Fiona Steele and Gabriele B. Durrant

# Alternative approaches to multilevel modelling of survey non-contact and refusal

## Article (Accepted version)
## (Refereed)

**Original citation:**
Steele, Fiona and Durrant, Gabriele B. (2011) *Alternative approaches to multilevel modelling of survey non-contact and refusal.* International statistical review, 79 (1). pp. 70-91. ISSN 0306-7734

DOI: 10.1111/j.1751-5823.2011.00133.x

# Alternative Approaches to Multilevel Modelling of Survey Noncontact and Refusal

**Fiona Steele[1] and Gabriele B. Durrant[2]**

[1]Centre for Multilevel Modelling, University of Bristol

[2]Southampton Statistical Sciences Research Institute, University of Southampton

**Summary**. We review three alternative approaches to modelling survey noncontact and refusal: multinomial, sequential and sample selection (bivariate probit) models. We then propose a multilevel extension of the sample selection model to allow for both interviewer effects and dependency between noncontact and refusal rates at the household and interviewer level. All methods are applied and compared in an analysis of household nonresponse in the UK, using a dataset with unusually rich information on both respondents and nonrespondents from six major surveys. After controlling for household characteristics, there is little evidence of residual correlation between the unobserved characteristics affecting noncontact and refusal propensities at either the household or the interviewer level. We also find that the estimated coefficients of the multinomial and sequential models are surprisingly similar, which further investigation via a simulation study suggests is due to noncontact and refusal having largely different predictors.

**Keywords**. Clustered categorical response data, discrete choice models, hierarchical models, survey nonresponse, interviewer effects.

**Résumé**. Nous passons en revue trois approches alternatives pour modéliser l'absence de contact et le refus dans les enquêtes: modélisation multinomial, séquentielles et par sélection d'échantillon (probit bivarié). Nous proposons une généralisation de type multi niveau pour la modélisation par sélection d'échantillon pour tenir compte des effets de l'enquêteur et de la dépendance entre les taux d'absence de contact et de refus au niveau des ménages et de l'enquêteur. Toutes les méthodes sont appliquées et comparées dans le cadre d'une analyse de non réponse de ménages dans le Royaume Uni, en utilisant des données de six enquêtes majeures contenants des informations riches à la foi sur les répondants et les non répondants. Après avoir effectué un contrôle sur les caractéristiques des ménages, nous constatons qu'il y a peu d'évidence d'une corrélation réelle entres les caractéristiques non observées qui affectent les probabilités de refus et d'absence de contact, à la foi au niveau des ménage et au niveau de l'enquêteur. Nous avons aussi constaté que les coefficients estimés pour les modèles multinomiaux et séquentielles sont étonnamment similaires. Après une analyse par simulations plus détaillées, il semble que cela soit du au fait qu'il y a très peu de chevauchement entre les prédicteurs d'absence de contact et de refus.

## 1. Introduction

Achieving participation of a sample unit, such as a household or individual, to a survey request has become an increasingly difficult task with nonresponse rates steadily rising over recent decades (Bethlehem et al., 2011; de Heer, 1999). Conceptually, for face-to-face and telephone surveys, the response process is commonly separated into two stages (Groves and Couper, 1998): first the survey agency needs to establish contact with the sample unit and then the sample unit needs to agree to participate in the survey. Previous research has shown that different characteristics of the sample unit may influence the two processes, which has led many survey researchers to advocate the treatment of noncontact and refusal as separate components of nonresponse (Groves and Couper, 1998; Lynn and Clarke, 2002; Lynn et al., 2002). Typical correlates of contact are proxies for the amount of time spent at home and lifestyle as well as the presence of physical impediments to the household and the timing of the call (Groves and Couper, 1998; Purdon et al., 1999). The decision to take part in a survey or to refuse, however, may be influenced by individual or household demographic and socio-economic characteristics and attitudes, the topic of the survey and the interaction between the individual and the interviewer (Groves and Couper, 1998; Durrant and Steele, 2009). The two processes may also have common correlates, which may work in the same or possibly opposite directions. Furthermore, some of these factors are likely to be unmeasured, leading to an unexplained correlation between a sample unit's ease of contact and their likelihood of participation. The potential relationship between the processes of noncontact and refusal has led to calls to consider them simultaneously, rather than focusing on only one component (Lynn et al., 2002).

Previous researchers have tended to use one of two approaches to model noncontact and refusal. The first is to define a composite noncontact-refusal outcome with three categories – (a) noncontact, (b) contact and refusal, and (c) contact and participation – and to estimate simultaneously pairwise contrasts between categories using a multinomial logit model (e.g. O'Muircheartaigh and Campanelli, 1999; Pickery and Loosveldt, 2002; Durrant and Steele, 2009).

In the second approach, the response process is viewed as the outcome of two sequential events, leading to two binary outcomes - one for noncontact and another for refusal, conditional on contact - that are typically modelled separately using binary logit or probit models (e.g. Hawkes and Plewis, 2006). One advantage of the multinomial logit model over the sequential model is that covariate effects on the probability of refusal and noncontact may be evaluated simultaneously and tested for equivalence. However, the coefficients of a multinomial model can be difficult to interpret because comparisons with category (c) combine outcomes of the contact and participation processes. The sequential model has a more intuitive appeal because it follows the two-stage process of securing cooperation and the parameters are directly interpretable in terms of each stage.

Neither the multinomial logit nor the sequential model, however, allows for the possibility that there may be residual correlation between a sample unit's ease of contact and likelihood of co-operating with the survey request, that is, dependency unexplained by the covariates in the model. This is of particular concern in view of the paucity of information that is usually available for nonrespondents. Failure to account for this correlation may lead to biased parameter estimates. Moreover, because the omitted variables may have opposing effects on noncontact and refusal, it is difficult to predict the direction of any bias. A third approach has therefore been proposed in which probit equations for noncontact and refusal are jointly estimated to allow for dependence between the two processes (Nicoletti and Peracchi, 2005).

In face-to-face surveys, it is the task of the interviewer to both establish contact and to persuade the selected sample member to take part in the survey. Previous research has therefore considered interviewer effects on the probabilities of noncontact and refusal, and clustering of nonresponse rates by interviewer due to effects of unmeasured interviewer characteristics (Groves and Couper, 1998; Pickery and Loosveldt, 2002; O'Muircheartaigh and Campanelli, 1999). Just as a sample unit's chance of being contacted might be correlated with their chance of participating in a survey, we might expect that interviewers with high contact rates will also have high participation rates. Moreover, if some of the interviewer characteristics affecting the probability of noncontact and refusal are unmeasured, there will be residual correlation between interviewer effects on the two processes.

Using a sequential model with independently estimated equations for noncontact and refusal, it is not possible to allow for cross-process correlation at the interviewer level. On the other hand, multilevel multinomial logit modelling has been used to allow for interviewer effects on the probabilities of noncontact and refusal, and it is straightforward in the multilevel framework to allow for correlation between interviewer effects on each process (O'Muircheartaigh and Campanelli, 1999; Pickery and Loosveldt, 2002; Durrant and Steele, 2009). In this paper, we propose an extension of Nicoletti and Peracchi's bivariate probit model that allows for interviewer effects on noncontact and refusal as well as residual correlation between processes both at the sample unit level and at the interviewer level. We argue that this approach overcomes the limitations of the sequential model, while retaining its convenient interpretation in terms of the two stages of securing an interview. Another aim of the paper is to provide a review of each of the three types of nonresponse models described above, emphasising links between them as well as highlighting differences in their underlying assumptions and in the interpretation of model parameters. Although examples of applications of each method can be found in the nonresponse literature, there has been no attempt to compare the different approaches and some researchers may be unaware of the alternatives and the differences between them.

We illustrate the application of multilevel versions of each method in an analysis of household and interviewer effects on noncontact and refusal using data from the UK Census Link Study. The key strengths of these data are that detailed information is available for both responding and nonresponding units across several surveys, including detailed information on interviewers. We compare the fit of the three models and, in particular, assess the evidence for residual correlation between noncontact and refusal propensities at the household and interviewer level. Finally, we provide practical guidelines on when one approach might be preferred over the others, drawing on evidence from our analysis of nonresponse in the UK and from a simulation study that explores when multinomial and sequential models yield similar parameter estimates.

While this paper focuses on approaches to modelling survey noncontact and refusal, the methods we describe are relevant to any application involving clustered multinomial response data where the response may be viewed as the outcome of one or more sequential processes. In demography, for example, we might treat partnership status (single, married or cohabiting) as the outcome of a two-stage process: (i) the decision to partner, and (ii) the decision to marry or cohabit a prospective partner. Clustering arises if we have longitudinal data on individuals (e.g. Steele et al., 2006), or if there is geographical variation in the rates of marriage and cohabitation, and residual correlation would be expected if the probabilities of entering marriage and cohabitation have shared unmeasured correlates (Hill et al., 1993).

The remainder of the paper is structured as follows. In Section 2 we describe in more detail why dependency between the noncontact and refusal processes might be expected, and the consequences of ignoring it. The alternative approaches to modelling nonresponse are reviewed in Section 3, and an extension of the bivariate probit model to allow for interviewer effects is proposed. These methods are then applied and compared in analyses of the UK Census Link Study in Section 4. This is followed in Section 5 by a simulation study to investigate the conditions under which the multinomial logit and sequential logit models are expected to have similar regression coefficients. We end in Section 6 with a summary of the main findings and their implications for survey research and practice.

## 2. Dependency between the Noncontact and Refusal Processes

While previous research suggests that the processes of noncontact and refusal are relatively distinct, there is evidence that they share predictors, with some factors affecting noncontact and refusal in the same direction and others working in opposite directions. For example, single male households are both difficult to contact and to persuade to participate (Groves and Couper, 1998), while unemployed sample members have been found to be easier to contact but to be more likely to refuse than those in full-time employment (Durrant and Steele, 2009; Durrant et al., 2011). In practice, and especially given that we usually have little if any information on nonrespondents, some of these shared influences will be unobserved leading to a residual correlation between the two processes. A cross-process correlation could also arise as a result of incorrect classification of refusals as noncontacts where individuals pretend to be not at home when an interviewer calls (Nicoletti and Peracchi, 2005). If individuals who avoid the interviewer also have a high chance of refusal, we would expect to find a positive residual correlation between the difficulty of contact and the propensity to refuse. In general, however, it is difficult to predict the direction of the residual correlation.

The few previous studies that have considered the relationship between noncontact and refusal have found little evidence of a correlation, but their findings are based on restricted samples. Lynn et al. (2002) compare aggregate noncontact and refusal propensities across a range of surveys over time, but do not have information on nonrespondents and therefore compare respondents classified according to their difficulty to contact and reluctance to participate. Nicoletti and Peracchi (2005) conduct a micro-level analysis of longitudinal data from several European countries, and find a small negative correlation, but focus on attrition after the first wave because of a lack of information on nonrespondents at wave 1. Generalisation of their results to cross-sectional surveys is problematic because, as noted by the authors and Lepkowski and Couper (2002), wave 1 unit-nonresponse differs in important ways from nonresponse at subsequent waves, and nonresponse in cross-sectional surveys is likely to be more similar to wave 1 nonresponse.

Unmeasured influences on noncontact and refusal are likely to include characteristics of both the sample unit and the interviewer, leading to residual correlations between a sample unit's propensities to be contacted and to participate and between an interviewer's contact and participation rates. For example, we might expect that more able, motivated or ambitious interviewers have a higher chance both of establishing contact and of securing participation. In the absence of good measures of interviewer characteristics that are associated with contact and participation, we would therefore expect a positive residual correlation between interviewer effects on noncontact and refusal. The existence of a correlation at the interviewer level has been speculated in the survey literature, but has been explored by only a few (O'Muircheartaigh and Campanelli, 1999; Pickery and Loosveldt, 2002; Durrant and Steele, 2009). However, we are unaware of any attempt to allow simultaneously for correlation at both the interviewer and sample unit level.

Failure to account for residual correlation between the different components of nonresponse may lead to biased parameter estimates. To illustrate the biases that may arise, suppose a person's ease of contact and chance of participation depend on factors relating to how busy they are, but that we observe only the number of hours worked by each head of household (X). If the effect of X, and that of the omitted factors, on the noncontact probability is positive, then 'busy' people will be under-represented in the contacted sample. Suppose also that busy people are more likely to refuse, then when we model the propensity of refusal conditional on contact (that is, fit the refusal equation to the contacted subsample) we would underestimate the effect of hours worked on refusal. In other words, the contacted subsample is a selected group containing households that may, on average, be less likely to refuse. One way to allow for this sample selection is to jointly model the propensity of noncontact and the *unconditional* propensity of refusal using a bivariate probit model (as in Nicoletti and Peracchi, 2005). This approach enables us to make inferences about the determinants of refusal in the whole target population, rather than in the contactable subpopulation which, in the above illustration, has an above-average participation rate.

It might be argued that survey practitioners are interested in identifying the determinants of refusal among contacted households rather than the intended population because, in practice, interviewers can only seek the participation of sample members with whom they make contact. The problem with basing inferences on the contacted subsample, however, is that this group may be population and survey specific. To the extent that the nature of sample selection varies across different target populations and surveys, for example because of differences in interviewing strategies and in fieldwork processes, restricting inferences to the population of contacts limits the external validity of a study. We argue that when the objective of nonresponse modelling is to gain an understanding of the underlying causal processes – for example, to identify determinants of participation that could potentially be manipulated by survey organisations – it is the correlates of refusal among *all* prospective respondents that is of interest.

## 3. Alternative Modelling Approaches

In this section, we describe three alternative strategies for analysing noncontact and refusal: multinomial, sequential and bivariate probit (sample selection) models. We consider multilevel models for two-level hierarchical structures where the response outcome is defined for households (at level 1) nested within interviewers (at level 2). All models can be extended to handle three-level structures. For example, household random effects may be added where there are multiple sample members per household, provided there is sufficient within-household variation in noncontact and refusal to permit identification of household effects. It is also possible to allow for non-hierarchical data structures, for instance where sample units are nested within a cross-classification of interviewers and areas (because an interviewer may work in more than one area and an area may be visited by more than one interviewer). (See Durrant et al. (2010) for an application of a cross-classified model in an analysis of refusal conditional on contact.)

Throughout the paper we denote the outcome of a successful contact by $C$, its complement noncontact by $\overline{C}$, participation by $P$ and refusal by $\overline{P}$. Using this notation, a contact followed by the household agreeing to participate is denoted by $C \cap P$ (contact <u>and</u> participation), contact followed by a refusal is denoted by $C \cap \overline{P}$, and $C$ is the union of $C \cap P$ and $C \cap \overline{P}$. To highlight the similarities and differences between the three methods, especially with respect to assumptions made about residual correlations, each model is expressed in terms of continuously-distributed latent propensities $y^*$ that underlie the observed categorical response outcomes $y$. This latent variable or threshold representation of discrete response models is commonly used in econometrics (e.g. Maddala, 1983) where the propensity to ‘choose’ a particular response category, $y^*$, is often called a utility function.

### 3.1 Multinomial models

Multinomial models have been used by several authors to examine simultaneously the predictors of noncontact and refusal (Pickery and Loosveldt, 2002; O'Muircheartaigh and Campanelli, 1999; Durrant and Steele, 2009). Using this approach the outcomes of the noncontact and refusal process are combined to define a single three-category response $y_{ij}$ for household $i$ of interviewer $j$ as follows:

$$y_{ij} = \begin{cases} 1 & \text{noncontact } (\bar{C}) \\ 2 & \text{contact and refusal } (C \cap \bar{P}) \\ 3 & \text{contact and participation } (C \cap P) \end{cases} \qquad (1)$$

Denote by $y_{ij}^{(\bar{C})*}$, $y_{ij}^{(C\cap\bar{P})*}$ and $y_{ij}^{(C\cap P)*}$ the latent propensities of, respectively, noncontact, contact and refusal, and contact and participation. The observed response outcome for a particular household is the category of $y_{ij}$ for which the underlying propensity is greatest. For example, a household is classified as a noncontact ($y_{ij}$ = 1) if $y_{ij}^{(\bar{C})*} > y_{ij}^{(C\cap\bar{P})*}$ and $y_{ij}^{(\bar{C})*} > y_{ij}^{(C\cap P)*}$. Similarly, a household is a refusal if $y_{ij}^{(C\cap\bar{P})*} > y_{ij}^{(\bar{C})*}$ and $y_{ij}^{(C\cap\bar{P})*} > y_{ij}^{(C\cap P)*}$. For this reason, we can model only the *differences* between the propensities. Because our interest is in noncontact and refusal versus participation, it is natural to take $C \cap P$ as the baseline category, leading to the following equations for noncontact and refusal:

$$y_{ij}^{(\bar{C})*} - y_{ij}^{(C\cap P)*} = \boldsymbol{\alpha}^{(\bar{C})}\mathbf{x}_{ij}^{(\bar{C})} + u_j^{(\bar{C})} + e_{ij}^{(\bar{C})} \qquad (2a)$$

$$y_{ij}^{(C\cap\bar{P})*} - y_{ij}^{(C\cap P)*} = \boldsymbol{\alpha}^{(C\cap\bar{P})}\mathbf{x}_{ij}^{(C\cap\bar{P})} + u_j^{(C\cap\bar{P})} + e_{ij}^{(C\cap\bar{P})} \qquad (2b)$$

where $\mathbf{x}_{ij}^{(\bar{C})}$ and $\mathbf{x}_{ij}^{(C\cap\bar{P})}$ are vectors of covariates for noncontact and refusal with coefficient vectors $\boldsymbol{\alpha}^{(\bar{C})}$ and $\boldsymbol{\alpha}^{(C\cap\bar{P})}$, $u_j^{(\bar{C})}$ and $u_j^{(C\cap\bar{P})}$ are bivariate normal random effects representing unobserved interviewer characteristics affecting each nonresponse process, and $e_{ij}^{(\bar{C})}$ and $e_{ij}^{(C\cap\bar{P})}$ are household-specific residuals. Equations (2a) and (2b) are estimated simultaneously which ensures that the probabilities associated with each category of $y_{ij}$ in (1) sum to 1 for a given household.

A multinomial logit model arises from the assumption that $e_{ij}^{(\bar{C})}$ and $e_{ij}^{(C\cap\bar{P})}$ follow independent standard Type I extreme value distributions, while a multinomial probit model assumes that they follow a bivariate normal distribution (Maddala, 1983). The assumption that the household-level residuals are uncorrelated is commonly known as the 'independence of irrelevant alternatives' (IIA). Concern about the IIA property of the multinomial logit model in situations where there may be similarities among some of the categories of $y$ has led to a preference towards the multinomial probit in many applications, for example in models for mode of transport choice (Ben-Akiva and Lerman, 1985) and voting choices (Gordon, 2002). In the present context, similarity between the noncontact and refusal outcomes of the response process may result from misclassification (as described in the previous section), leading to a positive correlation between $e_{ij}^{(\bar{C})}$ and $e_{ij}^{(C\cap\bar{P})}$, or from unmeasured household characteristics affecting both noncontact and refusal. Nevertheless, all published examples of multinomial modelling of nonresponse have used logit rather than probit models. The assumption that the interviewer-level random effects, $u_j^{(\bar{C})}$ and $u_j^{(C\cap\bar{P})}$, follow a bivariate normal distribution allows for the possibility that the unobserved interviewer influences on noncontact and refusal may be correlated. Previous studies that have used a multilevel multinomial model, and that tested for the presence of an interviewer-level residual correlation, all found evidence of a positive correlation which was partly or wholly explained by covariates (Pickery and Loosveldt, 2002; O'Muircheartaigh and Campanelli, 1999; Durrant and Steele, 2009).

To summarise, there are three main advantages of the multinomial modelling approach: it permits testing of the equality of covariate effects across contrasts, testing for the presence of an interviewer-level residual correlation and, using a probit formulation, it is possible to allow for correlation between a sample unit's probabilities of noncontact and refusal. However, a problem with using either a multinomial logit or probit model in the present application is the interpretation of the coefficients and associated significance tests in the noncontact equation, which contrasts the events $\bar{C}$ and $C \cap P$. In a logit model, for example, the exponentiated coefficient $\exp(\alpha_k^{(\bar{C})})$ in equation (2a) represents the multiplicative effect of a one-unit change in variable $x_k^{(\bar{C})}$ on the ratio of the probability of noncontact ($\bar{C}$) to the probability of contact and participation ($C \cap P$). The comparison of these two outcomes is somewhat awkward because the baseline category combines contact and participation, making it difficult to isolate the effects of covariates on noncontact from effects on participation. It would be more natural to compare $\bar{C}$ versus $C$, i.e. the union of $C \cap P$ and $C \cap \bar{P}$. This interpretational issue can be overcome to some extent by calculating predicted response probabilities from the estimated model, but this is time-consuming and there remains the problem of how to carry out significance tests for covariate effects on $\bar{C}$ versus $C$. Furthermore, although the approach allows for the correlation between $u_j^{(\bar{C})}$ and $u_j^{(C \cap \bar{P})}$, and in a probit model also for the correlation between $e_{ij}^{(\bar{C})}$ and $e_{ij}^{(C \cap \bar{P})}$, these correlations do not have a straightforward interpretation. Specifically, $\text{corr}(u_j^{(\bar{C})}, u_j^{(C \cap \bar{P})})$ and $\text{corr}(e_{ij}^{(\bar{C})}, e_{ij}^{(C \cap \bar{P})})$ are the residual correlations at the interviewer and household level between the *differences* $y_{ij}^{(\bar{C})*} - y_{ij}^{(C \cap P)*}$ and $y_{ij}^{(C \cap \bar{P})*} - y_{ij}^{(C \cap P)*}$ which will not in general equal the residual correlations between the noncontact and refusal propensities, $y_{ij}^{(\bar{C})*}$ and $y_{ij}^{(C \cap \bar{P})*}$.

### 3.2. Sequential models

An alternative to the multinomial approach is to model noncontact and refusal as the outcomes of two sequential processes where, at the first stage, an interviewer attempts to contact the household and, at the second (conditional on successful contact), a survey request is made. In the sequential approach we model the binary outcomes of the noncontact and refusal processes directly, rather than combining them in a single categorical response. Consequently, there is a clear separation of these two components of nonresponse and the estimated coefficients have a simple interpretation. As noted by Hawkes and Plewis (2006) the sequential model recognises the ordered nature of the response outcome $y$ in (1) in that refusal and participation (categories 2 and 3) both imply contact (the complement of category 1). The two binary responses are defined as follows:

$$y_{ij}^{(\bar{C})} = \begin{cases} 1 & \text{noncontact} \\ 0 & \text{contact} \end{cases} \qquad \text{and} \qquad y_{ij}^{(\bar{P}|C)} = \begin{cases} 1 & \text{refusal} \mid \text{contact} \\ 0 & \text{participation} \mid \text{contact} \end{cases}$$

where $y_{ij}^{(\bar{P}|C)}$ is observed only when $y_{ij}^{(\bar{C})} = 0$.

As for the multinomial response $y_{ij}$, we can think of two latent propensities underlying the observed binary variables, $y_{ij}^{(\bar{C})}$ and $y_{ij}^{(\bar{P}|C)}$, which we denote by $y_{ij}^{(\bar{C})*}$ and $y_{ij}^{(\bar{P}|C)*}$. The continuous latent variables and the binary observed responses are related as follows:

$$y_{ij}^{(\bar{C})} = \begin{cases} 1 & \text{if } y_{ij}^{(\bar{C})*} > 0 \\ 0 & \text{otherwise} \end{cases} \qquad \text{and} \qquad y_{ij}^{(\bar{P}|C)} = \begin{cases} 1 & \text{if } y_{ij}^{(\bar{P}|C)*} > 0 \\ 0 & \text{otherwise} \end{cases},$$

where zero is the arbitrarily chosen threshold.

A sequential model can be written in terms of these propensities as:

$$y_{ij}^{(\bar{C})*} = \boldsymbol{\beta}^{(\bar{C})}\mathbf{x}_{ij}^{(\bar{C})} + u_j^{(\bar{C})} + e_{ij}^{(\bar{C})} \qquad \text{(3a)}$$

$$y_{ij}^{(\bar{P}|C)*} = \boldsymbol{\beta}^{(\bar{P}|C)}\mathbf{x}_{ij}^{(\bar{P}|C)} + u_j^{(\bar{P}|C)} + e_{ij}^{(\bar{P}|C)} \qquad \text{(3b)}$$

where $\boldsymbol{\beta}^{(\bar{C})}$ and $\boldsymbol{\beta}^{(\bar{P}|C)}$ are coefficient vectors associated with covariates $\mathbf{x}_{ij}^{(\bar{C})}$ and $\mathbf{x}_{ij}^{(\bar{P}|C)}$, $u_j^{(\bar{C})}$ and $u_j^{(\bar{P}|C)}$ are normally distributed interviewer random effects, and $e_{ij}^{(\bar{C})}$ and $e_{ij}^{(\bar{P}|C)}$ are household-level residuals. The assumption that the household residuals follow independent standard normal distributions leads to a sequential probit model (Maddala, 1983, Chapter 2), while standard logistic distributional assumptions lead to a sequential logit model, more commonly known as a continuation ratio model (Agresti, 1996, p.218-220). Both the sequential probit and continuation ratio model are commonly applied in the analysis of ordered categorical responses that can be viewed as the result of a set of sequential 'decisions', for example progression through a series of educational transitions until a certain level of qualifications is achieved (e.g. Brien and Lillard, 1994).

The sequential model is the most commonly used method in the nonresponse modelling literature (Hawkes and Plewis, 2006; Lepkowski and Couper, 2002; Groves and Couper, 1998) although some of these authors estimated models for only one of the two processes (Pickery et al., 2001; de Leeuw and de Heer, 2002; Durrant and Steele, 2009; Durrant et al., 2011). However, previous research using a sequential modelling approach is based on an assumption of conditional independence between noncontact and refusal, and none has allowed for correlation between the interviewer random effects for the two processes.

When all predictors of nonresponse are categorical, both the multinomial logit model and continuation ratio model can be fitted as loglinear models to cell counts in the cross-classification of $y$ and the covariates, treating the multidimensional marginal totals for the covariates as fixed (Fienberg, 1980, Chapter 6). Moreover, Fienberg reports that, in certain cases, the deviance for a multinomial logit model will be identical to the sum of the deviances for the separate continuation ratio models, and gives an example where there is little or no difference between the two approaches in terms of the goodness of fit of various models fitted to the same cross-classification. In these situations, the multinomial logit and continuation ratio models can be viewed as alternative parameterisations of the same loglinear model, in which case they would be expected to yield similar predicted response probabilities. It follows that, under the same conditions, the multinomial probit and sequential probit models should also lead to similar predicted probabilities, provided $\text{corr}(e_{ij}^{(\bar{C})}, e_{ij}^{(C\cap\bar{P})}) = 0$ in (2a) and (2b). More generally, Fienberg states that "the choice between the two approaches will depend on the substantive context of the problem and on the interpretability of the resulting models" (p.116). In the application to nonresponse, the sequential model might therefore be preferred because coefficients in equation (3a) for the propensity of noncontact are more easily interpreted than coefficients in (2a) for the difference between the propensities of noncontact and participation. Nevertheless, Fienberg recommends applying both classes of models to a given dataset and comparing their goodness of fit.

In a continuation ratio or sequential probit model, the multinomial likelihood for the cell probabilities in the cross-classification of $y$ and categorical covariates factors into two independent binomial likelihoods. This implies that equations (3a) and (3b) can be estimated separately, with estimation of the refusal equation (3b) based on the subsample of contacted cases (with $y_{ij}^{(\bar{C})} = 0$). As in the multinomial model, however, we may wish to allow for correlation between the interviewer random effects in which case (3a) and (3b) must be estimated jointly. The joint model can be framed as a type of multivariate bivariate response model, where all households have a binary response for noncontact and contacted households have a second binary response for refusal. Such a model can be estimated using any statistical software that can handle multilevel binary responses (including most mainstream packages such as SAS, Stata and R). However, while we might also wish to allow for residual correlation between a household's noncontact and refusal propensities (for the reasons given in Section 2), it is not possible to specify a joint distribution for $e_{ij}^{(\bar{C})}$ and $e_{ij}^{(\bar{P}|C)}$ in the sequential model because $e_{ij}^{(\bar{P}|C)}$ is defined only for the subset of contacted households.

### 3.3. Sample selection models

In the sequential model the propensity to refuse, $y_{ij}^{(\bar{P}|C)*}$, is defined only for households that are successfully contacted. Although the outcome of the participation decision is *observed* only for contacted households, we can think of an underlying (unconditional) refusal propensity, $y_{ij}^{(\bar{P})*}$, that is defined for *all* households regardless of whether or not they were contacted. This refusal propensity underlies a binary response $y_{ij}^{(\bar{P})}$ such that

$$y_{ij}^{(\bar{P})} = \begin{cases} 1 & \text{if } y_{ij}^{(\bar{P})*} > 0 \\ 0 & \text{otherwise} \end{cases}$$

where $y_{ij}^{(\bar{P})}$ is observed only when $y_{ij}^{(\bar{C})} = 0$.

A joint model for the noncontact propensity and the unconditional refusal propensity can be written

$$y_{ij}^{(\bar{C})*} = \boldsymbol{\beta}^{(\bar{C})}\mathbf{x}_{ij}^{(\bar{C})} + u_j^{(\bar{C})} + e_{ij}^{(\bar{C})} \qquad \text{(4a)}$$

$$y_{ij}^{(\bar{P})*} = \boldsymbol{\beta}^{(\bar{P})}\mathbf{x}_{ij}^{(\bar{P})} + u_j^{(\bar{P})} + e_{ij}^{(\bar{P})} \qquad \text{(4b)}$$

where (4a) is identical to (3a) of the sequential model. Equations (4a) and (4b) together define a multilevel version of a sample selection model where (4a) is the selection equation determining whether $y_{ij}^{(\bar{P})}$ is observed. While in the sequential model the correlation between $e_{ij}^{(\bar{C})}$ and $e_{ij}^{(\bar{P}|C)}$ cannot be estimated because $y_{ij}^{(\bar{P}|C)*}$ is defined only for contacted households, the correlation between $e_{ij}^{(\bar{C})}$ and $e_{ij}^{(\bar{P})}$ is defined and estimable (under the conditions outlined below). The model given by (4a) and (4b) is a generalisation of the single-level model used by Nicoletti and Peracchi (2005) in the context of nonresponse modelling to include random interviewer effects that may be correlated for noncontact and refusal.

Sample selection models were first developed for a continuous outcome of interest and binary selection variable (Heckman, 1979), and later generalised to binary outcomes (van de Ven and van Praag, 1981). Single-level sample selection models are now used routinely in economics and the social sciences in situations where an outcome variable is observed only for a non-random subset of the sample, due to selective nonresponse or self-selection of individuals (see Wooldridge, 2002 for an overview). Other examples of sample selection include female wages observed only for those who are in employment (Dolton and Makepeace, 1986) and the number of children born to women with at least one child (Billari and Borgoni, 2005). Sample selection models have also been proposed to allow for non-ignorable nonresponse, for example in longitudinal studies where the outcome of primary interest is unavailable for some sample members due to attrition (Diggle and Kenward, 1994). In each case the model consists of two components: a probit equation for selection into some sub-sample (e.g. for female labour force participation, having children or, in the present case, making contact with a household a survey) and an equation for an outcome which is observed only for that sub-sample (e.g. wages, number of children, or, here, a household agreeing to participate in a survey). Importantly, although the second outcome is partially observed, the corresponding component of the model is for the outcome of *each* sample member (not just the observed sub-sample). The two equations are linked through their residual terms which are usually assumed to follow a bivariate normal distribution. In our application to nonresponse, a non-zero residual correlation would suggest that contacted households differ systematically from those that are not contacted; that is, the unmeasured determinants of contact are not independent of participation. For example, Nicoletti and Peracchi (2005) anticipate a positive correlation if selection is driven primarily by misclassification of refusals as noncontacts. If sample selection is ignored, the effects of covariates on participation will be biased, leading to incorrect inferences about the refusal process.

**Identification and estimation of the multilevel sample selection model**

Identification of the sample selection model requires that the selection equation (4a) contains at least one covariate (called an instrument) that is not included in the equation for the outcome of primary interest (4b), a condition which is commonly referred to as a covariate exclusion restriction. Without instruments, identification relies on correct specification of the functional form of the model, specifically the nonlinearity of the probit transformation and the bivariate normal assumption (Wooldridge, 2002, Section 17.4.3). In our application, the covariate exclusion restriction translates to a requirement that $\mathbf{x}_{ij}^{(\bar{C})}$ contains at least one variable that is not in $\mathbf{x}_{ij}^{(\bar{P})}$, i.e. a variable that predicts noncontact but does not have a direct effect on the probability of refusal. In many applications of sample selection models and related instrumental variable methods, it is difficult to find good instruments that can be justified on theoretical grounds, leading to weak identification (e.g. Bound et al., 1995) which in turn may induce more serious bias than if selection were ignored (Brandt and Schneider, 2007) and large standard errors (Wooldridge, 2002).

In the present case, theories about the processes of noncontact and refusal (supported by empirical evidence from a range of surveys and populations) point towards several plausible instruments, i.e. variables which predict whether a household can be contacted but which do not predict whether household members are willing to participate in the survey (Groves and Couper, 1998). Previous research suggests that the probability of contact is influenced mainly by the ease of access to a household and factors related to the likelihood of finding someone at home, while the probability of participation depends more on the characteristics and attitudes of household members. In particular, the presence of physical barriers to gaining access to a household would not be expected to affect participation, after controlling for socio-economic characteristics such as economic activity and education, and this is indeed supported by empirical research on the correlates of noncontact and refusal (e.g. Durrant and Steele, 2009). Indicators of the ease of physical access to a household are observed by the interviewer at each call to a household (so-called interviewer observation variables and interviewer call record data). Other examples of call record data are indicators of the type of house and area (e.g. residential or commercial), and the time and frequency of calls. Such data are a rich source of information for modelling the nonresponse process and are commonly found to be good predictors of noncontact but not of refusal (after controlling for individual-level characteristics and attitudes) (Purdon et al., 1999; Kreuter et al., 2010; Kulka and Weeks, 1988; Durrant et al., 2010; Durrant et al., 2011).

In the application that follows, we consider several potential instruments: whether the sample household is located in a house rather than a flat or other multi-occupancy building, various indicators of physical impediments to access (a locked common entrance, locked gates, entry phones or security devices) and household type (single person, couple or multiple individuals). The first variable is a proxy for the ease of physical access while household type is an indicator for the propensity for someone to be at home when the interviewer calls. (Further details of these variables and other predictors of the noncontact and refusal processes are given in Section 4.2.)

The equations of a single-level sample selection model are traditionally estimated using a two-step procedure or jointly using maximum likelihood. Both Heckman's original two-step method and maximum likelihood rely on the assumption that the residuals follow a bivariate normal distribution, although these assumptions have been relaxed to include non-normal distributions and semi-parameteric and nonparametric two-step methods have also been developed (see Vella, 1998 for a review). In practice, however, most applications of sample selection models have assumed bivariate normality which Vella (1998) attributes both to difficulty in implementing other approaches and to research that suggests specification of the regression function is more important than distributional assumptions. Given that the model we propose has correlated random effects at both the household and the interviewer level, we follow the common practice of assuming bivariate normality.

Equations (4a) and (4b) define a multilevel censored bivariate probit model. Multivariate responses can be viewed as a type of two-level hierarchical structure, in which the responses form the level 1 units nested within sample members at level 2 (Goldstein, 2003). Consequently, a bivariate probit model can be framed as a two-level model for a pair of binary responses, and it follows that the extension to include random interviewer effects can be framed as a three-level model. Importantly, no adjustment is needed to handle the fact that $y_{ij}^{(\bar{P})}$ is observed only when $y_{ij}^{(\bar{C})} = 0$ because in a multilevel model there is no requirement for the data to be balanced. The multilevel censored bivariate probit model can be estimated as a standard 3-level probit model under the commonly made assumption that $y_{ij}^{(\bar{P})}$ is missing at random (MAR) (Little and Rubin, 2002). Here, MAR implies that, conditional on $\mathbf{x}_{ij}^{(\bar{C})}$ and $\mathbf{x}_{ij}^{(\bar{P})}$, the difference in the refusal propensities for two households are random draws from a bivariate normal distribution, regardless of whether or not they were contacted. Under MAR, a multilevel sample selection model can be estimated in a range of software packages using maximum likelihood via numerical quadrature (e.g. proc nlmixed in SAS, aML and sabre) or Markov chain Monte Carlo methods (e.g. MLwiN and WinBUGS). All analyses presented in this paper were carried out using aML (Lillard and Panis, 1998-2003).

## 4. Application of Alternative Modelling Approaches

### 4.1 The UK 2001 Census Link Study

The analysis is based on data from the UK 2001 Census Link Study which includes the response outcome of six major UK household surveys. These surveys vary in their design and subject matter, but all are face-to-face surveys administered via interviewers. The six surveys are: the Expenditure and Food Survey (EFS), the Family Resources Survey (FRS), the General Household Survey (GHS), the Omnibus Survey (OMN), the National Travel Survey (NTS) and the Labour Force Survey (LFS). The key advantage of the database is that the response outcome has been linked to various data sources, providing unusually rich information on both responding and nonresponding households. The first key source is the UK 2001 Census which gives detailed demographic and socio-economic information on households and individuals. In addition, observations about each household and the immediate neighbourhood were recorded by the interviewer during fieldwork, providing further information on both responding and nonresponding households, even if the interviewer did not establish contact with the household. This information was recorded in an interviewer observation questionnaire. Examples of information obtained are characteristics about the accommodation, whether the household lives in a house or flat, the presence of security measures and physical impediments such as locked gates, information about the household composition (e.g. indications of the presence of children), and information about the condition of the housing and the neighbourhood. Furthermore, the dataset includes information about each interviewer, collected in a separate interviewer attitudes survey in 2001. The timing of the collection of the different data sources was chosen to coincide with the last UK Census in 2001.

The response outcome recorded for the six surveys distinguishes the two main components of nonresponse: i) noncontact, where it was not possible for the interviewer to establish contact with a selected household and ii) refusal, where contact with at least one responsible resident was made but the household refused an interview. Participation is defined when the interviewer was able to carry out an interview with *at least one* member of the household. All surveys included in this study, apart from the Omnibus survey, require that all household members participate in the survey request, referred to as *full response*. The case where not all household members respond is referred to as *partial response*. Both full and partial responses are classified as responding households. The linked data are available for 18,530 households and 565 interviewers after excluding vacant and non-residential addresses, re-issues, unusable records, and records that were not linked (as described in detail in Durrant and Steele, 2009). For further information about the Census Link Study, the surveys and the linked data sources, see Durrant and Steele (2009) and Beerten and Freeth (2004).

### 4.2. Explanatory variables

The choice of explanatory variables follows Durrant and Steele (2009), who were guided by current conceptual frameworks for survey nonresponse. The effects on nonresponse are mostly based on psychological concepts such as social exchange, civic engagement and social isolation and integration.  These theories are concerned with influences on access to the sample unit, cooperation of the sample unit with the survey request, influence of the social context on individual action and the interplay of multiple effects on survey participation. Based on these theories the explanatory variables expected to be related to refusal, and available in our dataset, are indicators of the demographic and socio-economic background of the household and its occupants, including indicators of location and area, as well as attitude of householders (see also Groves and Couper, 1998). We expect noncontact to depend primarily on household characteristics (such as the presence of physical impediments to access) and lifestyle characteristics of the householders (such as proxies of the time spent at home).

Variables that were included in both the final noncontact and the refusal equations (see Tables 2 and 3 for a full listing) comprised characteristics of the household (such as indicators of moving house during the last year and car ownership), characteristics of the household representative (age, gender, highest qualification, economic activity and perceived health status), household composition (presence of children, pensioners and carers), and the location and area of residence (London versus the rest of the UK, and interviewer observation variables describing the condition of the house relative to others in the area and whether the interviewer would feel safe walking in the area after dark).

Identification of the bivariate model requires that some variables in the noncontact (selection) equation are excluded from the refusal equation (see Section 3.3). The following variables were used as instruments, i.e. predictors of noncontact but not of refusal (conditional on other factors included in the model): building type (house vs. other, including flat), household type (single, couple or multiple-occupancy), and their interactions with survey. Indicators of the presence of physical barriers to entry (e.g. locked gates and security devices) were also considered as instruments as these have been hypothesised to predict noncontact but not refusal (see Groves and Couper, 1998). However, these variables were not found to have significant effects on noncontact after adjusting for the effects of other variables such as the type of accommodation. Further exclusions from the refusal equation, based on the findings of Durrant and Steele (2009), were rural vs. urban residence and the number of employed adults in the household.

### 4.3. Results

The three types of model described in Section 3 were fitted to data from the UK Census Link Study. The aim of the analysis is to compare these approaches to modelling nonresponse, in terms of both goodness of fit and interpretation of the parameter estimates. For ease of comparison, we estimated the probit form of each model, although we note that previous investigations of nonresponse have tended to use the multinomial logit model which constrains the household-level random effect correlation to equal zero. Table 1 shows estimates of standard deviations and correlations (where estimated) of the interviewer and household-level random effects, together with the deviance statistic for each model. To explore the extent to which any residual correlation may be explained by covariates, we considered two specifications: a model with the full set of covariates (upper panel) and a reduced model without household characteristics and interviewer observations.

The deviance for the full specification is almost identical for all three models. Following Fienberg (1980) we would expect the multinomial and sequential models to have similar deviances. Of particular interest is the comparison of the sequential model and its extension, the (censored) bivariate probit model, which allows for the possibility that contacted households may differ from noncontacted households on unmeasured characteristics. A likelihood ratio test for the comparison suggests, however, that the addition of the household-level residual correlation between noncontact and refusal propensities does not improve the fit of the model ($2\ \Delta$ deviance = 0.4, 1 d.f., $p = 0.53$). We therefore conclude that there is little evidence of sample selection in the full model, which implies that contacted households can be treated as a random sample from the whole target population. A model with survey-specific household-level correlations was also considered (results not shown). Comparing this to the sequential model, there was virtually no change in deviance for an additional five parameters, indicating that there is also no sample selection for any of the six surveys.

Because it is possible that any sample selection has been explained by the covariates included in the nonresponse models, we also considered a reduced model with only the survey indicators (in both equations) and instruments. From the deviances given in the lower panel of Table 1, we find that the household-level residual correlation is now borderline significant ($2\ \Delta$ deviance = 3.8, 1 d.f., $p = 0.051$). The negative direction of the correlation indicates that the type of household that is easier to contact ($e_{ij}^{(\bar{C})} < 0$) tends to be more likely to refuse ($e_{ij}^{(\bar{P})} > 0$). A negative residual correlation is consistent with the findings of Nicoletti and Perrachi's (2005) study of attrition in a longitudinal survey (among sample members who had responded at the first wave). In the present study, however, with rich information on both respondents and nonrespondents, this selection is explained by the household characteristics included in the full model.

[TABLE 1 ABOUT HERE]

Turning to the estimates of the residual standard deviation and correlation at the interviewer level we find evidence of significant unobserved interviewer heterogeneity in noncontact and refusal for both the full and reduced model (Table 1). However, none of the residual interviewer-level correlations, for either the full or the reduced model, are significantly different from zero, which implies that interviewers with above-average contact rates are no more or less likely to persuade households to participate than interviewers with below-average contact rates. O'Muircheartaigh and Campanelli (1999) and Pickery and Loosveldt (2002) also reported a nonsignificant correlation at the interviewer level after controlling for other variables in their (multinomial logit) models.

The estimated coefficients for models with the full set of covariates are given in Table 2 (noncontact) and Table 3 (refusal). As noted in Sections 3.1 and 3.2 coefficients for the noncontact equation have a different interpretation for a multinomial model than for a sequential model (regardless of whether there is any adjustment for selection). In the multinomial model the propensity of noncontact $y_{ij}^{(\bar{C})}$ is contrasted with the propensity to be contacted *and* to participate $y_{ij}^{(C \cap P)}$, while in the sequential and sample selection models $y_{ij}^{(\bar{C})}$ is contrasted with the contact propensity $y_{ij}^{(C)}$. Given the difference in reference category, it is perhaps then surprising that the coefficients from the noncontact equation are, for most covariates, very similar for the multinomial and sequential models (see Table 2). To explore the conditions under which the noncontact coefficients would be expected to be similar for the multinomial and sequential models we conducted a simulation study (see Section 5). The sequential and bivariate (sample selection) model estimates for noncontact are expected to be close because, even if there had been support for sample selection, this should have little effect on estimates in the refusal equation.

[TABLES 2 AND 3 ABOUT HERE]

The substantive conclusions about the predictors of noncontact are the same whatever model is used. The probability of noncontact is highest in the Omnibus survey, while the following characteristics are associated with an increased chance of contact: living in a house rather than a flat or any other type of building (but only in the EFS, FRS and Omnibus), households with dependent children or pensioners, couple rather than single-person households (except in the Omnibus), and households where the household reference person (HRP) is aged 50-79 years. The noncontact probability is also lower when the building is noted by the interviewer as being in a better condition than others in the area. To summarise, noncontact is determined primarily by household characteristics, such as the presence of physical impediments, and lifestyle, such as proxies of the time spent at home. The differences between surveys could be attributed to differences in the length of fieldwork, interviewer workload, and type of interviewer training provided.

Turning to the refusal equation, the estimated coefficients are almost identical across the different models (Table 3), but this is expected for the following reasons. The similarity between the multinomial and sequential estimates is easiest to see if we consider a logit link. Coefficients from the refusal component of a multinomial logit model are effects on the log of the ratio of $\Pr(C \cap \overline{P})$ to $\Pr(C \cap P)$, while their counterparts in a sequential logit (continuation ratio) model are effects on the log of the ratio of $\Pr(\overline{P} \mid C)$ to $\Pr(P \mid C)$. Using the facts that $\Pr(C \cap \overline{P}) = \Pr(C) \cdot \Pr(\overline{P} \mid C)$ and $\Pr(C \cap P) = \Pr(C) \cdot \Pr(P \mid C)$, it can be seen that these ratios are equal. The closeness of the sequential and bivariate model estimates is anticipated because of the non-significance of the household residual correlation in the bivariate model, i.e. absence of sample selection.

The probability of refusal is highest in the EFS and lowest in the LFS, and higher in London than in other areas for all surveys except for the LFS. The following household characteristics are associated with an increased chance of refusal: HRPs with a low level of education, self-employed HRPs (especially in the FRS), and not having dependent children. Living in a building that is noted as being in a poor condition by the interviewer is also associated with a higher chance of refusal. The variation among surveys may be due to differences in the survey topics and in the response burden, such as use of diaries and interview length. For a more detailed discussion of the substantive findings and their links to psychological and sociological concepts and response theories see Durrant and Steele (2009).

## 5. Simulation study

In the previous section multinomial and sequential probit models for nonresponse were fitted to the Census Link Study dataset, and it was noted that the estimated coefficients for the noncontact equations were very similar for the two models. This result was unexpected given that the two sets of coefficients represent effects on different contrasts: noncontact versus contact and participation in the multinomial model, and noncontact versus contact in the sequential model. A simulation study was therefore carried out to explore the conditions under which coefficients for the multinomial and sequential models are similar, and when they are expected to be different.

Following the two-stage nature of the survey response process, data were simulated from a sequential model: for each unit a binary noncontact indicator, $y^{(\bar{C})}$, was generated and then a binary indicator for refusal was generated for units with $y^{(\bar{C})} = 0$. A total of 100 datasets were created, each with sample size 10,000. A logit link was used because estimation of multinomial probit models is highly computationally intensive, but the same conclusions would be reached whatever the choice of link function. To further simplify the simulations, an unclustered data structure was assumed because our interest centres on the regression coefficients rather than the random effect parameters. The binary responses $y^{(\bar{C})}$ and $y^{(\bar{P}|C)}$ were generated from the following single-level sequential logit model:

$$\log\left(\frac{\Pr(y^{(\bar{C})}=1)}{\Pr(y^{(\bar{C})}=0)}\right)=\alpha_0+\alpha_1 x_1+\alpha_2 x_2$$

$$\log\left(\frac{\Pr(y^{(\bar{P})}=1\mid y^{(\bar{C})}=0)}{\Pr(y^{(\bar{P})}=0\mid y^{(\bar{C})}=0)}\right)=\beta_0+\beta_1 x_1+\beta_2 x_2$$

where $x_1$ is a binary variable with 50% in each category and $x_2$ follows a standard normal distribution.

A total of 16 simulation conditions were considered, based on different combinations of values assumed for $\alpha_1, \alpha_2, \beta_1$ and $\beta_2$. In all simulations, $\alpha_0$ and $\beta_0$ were fixed to give baseline noncontact and refusal probabilities of 0.04 and 0.23 respectively (the average noncontact and refusal rates across the six surveys in our dataset). $\alpha_1$ was chosen such that, for $x_2$ = 0, the probability of noncontact was either 0.04 ($\alpha_1$ = 0) or 0.08 ($\alpha_1$ = 0.736). In a similar way, $\beta_1$ was chosen such that, for $x_2$ = 0, the probability of refusal was either 0.23 ($\beta_1$ = 0) or 0.28 ($\beta_1$ = 0.264). The coefficients of $x_2$, $\alpha_2$ and $\beta_2$, were fixed at 0 or 0.5. For each simulated $y^{(\bar{C})}$ and $y^{(\bar{P}|C)}$, a multinomial response $y$ was created and coded as in equation (1). A sequential logit model was then fitted to $y^{(\bar{C})}$ and $y^{(\bar{P}|C)}$, and a multinomial logit model fitted to $y$.

Table 4 shows the mean of the estimated coefficients across the 100 simulations for each simulation condition. As anticipated, given that the data were generated under a sequential logit model, the 'sequential' estimates of all parameters are close to their true values for all simulation conditions. However, we would not in general expect the multinomial estimates of $\alpha_1$ and $\alpha_2$ to be close to the true values (or to their 'sequential' estimates) because of the difference in the base category for the two models ($C$ in the sequential model and $C \cap P$ in the multinomial model). In fact, we find that the multinomial model produces unbiased estimates of $\alpha_1$ under two scenarios: (i) $x_1$ predicts noncontact but not refusal ($\alpha_1$ = 0.736, $\beta_1$ = 0, i.e. conditions 5-8), or (ii) $x_1$ predicts neither noncontact nor refusal ($\alpha_1$ = $\beta_1$ = 0, i.e. conditions 13-16). Consistent estimates of $\alpha_2$ are obtained under similar conditions for the effects of $x_2$ on noncontact and refusal ($\alpha_2$ and $\beta_2$).

[TABLE 4 ABOUT HERE]

Returning to the results from the Census Link Study (Tables 2 and 3), we find that there is little overlap between the predictors of noncontact and refusal; for most covariates that appear in both equations, one of the above scenarios holds. In the few cases where a coefficient is significantly different from zero in both the noncontact and refusal equations (dummies for the Omnibus survey, dependent children and poor building condition) the covariate does not have a strong effect on both components of nonresponse. (The overall noncontact rate is 4%, so apparently large coefficients translate to small effects on the probability of noncontact.) The lack of significance of the household and interviewer-level residual correlations provides further evidence that the processes of noncontact and refusal are almost independent.

## 6. Conclusions

We have reviewed two widely used approaches to modelling survey noncontact and refusal – multinomial and sequential models – and described multilevel versions of each that allow for clustering in each component of nonresponse by interviewer. A more recently applied method – the sample selection model, which allows for residual correlation between a sample unit's noncontact and refusal propensities – was also considered and extended to incorporate interviewer effects. All three methods have been compared in terms of their assumptions about the underlying nonresponse process and the interpretation of regression coefficients. In particular we note that, although the predicted probabilities derived from multinomial and sequential models will tend to be very similar, the coefficients and random effect parameters for the refusal equation have a different interpretation depending on the type of model used. Moreover, failure to account for residual dependency between sample units' ease of contact and chance of participation could lead to biased estimates if a multinomial (logit) or sequential model is used. We have also compared the three methods empirically in an analysis of household nonresponse in the UK, using a rich dataset with information on both respondents and nonrespondents.

In spite of the fact that the reference category for the noncontact equation differs for the multinomial model (contact *and* participation) and the sequential model (contact), we find that the estimated coefficients are surprisingly similar in our application. Further examination of the predictors of noncontact and refusal, supported by simulation results, indicates that this similarity arises because the two processes are largely distinct: variables that have significant effects on noncontact tend to be unimportant for refusal, and vice versa. These findings suggest that while estimates from a sequential model have a simpler interpretation, estimated coefficients from a multinomial model are a good approximation to effects on the log-odds of noncontact (versus contact). Whichever model is used, however, calculation of predicted probabilities of noncontact and refusal is generally recommended in order to assess the magnitude of effects.

We find that, after controlling for a range of household characteristics, there is little evidence of residual correlation between noncontact and refusal at either the household or the interviewer level. In other words, conditional on covariates, contacted households are no more or less likely to refuse to participate than noncontacted households, and it is therefore reasonable to make inferences about the determinants of refusal in the whole target population based on analysis of refusal among those contacted using either a multinomial or sequential model. We note, however, that our dataset is unusual in that it provides detailed information on nonrespondents. Before adjusting for household characteristics, there is some evidence of a negative residual correlation at the household level, suggesting that households that are easier to contact might be more likely to refuse. In many applications there is a paucity of information about nonrespondents, leading to a lack of variables available for inclusion in the nonresponse models and therefore possibly unexplained correlation between noncontact and refusal at the household level. However, identification of the bivariate probit model that allows for such correlation depends on finding variables (instruments) that predict contact but not of participations. Good candidates for instruments are interviewer observation and call record data (Purdon et al., 1999; Durrant et al., 2011), such as indicators of ease of access or at-home patterns.

The lack of dependency between noncontact and refusal propensities is consistent with previous research that has found noncontact and refusal to be quite distinct processes. This paper therefore provides further evidence to support their treatment as separate outcomes rather than as a single nonresponse outcome. The implications of this finding for survey practice are wide ranging. Survey agencies aiming to reduce nonresponse should implement different strategies to reduce both components of nonrespondents, such as increasing the length of the fieldwork period to improve contact rates and training of interviewers to better target groups with lower participation rates. The findings also have implications for adjustment and estimation at the data analysis stage to reduce nonresponse bias. For example, the treatment of nonresponse as the outcome of a two-stage sequential process may lead to a sequential weighting method to adjust for nonresponse, first weighting for noncontact and then for refusal, with the final weights reflecting the selection in the two stages of the response process (as proposed by Groves and Couper, 1998). Alternative approaches to nonresponse weighting based on different nonresponse models have also been explored in Cobben (2009, Chapter 8). Further research is currently underway to investigate how best to allow for interviewer effects and other correlation structures in nonresponse weighting models.

**Acknowledgements**

The research was funded by the Economic and Social Research Council (ESRC), UK, grant number: RES-062-23-0458. The authors appreciate comments on an earlier draft from Paul Clarke.

The authors would like to thank the UK Office for National Statistics (ONS) for allowing access to the dataset. Any views expressed in this paper are those of the authors.

## References

Agresti, A. (1996) *An Introduction to Categorical Data Analysis,* John Wiley & Sons, Inc, New York.

Beerten, R. and Freeth, S. (2004) Exploring survey nonresponse in the UK: the Census-Survey Nonresponse Link Study, In *Office for National Statistics, working paper 1-16*.

Ben-Akiva, M. E. and Lerman, S. R. (1985) *Discrete Choice Analysis: Theory and Application to Travel Demand,* MIT Press, Cambridge, Massachusetts.

Bethlehem, J., Cobben, F. and Schouten, B. (2011) *Handbook in Nonresponse in Household Surveys,* Wiley, New York.

Billari, F. C. and Borgoni, R. (2005) Assessing the use of sample selection models in the estimation of fertility postponement effects, *Statistical Methods and Applications,* **14,** 389-402.

Bound, J., Jaeger, D. A. and Baker, R. M. (1995) Problems with instrumental variable estimation when the correlation between the instruments and the endogenous explanatory variable is weak, *Journal of the American Statistical Association,* **90,** 443-450.

Brandt, P. T. and Schneider, C. J. (2007) So the reviewer told you to use a selection model? Selection models and the study of international relations, University of Texas at Dallas.

Brien, M. J. and Lillard, L. A. (1994) Education, marriage and first conception in Malaysia, *The Journal of Human Resources,* **29**.

Cobben, F. (2009) Nonresponse in sample surveys: Methods for analysis and adjustment, In *PhD thesis, University of Amsterdam and Central Bureau of Statistics Netherlands*Amsterdam.

de Heer, W. (1999) International response trends: results of an international survey, *Journal of Official Statistics,* **15,** 129-142.

de Leeuw, E. and de Heer, W. (2002) Trends in household survey nonresponse: a longitudinal and international comparison, In *Survey Nonresponse*(Eds, Groves, R. M., Dillman, D. A., Eltinge, J. L. and Little, R. J. A.) Wiley, New York, pp. 41-54.

Diggle, P. and Kenward, M. G. (1994) Informative drop-out in longitudinal data analysis, *Applied Statistics,* **43,** 49-93.

Dolton, P. J. and Makepeace, G. H. (1986) Sample selection and male-female earnings differentials in the graduate labour market, *Oxford Economic Papers,* **38,** 317-341.

Durrant, G. B., D'Arrigo, J. and Steele, F. (2011) Using field process data to predict best times of contact conditioning on household and interviewer influences, *Journal of the Royal Statistical Society A (Statistics in Society) (forthcoming)*.

Durrant, G. B., Groves, R. M., Staetsky, L. and Steele, F. (2010) Effects of interviewer attitudes and behaviors on refusal in household surveys, *Public Opinion Quarterly,* **74,** 1-36.

Durrant, G. B. and Steele, F. (2009) Multilevel modelling of refusal and noncontact nonresponse in household surveys: evidence from six UK government surveys, *Journal of the Royal Statistical Society A (Statistics in Society),* **172,** 361-381.

Fienberg, S. E. (1980) *The Analysis of Cross-Classified Categorical Data,* The MIT Press, Cambridge, Massachusetts.

Goldstein, H. (2003) *Multilevel Statistical Models,* Arnold, London.

Gordon, S. C. (2002) Stochastic dependence in competing risks, *American Journal of Political Science,* **46,** 200-217.

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

Hawkes, D. and Plewis, I. (2006) Modelling non-response in the National Child Development Study, *Journal of the Royal Statistical Society A (Statistics in Society),* **169,** 479-491.
Heckman, J. J. (1979) Sample selection bias as a specification error, *Econometrica,* **47,** 153-161.
Hill, D. H., Axinn, W. G. and Thornton, A. (1993) Competing hazards with shared unmeasured risk factors, *Sociological Methodology,* **23,** 245-277.
Kreuter, F., Olson, K., Wagner, J., Yan, T., Ezzati-Rice, T. M., Casas-Cordero, C., Lemay, M., Peytchev, A., Groves, R. M. and Raghunathan, T. E. (2010) Using proxy measures and other correlates of survey outcomes to adjust for non-response: examples from multiple surveys, *Journal of the Royal Statistical Society A (Statistics in Society), to appear,* **173,** 389-407.
Kulka, R. A. and Weeks, M. F. (1988) Towards the development of optimal calling protocols for telephone surveys: a conditional probabilities approach, *Journal of Official Statistics,* **4,** 319-358.
Lepkowski, J. M. and Couper, M. P. (2002) Nonresponse in longitudinal household surveys, In *Survey Nonresponse*(Eds, Groves, R. M., Dillman, D. A., Eltinge, J. L. and Little, R. J. A.) Wiley, New York, pp. 259-272.
Lillard, L. A. and Panis, C. W. A. (1998-2003) *aML User's Guide and Reference Manual, Version 2,* EconWare, Los Angeles.
Lynn, P. and Clarke, P. (2002) Separating refusal bias and non-contact bias: evidence from UK national surveys, *The Statistician,* **51,** 319-333.
Lynn, P., Clarke, P., Martin, J. and Sturgis, P. (2002) The effects of extended interviewer efforts on non-response bias, In *Survey Nonresponse*(Eds, Groves, R. M., Dillman, D. A., Eltinge, J. L. and Little, R. J. A.) Wiley, Chichester, pp. 135-147.
Maddala, G. S. (1983) *Limited Dependent and Qualitative Variables in Econometrics* Cambridge University Press, Cambridge.
Nicoletti, C. and Peracchi, F. (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,** 763-781.
O'Muircheartaigh, C. and Campanelli, P. (1999) A multilevel exploration of the role of interviewers in survey nonresponse, *Journal of the Royal Statistical Society A (Statistics in Society),* **162,** 437-446.
Pickery, J. and Loosveldt, G. (2002) A multilevel multinomial analysis of interviewer effects on various components of unit nonresponse, *Quality and Quantity,* **36,** 427-437.
Pickery, J., Loosveldt, G. and Carton, A. (2001) The effects of interviewer and respondent characteristics on response behavior in panel surveys, *Sociological Methods and Research,* **29,** 509-523.
Purdon, S., Campanelli, P. and Sturgis, P. (1999) Interviewers' calling strategies on face-to-face interview surveys, *Journal of Official Statistics,* **15,** 199-216.
Steele, F., Kallis, C. and Joshi, H. (2006) The formation and outcomes of cohabiting and marital partnerships in early adulthood: the role of previous partnership experience, *Journal of the Royal Statistical Society, Series A,* **169,** 757-780.
van de Ven, W. P. M. M. and van Praag, B. M. S. (1981) The demand for deductibles in private health insurance: A probit model with sample selection, *Journal of Econometrics,* **17,** 229-252.
Vella, F. (1998) Estimating models with sample selection bias: a survey, *The Journal of Human Resources,* **33,** 127-169.
Wooldridge, J. M. (2002) *Econometric Analysis of Cross Section and Panel Data,* The MIT Press, Cambridge, Massachusetts.

*Table 1*. Estimates (and standard errors) of residual standard deviations and correlations from nonresponse models

| | Multinomial probit Estimate (SE) | Sequential probit Estimate (SE) | Bivariate probit Estimate (SE) |
|---|---|---|---|
| **Full model** | | | |
| *Interviewer-level* | | | |
| Noncontact st. dev. | 0.319** (0.033) | 0.309** (0.032) | 0.308** (0.032) |
| Refusal st. dev. | 0.167** (0.018) | 0.169** (0.018) | 0.169** (0.018) |
| Noncontact-refusal correlation | 0.221 (0.215) | 0.170 (0.163) | 0.145 (0.182) |
| *Household-level*[a] | | | |
| Noncontact-refusal correlation | 0.141 (0.399) | - | -0.101 (0.319) |
| *Model deviance (- log-likelihood)* | 12029.4 | 12029.1 | 12028.9 |
| **Reduced model** | | | |
| *Interviewer-level* | | | |
| Noncontact st. dev. | 0.320** (0.029) | 0.312** (0.026) | 0.310** (0.026) |
| Refusal st. dev. | 0.172** (0.015) | 0.177** (0.015) | 0.172** (0.015) |
| Noncontact-refusal correlation | 0.123 (0.169) | 0.218 (0.136) | 0.103 (0.147) |
| *Household-level*[a] | | | |
| Noncontact-refusal correlation | -0.291 (0.455) | - | -0.439 (0.298) |
| *Model deviance (- log-likelihood)* | 12344.6 | 12346.5 | 12344.6 |

[a] Household-level residual standard deviations fixed at 1 in all models.

** $p < 0.01$; * $0.01 \leq p < 0.05$

*Table 2*. Estimated coefficients (and standard errors) for the noncontact equation of multilevel models for nonresponse

| Variable | Multinomial probit† Estimate | (SE) | Sequential probit† Estimate | (SE) | Bivariate probit† Estimate | (SE) |
|---|---|---|---|---|---|---|
| Constant | -0.963 | (0.272) | -1.128 | (0.219) | -1.132 | (0.220) |
| **Household-level** | | | | | | |
| Survey (ref=EFS) [a] | | | | | | |
| FRS | 0.152 | (0.197) | 0.196 | (0.190) | 0.195 | (0.191) |
| GHS | -0.183 | (0.195) | -0.112 | (0.187) | -0.114 | (0.188) |
| Omnibus | 0.363* | (0.161) | 0.400** | (0.144) | 0.403** | (0.145) |
| NTS | -0.386 | (0.198) | -0.326 | (0.195) | -0.325 | (0.194) |
| LFS | -0.213 | (0.178) | -0.099 | (0.159) | -0.101 | (0.159) |
| Highest qualification of HRP [b] (ref=no academic) | | | | | | |
| O/A levels, GCSEs | -0.094 | (0.634) | -0.076 | (0.062) | -0.074 | (0.062) |
| First/higher degree | -0.066 | (0.092) | -0.017 | (0.083) | -0.016 | (0.084) |
| Other | -0.066 | (0.114) | -0.046 | (0.110) | -0.045 | (0.110) |
| House (ref=other, e.g. flat) [c] | -0.560** | (0.124) | -0.544** | (0.121) | -0.546** | (0.121) |
| Dependent children | -0.309** | (0.058) | -0.283** | (0.056) | -0.284** | (0.057) |
| London[c] | 0.346 | (0.204) | 0.272 | (0.201) | 0.272 | (0.272) |
| Rural area | -0.121 | (0.097) | -0.103 | (0.096) | -0.110 | (0.096) |
| Female (HRP) | -0.104 | (0.058) | -0.111 | (0.058) | -0.110 | (0.058) |
| Economic activity of HRP (ref=employee) | | | | | | |
| Self-employed | 0.053 | (0.083) | 0.022 | (0.081) | 0.023 | (0.080) |
| Unemployed | 0.123 | (0.170) | 0.109 | (0.169) | 0.110 | (0.169) |
| Retired | 0.098 | (0.165) | 0.108 | (0.158) | 0.114 | (0.159) |
| Looking after family | -0.236 | (0.205) | -0.226 | (0.201) | -0.221 | (0.203) |
| Other (including student) | 0.019 | (0.143) | 0.016 | (0.142) | 0.018 | (0.142) |
| Pensioner in household | -0.283* | (0.143) | -0.291* | (0.133) | -0.291* | (0.136) |
| Perception of health of HRP (ref=good) | | | | | | |
| Fairly good | -0.032 | (0.055) | -0.040 | (0.054) | -0.041 | (0.054) |
| Not good | -0.029 | (0.080) | -0.038 | (0.078) | -0.037 | (0.079) |
| Carer in household | -0.027 | (0.063) | -0.015 | (0.062) | -0.016 | (0.062) |
| Household type (ref=single person) [c] | | | | | | |
| Couple | -0.573** | (0.160) | -0.568** | (0.156) | -0.563** | (0.156) |
| Multiple | -0.425 | (0.317) | -0.021 | (0.312) | -0.022 | (0.310) |
| Number adults employed (ref=none) | | | | | | |
| One | 0.241 | (0.126) | 0.236 | (0.124) | 0.240 | (0.124) |
| Two or more | 0.229 | (0.139) | 0.223 | (0.138) | 0.228 | (0.137) |
| Age of HRP (ref=16-34) | | | | | | |
| 35-49 | -0.099 | (0.068) | -0.110 | (0.066) | -0.110 | (0.070) |
| 50-64 | -0.265** | (0.077) | -0.272** | (0.073) | -0.275** | (0.073) |
| 65-79 | -0.367* | (0.182) | -0.365* | (0.180) | -0.367* | (0.181) |
| 80 or older | -0.347 | (0.252) | -0.362 | (0.250) | -0.362 | (0.251) |
| Household has no car | 0.076 | (0.059) | 0.071 | (0.058) | 0.069 | (0.058) |
| Household moved in last year | -0.001 | (0.079) | 0.009 | (0.076) | 0.009 | (0.076) |
| **Interviewer observations** | | | | | | |
| Condition of building relative to others in area (ref=better) | | | | | | |
| Worse | 0.360** | (0.101) | 0.314** | (0.097) | 0.314** | (0.098) |
| About the same | 0.022 | (0.085) | 0.014 | (0.084) | 0.013 | (0.084) |
| Would feel safe walking in area after dark | -0.128 | (0.077) | -0.119 | (0.075) | -0.119 | (0.076) |

*continued…*

| **Interactions between survey and household-level variables** | | | | | | |
|---|---|---|---|---|---|---|
| Survey × London | | | | | | |
| FRS and London | -0.534 | (0.302) | -0.470 | (0.292) | -0.475 | (0.293) |
| GHS and London | -0.471 | (0.283) | -0.415 | (0.276) | -0.419 | (0.276) |
| Omnibus and London | -0.021 | (0.270) | 0.023 | (0.254) | 0.021 | (0.254) |
| NTS and London | 0.006 | (0.351) | 0.014 | (0.348) | 0.012 | (0.346) |
| LFS and London | -0.296 | (0.284) | -0.208 | (0.278) | -0.208 | (0.278) |
| Survey × House (vs. flat and other) | | | | | | |
| FRS and House | 0.022 | (0.225) | 0.015 | (0.222) | 0.016 | (0.223) |
| GHS and House | 0.419* | (0.203) | 0.406* | (0.200) | 0.407* | (0.201) |
| Omnibus and House | 0.254 | (0.157) | 0.252 | (0.154) | 0.250 | (0.156) |
| NTS and House | 0.472* | (0.205) | 0.456* | (0.200) | 0.458* | (0.200) |
| LFS and House | 0.362* | (0.164) | 0.343* | (0.162) | 0.347* | (0.162) |
| Survey × Household type | | | | | | |
| FRS and Couple | 0.113 | (0.225) | 0.094 | (0.222) | 0.094 | (0.223) |
| GHS and Couple | 0.038 | (0.207) | 0.024 | (0.204) | 0.021 | (0.205) |
| Omnibus and Couple | 0.416* | (0.176) | 0.415* | (0.170) | 0.411* | (0.172) |
| NTS and Couple | 0.082 | (0.208) | 0.082 | (0.206) | 0.077 | (0.207) |
| LFS and Couple | 0.221 | (0.186) | 0.210 | (0.184) | 0.207 | (0.184) |
| FRS and Multiple | -0.230 | (0.522) | -0.246 | (0.514) | -0.247 | (0.513) |
| GHS and Multiple | -0.713 | (0.567) | -0.759 | (0.556) | -0.742 | (0.560) |
| Omnibus and Multiple | -0.047 | (0.453) | -0.070 | (0.443) | -0.068 | (0.444) |
| NTS and Multiple | -0.264 | (0.497) | -0.291 | (0.481) | -0.285 | (0.485) |
| LFS and Multiple | -0.532 | (0.500) | -0.557 | (0.497) | -0.560 | (0.495) |

Notes:

† Estimates are effects on propensity of noncontact versus: propensity of contact and participate (multinomial), and propensity of contact (sequential and bivariate)

** $p < 0.01$; * $0.01 \leq p < 0.05$

[a] EFS = Expenditure and Food Survey, FRS = Family Resources Survey, GHS = General Household Survey, NTS = National Travel Survey, LFS = Labour Force Survey

[b] HRP is household reference person

[c] Variable interacts with survey, so the main effect represents the effect in the EFS (the reference category for Survey)

*Table 3.* Estimated coefficients (and standard errors) for the refusal equation of multilevel models for nonresponse

| Variable | Multinomial probit† Estimate | (SE) | Sequential probit† Estimate | (SE) | Bivariate probit† Estimate | (SE) |
|---|---|---|---|---|---|---|
| Constant | -0.550 | (0.084) | -0.529 | (0.079) | -0.542 | (0.088) |
| **Household-level** | | | | | | |
| Survey (ref=EFS) [a] | | | | | | |
| FRS | -0.273** | (0.066) | -0.272** | (0.067) | -0.273** | (0.067) |
| GHS | -0.407** | (0.058) | -0.408** | (0.058) | -0.409** | (0.058) |
| Omnibus | -0.245** | (0.059) | -0.235** | (0.055) | -0.244** | (0.060) |
| NTS | -0.360** | (0.051) | -0.365** | (0.051) | -0.364** | (0.051) |
| LFS | -0.684** | (0.059) | -0.687** | (0.058) | -0.689** | (0.059) |
| Highest qualification of HRP [b] (ref=no academic) | | | | | | |
| O/A levels, GCSEs | -0.110** | (0.033) | -0.114** | (0.034) | -0.113** | (0.034) |
| First/higher degree | -0.278** | (0.044) | -0.281** | (0.045) | -0.281** | (0.045) |
| Other | -0.122* | (0.057) | -0.125* | (0.058) | -0.124* | (0.058) |
| Dependent children | -0.139** | (0.037) | -0.150** | (0.034) | -0.143** | (0.038) |
| London[c] | 0.273** | (0.104) | 0.284** | (0.107) | 0.281** | (0.108) |
| Female (HRP) | 0.023 | (0.028) | 0.022 | (0.028) | 0.022 | (0.028) |
| Economic activity HRP (ref=employee) [c] | | | | | | |
| Self-employed | 0.231** | (0.063) | 0.232** | (0.063) | 0.230** | (0.064) |
| Unemployed | 0.148 | (0.090) | 0.155 | (0.091) | 0.153 | (0.092) |
| Retired | -0.016 | (0.077) | -0.017 | (0.078) | -0.018 | (0.078) |
| Looking after family | 0.010 | (0.096) | 0.007 | (0.096) | 0.010 | (0.096) |
| Other (including student) | 0.070 | (0.071) | 0.074 | (0.071) | 0.073 | (0.071) |
| Pensioner in household | 0.036 | (0.062) | 0.031 | (0.061) | 0.036 | (0.062) |
| Perception of health of HRP (ref=good) | | | | | | |
| Fairly good | 0.069* | (0.030) | 0.069* | (0.031) | 0.069* | (0.031) |
| Not good | 0.064 | (0.040) | 0.063 | (0.040) | 0.064 | (0.040) |
| Carer in household | -0.070* | (0.033) | -0.073* | (0.034) | -0.072* | (0.034) |
| Age of HRP (ref=16-34) | | | | | | |
| 35-49 | 0.073 | (0.040) | 0.072 | (0.040) | 0.075 | (0.041) |
| 50-64 | 0.077 | (0.048) | 0.070 | (0.046) | 0.076 | (0.049) |
| 65-79 | 0.027 | (0.085) | 0.016 | (0.084) | 0.022 | (0.086) |
| 80 or older | 0.085 | (0.111) | 0.073 | (0.111) | 0.079 | (0.111) |
| Household has no car | 0.037 | (0.034) | 0.045 | (0.033) | 0.042 | (0.035) |
| Household moved in last year | -0.084 | (0.049) | -0.088 | (0.050) | -0.089 | (0.050) |
| **Interviewer observations** | | | | | | |
| Condition of building relative to others in area (ref=better) | | | | | | |
| Worse | 0.229** | (0.064) | 0.242** | (0.062) | 0.236** | (0.065) |
| About the same | 0.055 | (0.041) | 0.057 | (0.041) | 0.056 | (0.041) |
| Would feel safe walking in area after dark | -0.045 | (0.041) | -0.050 | (0.041) | -0.047 | (0.042) |

*continued…*

| **Interactions between survey and household-level variables** | | | | | | |
|---|---|---|---|---|---|---|
| Survey × Economic activity | | | | | | |
| FRS and Self-employed | 0.184* | (0.083) | 0.189* | (0.085) | 0.188* | (0.085) |
| GHS and Self-employed | 0.111 | (0.074) | 0.107 | (0.074) | 0.106 | (0.074) |
| Omnibus and Self-employed | -0.005 | (0.085) | 0.006 | (0.084) | 0.0003 | (0.086) |
| NTS and Self-employed | 0.151 | (0.079) | 0.154 | (0.080) | 0.154 | (0.080) |
| LFS and Self-employed | 0.140 | (0.078) | 0.140 | (0.078) | 0.139 | (0.078) |
| Survey × London | | | | | | |
| FRS and London | -0.120 | (0.164) | -0.134 | (0.168) | -0.131 | (0.168) |
| GHS and London | -0.091 | (0.170) | -0.104 | (0.168) | -0.098 | (0.174) |
| Omnibus and London | -0.094 | (0.148) | -0.069 | (0.156) | -0.080 | (0.155) |
| NTS and London | 0.044 | (0.164) | 0.040 | (0.167) | 0.040 | (0.168) |
| LFS and London | -0.348* | (0.176) | -0.357* | (0.178) | -0.355 | (0.180) |

Notes:

†Estimates are effects on propensity of refusal versus: propensity of contact and participation (multinomial), propensity of participation given contact (sequential) and unconditional propensity of participation (bivariate)

** $p < 0.01$; * $0.01 \leq p < 0.05$

[a] EFS = Expenditure and Food Survey, FRS = Family Resources Survey, GHS = General Household Survey, NTS = National Travel Survey, LFS = Labour Force Survey

[b] HRP is household reference person

[c] Variable interacts with survey, so the main effect represents the effect in the EFS (the reference category for Survey)

*Table 4*. Results from a simulation study comparing coefficients from the noncontact equation of sequential logit and multinomial logit models

| | Coefficient of $X_1$ | | | | Coefficient of $X_2$ | | | |
|---|---|---|---|---|---|---|---|---|
| | Refusal | Noncontact | | | Refusal | Noncontact | | |
| Condition | True $\beta_1$ | True $\alpha_1$ | Sequential | Multinomial | True $\beta_2$ | True $\alpha_2$ | Sequential | Multinomial |
| 1 | 0.264 | 0.736 | 0.726 | 0.807 | 0.5 | 0.5 | 0.500 | 0.657 |
| 2 | 0.264 | 0.736 | 0.738 | 0.802 | 0 | 0.5 | 0.509 | 0.509 |
| 3 | 0.264 | 0.736 | 0.726 | 0.797 | 0.5 | 0 | 0.001 | 0.139 |
| 4 | 0.264 | 0.736 | 0.744 | 0.812 | 0 | 0 | 0.001 | 0.001 |
| 5 | 0 | 0.736 | 0.735 | 0.736 | 0.5 | 0.5 | 0.492 | 0.630 |
| 6 | 0 | 0.736 | 0.747 | 0.746 | 0 | 0.5 | 0.500 | 0.500 |
| 7 | 0 | 0.736 | 0.720 | 0.721 | 0.5 | 0 | 0.005 | 0.125 |
| 8 | 0 | 0.736 | 0.735 | 0.734 | 0 | 0 | 0.007 | 0.007 |
| 9 | 0.264 | 0 | 0.007 | 0.090 | 0.5 | 0.5 | 0.498 | 0.653 |
| 10 | 0.264 | 0 | 0.019 | 0.084 | 0 | 0.5 | 0.503 | 0.503 |
| 11 | 0.264 | 0 | 0.004 | 0.074 | 0.5 | 0 | -0.004 | 0.130 |
| 12 | 0.264 | 0 | -0.011 | 0.055 | 0 | 0 | -0.004 | -0.004 |
| 13 | 0 | 0 | 0.003 | 0.001 | 0.5 | 0.5 | 0.504 | 0.647 |
| 14 | 0 | 0 | 0.012 | 0.012 | 0 | 0.5 | 0.499 | 0.499 |
| 15 | 0 | 0 | -0.003 | -0.005 | 0.5 | 0 | 0.002 | 0.123 |
| 16 | 0 | 0 | 0.015 | 0.014 | 0 | 0 | -0.001 | -0.001 |

Notes: (i) Coefficients in 'sequential' and 'multinomial' columns are the means of the coefficient estimates from fitting sequential and multinomial logit models to 100 simulated datasets, (ii) the 'sequential' coefficients are interpreted as effects of $X_1$ and $X_2$ on the log-odds of noncontact versus contact, while the 'multinomial' coefficients are effects on the log-odds of noncontact versus contact and participate.