variables
- comprising individual and household level census variables
- information obtained from interviewers
- area information
- ...



# Survey Nonresponse Census Link Data

Survey Respondents

R=1

Individ.  
Census

Household  
Census

Inter-  
viewer  
obser-  
vation

Inter-  
viewer  
attitude

Area

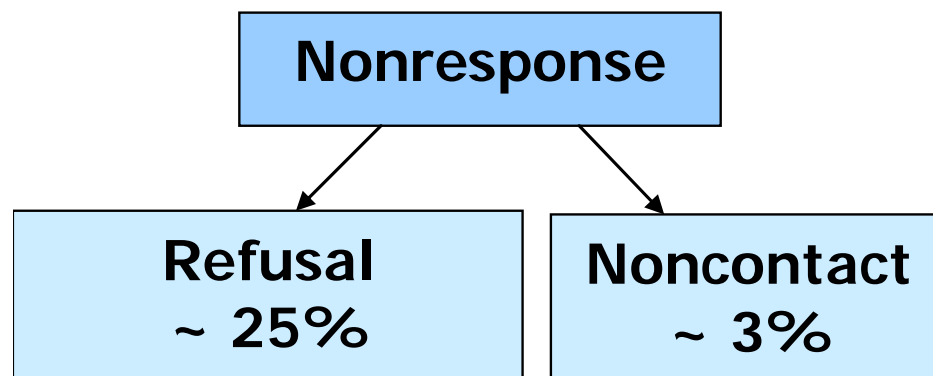
Survey Nonrespondents

R=0



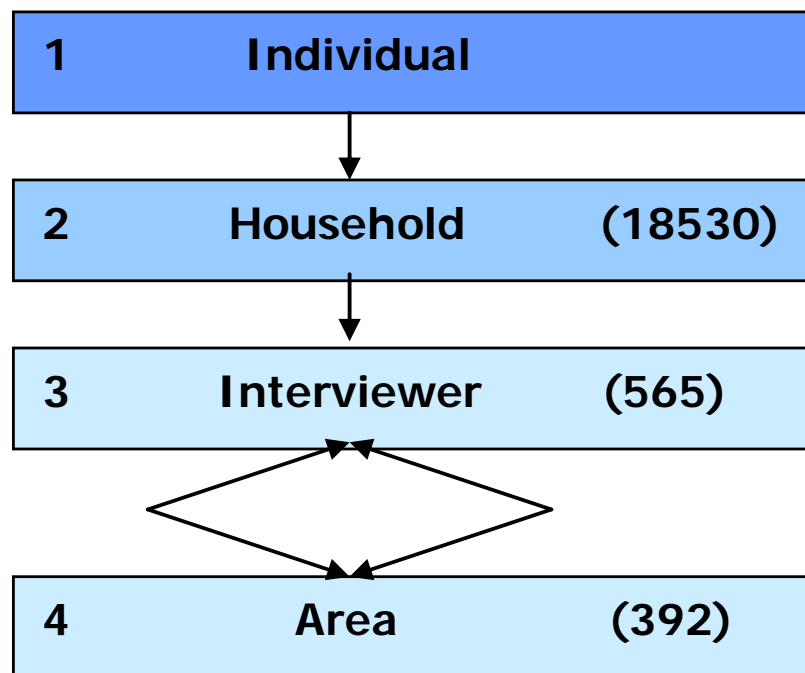
# Types of Nonresponse

- Focus on **refusal** and **non-contact** as the two main components of nonresponse



- Here: **household (unit-) nonresponse** only

# Multilevel Structure of the Data



- Here for modelling we only consider levels 2-4
- Note: this is not an interpenetrated sampling design

→ **Multinomial multilevel models:**

$$\log \left( \frac{\pi_{ij}^{(1)}}{\pi_{ij}^{(0)}} \right) = \beta_0^{(1)} + \sum_{p=1}^P \beta_p^{(1)} x_{pij} + u_j^{(1)}$$
$$y_{ij} = \begin{cases} 0 & \text{cooperation} \\ 1 & \text{refusal} \\ 2 & \text{noncontact} \end{cases}$$

$$\log \left( \frac{\pi_{ij}^{(2)}}{\pi_{ij}^{(0)}} \right) = \beta_0^{(2)} + \sum_{p=1}^P \beta_p^{(2)} x_{pij} + u_j^{(2)}$$
$$\Omega_u = \begin{bmatrix} \sigma_u^{2(1)} & \\ \sigma_u^{(1,2)} & \sigma_u^{2(2)} \end{bmatrix}$$

- Analysis of interviewer and area effects
- Advantage: Estimation of covariance between random term for refusals and non-contacts (eg for interviewers)



# Methodology (cont.)

- **Cross-classified logistic and multinomial models**  
(interviewers ↔ areas)

$$\log \left( \frac{\pi_{i(jk)}^{(1)}}{\pi_{i(jk)}^{(0)}} \right) = \beta_0^{(1)} + \sum_{p=1}^P \beta_p^{(1)} x_{pi(jk)} + u_j^{(1)} + v_k^{(1)} \quad (\sigma_u^{2(1)}, \sigma_v^{2(1)})$$

$$\log \left( \frac{\pi_{ij}^{(2)}}{\pi_{ij}^{(0)}} \right) = \beta_0^{(2)} + \sum_{p=1}^P \beta_p^{(2)} x_{pi(jk)} + u_j^{(2)} + v_k^{(2)} \quad (\sigma_u^{2(2)}, \sigma_v^{2(2)})$$

- Modelling interviewer and area effects simultaneously with the aim of separating interviewer and area effects



# Results: Household level variables

- Both refusal and noncontact systematically related to demographic and/or attitude variables
- Survey indicator highly significant (to account for design and response rate differences across the 6 surveys); survey dependent and independent effects (interactions)
- Noncontact 'easier' to explain than refusal
- Refusal and noncontact are two different processes
- Both components require different treatment



# HH level results: Refusal

- characteristics of hh representative generally more significant than hh-level variables
- **attitude variables** important rather than purely 'factual' variables (e.g. perception on health)
- strongly related to qualifications and socio-economic status
  
- relevance for how best to approach respondents
- more emphasis on attitude variables in later studies



# HH level results: Noncontact

- **factual variables** (type of house, physical barriers etc) important
- once controlled for characteristics of house basic area variables such as urban-rural, London, etc not or almost not significant any more
- variables clearly relate to the propensity of being physically at home (e.g. pensioner, single hh,...)
- some opposite effects to refusal (e.g. people that move more frequently, no adult in employment)



# Interviewer Effects



# Analysis of Interviewer Effects

- generally recognised that interviewers can have an effect on survey response ('interviewer variance')
- both refusal and noncontact are subject to interviewer effects (significant interviewer variance)
- interviewer-level variables can explain part of interviewer-level variation
- implications for survey practice, e.g. interviewer training, recruitment, evaluation

<b>Interviewer level variables</b> (0 = reference category)	<b>Categories</b>	$\hat{\beta}$ ( <i>ste</i> ( $\hat{\beta}$ ))
Years of experience (0 Less than 1 year)	1 1 to 2 years	-0.042 (0.075)
	2 3 to 8 years	0.054 (0.095)
	3 9 years or more	0.319 (0.112)*
Pay grade (0 Interviewer)	1 Advanced Interviewer	-0.004 (0.096)
	2 Merit 1	-0.085 (0.089)
	3 Merit 2	-0.211 (0.100)*
	4 Merit 3	-0.449 (0.103)*
	5 Field Manager	-1.227 (0.805)
	6 Missing	-0.045 (0.560)
Daily hours worked previous year at weekends (0 0 Hours)	1 1 to 4 hours inclusive	-0.079 (0.051)
	2 5 to 10 hours inclusive	-0.158 (0.077)*
	3 Missing	0.042 (0.109)
Can convince reluctant respondents (0 Less confident)	1 More confident	-0.234 (0.069)*
Should persuade most reluctant respondents (0 Strongly agree)	1 Agree	0.071 (0.054)
	2 Neither agree nor disagree	-0.126 (0.090)
	3 Disagree	0.113 (0.081)
	4 Strongly disagree	0.261 (0.143)*
<b>Cross-level interaction</b> (0=male)	Gender of household reference person x interviewer gender	-0.194 (0.077)*



# Interviewer Effects: Refusal

- Interviewer variance significant but small (VPC 2-3%)
- **Main effects** of basic interviewer characteristics:
  - Age, gender etc not significant
  - Pay grade significant
  - Interviewer experience (in years) just significant
  - indication that a more flexible approach of interviewers is more successful
- Importance of **interactions** between household and interviewer important (e.g. gender interaction) not just main effects

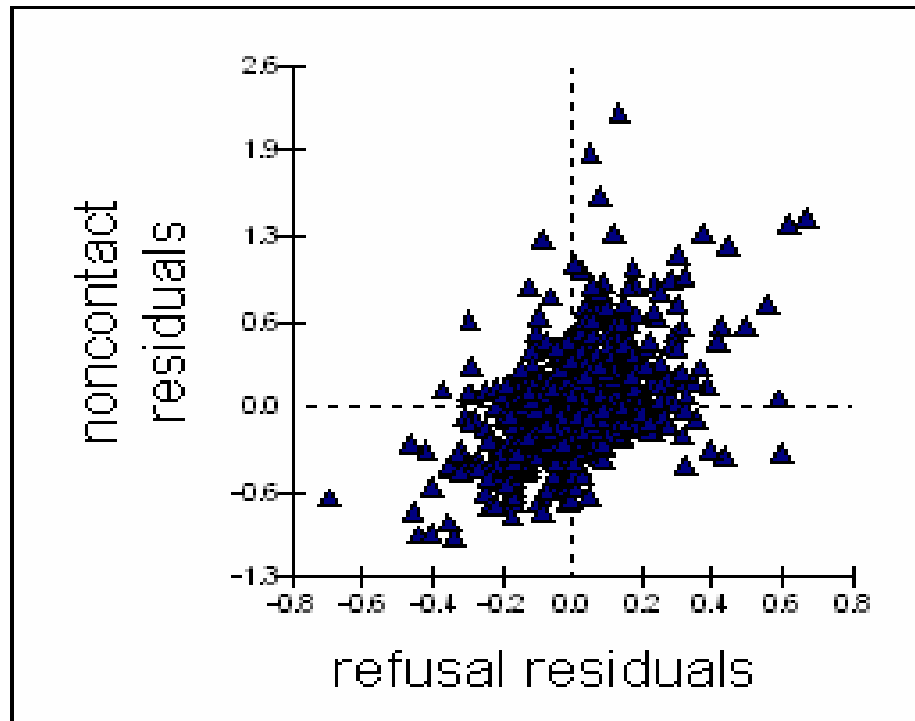


# Interviewer Effects: Refusal

- Attitudes of interviewers significant:
  - Confidence of interviewer regarding persuasion
  - Use of persuasion strategies
  - Relevance of positive interviewer expectations in gaining cooperation
- Interviewer behaviour as main effects basically not significant
  - Stresses potential importance of interaction between household and interviewer
  - Indication that tailoring of approach to household important rather than general approach

# Results from multinomial models

- Covariance term between random term for refusals and non-contacts for interviewers: positive and significant
- Means that after controlling for other factors in the model interviewers that perform well in establishing contact also perform well in gaining cooperation





# Cross-classified models

- Allows the separation of **interviewer and area effects** (variances) (fitted simultaneously)
- Results show interviewer variances about twice as large as area variances
- Hypotheses:
  - Area variables are ‘just’ proxies for household level characteristics (e.g unemployment rate; social housing,...)
  - Once characteristics of households and/or interviewers are controlled for area effects may disappear



# Summary of main findings

- Clear differences in **refusal and noncontact**
- Refusal better explained by socio-economic and attitudes variables
- Noncontact based on factual variables (propensity to be physically at home)
- **Interviewer** have a significant effect
- Important: interviewer attitudes and interaction with household, not necessarily main effects
- Socio-demographic characteristics of interviewers and main effects of behaviour less important



# Further Research

- Further investigation of interviewer effects (e.g. survey dependent and independent effects)
- Analysis of individual level nonresponse (4 levels)
- Analysis of interviewer calling patterns
- Interaction at 'doorstep'
- Investigation of non-response bias
- Investigation of performance of different adjustment methods (weighting, different weighting models, hierarchical; different types of estimators, ...)

- Potential implications for survey practice
  - How best to approach sample survey members
  - How best to target certain subgroups of the population
  - Tailoring of approaches
  - Interviewer training, recruitment and evaluation