Explaining differences across countries in survey contact rates

Annelies G. Blom[1]
[1]GESIS, P.O. Box 12 21 55, 68072 Mannheim, Germany

Abstract
Understanding the process leading to differential nonresponse rates across countries is a first step towards understanding cross-national differences in nonresponse bias. Recent developments in cross-national survey methodology have made auxiliary data available with which the processes leading to nonresponse can be researched. The European Social Survey (ESS), for example, collects detailed contact and neighbourhood data for each sample unit, i.e. respondents and nonrespondents. Using these data and applying decomposition analyses to within-country logit models of contact, this paper looks at differences in the correlates of non-contact across countries. Specifically it investigates whether differential contact rates can be attributed to manipulable and non-manipulable characteristics, the effects these characteristics have on the probability of contact (i.e. the coefficients) or a combination of the two.

Key Words: decomposition, cross-national surveys, non-response, survey process

1. Introduction

Inferences about differences between groups rely on the assumption that survey errors are comparable across groups. If this assumption fails and sampling or non-sampling errors (of coverage, nonresponse and measurement) differ, then any differences detected between groups may merely be artifacts of the data. This problem arises for any group comparison, but it is accentuated for cross-national comparisons, where the scope for differential errors between groups is magnified.

One aspect that might hinder cross-national comparisons is differential nonresponse bias. Statistical inference assumes that the data were drawn from the population by means of probability sampling and that each sampled unit is actually interviewed and their data observed. In survey reality, however, there are always sample units that can either not be reached or that are reached but not interviewed. If the data missing due to nonresponse are missing systematically, there will be nonresponse bias.

The magnitude of nonresponse bias is defined by the nonresponse rate and the respondent and non-respondent population estimates. Following Bethlehem (2002) the bias in the mean of a sample estimate is a function of the correlation (σ) of a survey variable y with response propensity ρ to be measured in the target population. As Lynn and Clarke (2002) have demonstrated that the two components of nonresponse, non-contact and refusal, have different drivers and need to be considered separately.

Nonresponse bias can be removed from estimates if additional data items are available that correlate both with the survey estimate and with the propensity to respond. The bias in an estimate can only be removed entirely if the additional items explain all of the difference between respondents and non-respondents in this estimate. But to approach this situation, the additional items will need to explain both the contact and the cooperation stages of the nonresponse process. In the following I will focus on the non-contact aspect of nonresponse only, though similar analyses could be carried out to explain cross-national differences in refusal rates.

Nonresponse bias in cross-national surveys is relevant, since the purpose of the data is to compare countries with regards to a certain estimate. If there is nonresponse bias is one of the countries or in both countries, but of different magnitude or direction, then the cross-national comparison will be biased. The bias due to non-contact in a difference in means between two countries A and B is

$$B(\bar{y}_{non-contact}) \approx \frac{\sigma_{y\rho_{contact(A)}}}{\bar{\rho}_{contact(A)}} - \frac{\sigma_{y\rho_{contact(B)}}}{\bar{\rho}_{contact(B)}}. \quad (1)$$

Analyzing the processes leading to difference in contact rates across countries can give some indication of differences in non-contact bias across countries. The paper examines the determinants of contact and the extent to which these differ across several countries participating in the ESS. The analyses decompose differences in contact rates across countries into (1) differences due to differential respondent characteristics, interviewer characteristics and fieldwork characteristics and (2) differences due to a differential impact of these characteristics on contact and cooperation propensities (i.e. the model coefficients). In a logistic regression I model the probability of contact in various countries of the ESS taking fieldwork and respondent characteristics into account. The contact rate differs across countries; as do the model coefficients. Subsequently, I carry out a decomposition analysis to disentangle the effect on contact propensity of differences across countries in fieldwork and respondent characteristics on the one hand and the effect of these characteristics on the contact propensity (coefficients) on the other hand. The findings shed light on the extent to which the manipulable aspects of the contact processes, i.e. interviewer characteristics and fieldwork procedures, are comparable across countries, whether standardized fieldwork procedures actually lead to equivalent contact rates, and more generally, whether bias associated with differential contact rates differs across countries.

2. Data

The ESS is a biennial cross-national survey of social and political attitudes across more than twenty countries in Europe. It was first fielded in 2002. In addition to the data collected in the main interview the ESS also collects information on the contacting and cooperation process and on the neighborhood of the selected sample unit. By means of standardized comparative contact forms the interviewer collects the contact history and neighborhood information of each sample unit, whether interviewed or not. These contact and neighborhood data of the first round of the ESS are the main data source of the subsequent analyses. Analyzing these data the paper looks at the correlates of differences in contact rates across countries. The first set of models analyze contact rates in seven ESS countries, where contact and neighborhood data of sufficient quality were available and predictor variables exhibited sufficient variation: Belgium, Finland, Greece, Ireland, Portugal, Spain and the UK. In addition to the standard ESS data some countries made data on interviewer demographics available. For a subset of countries (Belgium, Finland and the UK) these interviewer characteristics are included in a second set of models.

3. Method

In analyzing differences in contact propensity across countries I distinguish between factors that are simply different across countries (i.e. that are not manipulable) and those that can be influenced (i.e. that are manipulable). This serves to find out whether standardized optimal fieldwork would lead to comparable contact rates across countries. Following the state-of-the-art literature on the determinants of contact (c.f. Groves and Couper, 1998; Campanelli et al., 1999; Purdon et al., 1999; de Leeuw and de Heer, 2002; Lynn and Clarke, 2002; Nicoletti and Buck, 2004; Stoop, 2005)the variables of manipulable and non-manipulable characteristics listed in Table 1 were included in the general models of contact in the seven ESS countries .

Table 1: Covariates in the general model of contact

	Description (Variable name)
Manipulable factors	• Whether contact was ever attempted on a weekday evening (everweve), on a Saturday (eversat) and on a Sunday (eversun) • Natural log of the number of in-person contact attempts to contact (lninperson) • Interaction terms: everweve * lninperson (wevelninp), eversat * lninperson (satlninp), eversun * lninperson (sunlninp) • Workload: number of sample units the interviewer worked on (nworked) • Interviewer ability: the ESS interviewer cooperation rate in % (coopr) • Interviewer calling strategy: the percentage of call attempts the interviewer made on a weekday evening (intpcweve), on a Saturday (intpcsat) and on a Sunday (intpcsun) • The percentage of sample units that the interviewer attempted by phone (intpctel)
Non-manipulable factors	• Whether housing unit was a farm or single-unit housing (farmsingle); omitted category is multiunit housing • Whether there was any intercom a the housing unit (inter) • Whether there were any security features at the housing unit (security) • Physical state of buildings in the area: satisfactory state (physcsat) or bad state (physcbad); omitted category is good state • State of the sampled housing unit compared to other housing units in the area: better • (phcombet) or worse (phcomwor); omitted category is same • Urbanicity: percentage of single housing units (pcsingle) and percentage of farms (pcfarm) in the assignment of the interviewer making the first contact attempt

Note: Interviewer-related variables: unless stated otherwise, where more than one interviewer worked on a sample unit, the interviewer who made who made contact or, if no contact was made, the last interviewer is considered

3.1 Logit models

In Table 2 the logit coefficients of eight models of contact are displayed. The first model is a pooled model across the seven ESS countries (Belgium, Finland, Greece, Ireland, Portugal, Spain and the UK). Columns two to eight display the results of models of contact within the respective countries.

Examining the logit models one notices several issues that are central to the purpose of this paper. First of all, the coefficients for the same variables can differ substantially across countries. This means that the effect which a characteristic has on the probability of contact can differ across countries. Second, some of the variables significantly predict contact in several countries. And third, more of the manipulable than of the non-manipulable factors significantly predict the probability of contact in the models. This might well be due to the non-manipulable factors being worse measurements of the determinants of contact. However, this is not too serious a problem for the research at hand, since the aim of the paper is to find out what influence the manipulable factors have on differences in contact rates across countries. Furthermore, a likelihood ratio test of a reduced model with only manipulable factors versus the full model shows that the manipulable factors as a whole significantly improve the model.

Consequently, this first look at the logit models of contact shows that the *effects* of characteristics on contact propensity (i.e. the coefficients) vary across countries. However, also the *distributions* of the characteristics can vary across countries. The interpretation of the logit models is limited since they only show differences in the processes correlated with contact but not in the characteristics. Finding out whether differences in contact rates are due to differences in characteristics or due to differences in coefficients is the aim of the decomposition analysis following below.

Table 2: Logit models of contact

	Pool	Belgium	Finland	Greece	Ireland	Portugal	Spain	UK
everweve	0.073	0.600	0.336	-1.173	0.384	-5.124 ***	-0.374	1.486
eversat	-0.196	-0.019	-0.999	-2.882 *	-0.108	-0.742	-0.010	5.341 ***
eversun	-0.374	0.596	1.317	-3.055 *	-1.442 *	-5.312 ***	0.881	-0.494
lninperson	-1.715 ***	-1.631 ***	-1.388 **	-4.960 ***	-2.470 ***	-6.689 ***	-1.157 ***	-0.644 **
wevelninp	-0.431 ***	-1.002 **	-1.894 **	0.340	-0.460	3.115 ***	-0.427	-1.311 **
satlninp	0.028	-0.115	-0.272	1.694	0.220	0.488	-1.156 **	-3.017 ***
sunlninp	-0.023	-0.790	-2.101	1.497	1.340 *	3.049 **	-1.536 **	-0.085
nworked	0.004 *	-0.021 *	-0.005	0.000	0.011	0.007	0.003	0.002
coopr	0.011 ***	0.036 ***	0.056 ***	0.017 *	0.007	-0.018	-0.005	0.007
intpcweve	0.009 **	0.005	0.014	-0.006	0.017 *	-0.042 *	0.010	0.033 **
intpcsat	0.009 **	0.000	-0.025	0.000	0.020 *	-0.017	0.005	-0.009
intpcsun	0.024 ***	-0.012	0.070	-0.010	0.015	-0.033	0.021	0.009
intpctel	-0.011 ***	0.009	-0.044 *	-0.093	0.020	0.003	0.041	-0.035
farmsingle	0.062	-0.172	0.481	-0.536	-0.268	0.517	0.208	0.228
inter	-0.179	-0.234	0.263	0.495	-0.056	-0.593	0.443	-0.377
security	0.156	-0.231	0.682	-0.197	0.329	-0.324	0.278	0.350
physsat	-0.010	-0.206	-0.111	-0.088	0.016	-0.295	-0.057	0.004
physbad	-0.423 **	-0.103	-1.312 *	0.592	-0.693	0.207	-0.120	-1.113 *
phcombet	0.434 ***	1.300 ***	-0.234	1.041 *	0.969 ***	0.187	0.089	0.161
phcomwor	-0.014	-0.096	0.091	0.026	0.512	-0.961	-0.521	0.524
pcsingle	0.000	0.002	0.013	-0.017 **	0.021 *	-0.018 **	-0.002	-0.008
pcfarm	-0.015 ***	0.061	-0.016	-0.072 *	0.030 **	0.004	-0.005	0.017
constant	3.385 ***	2.183 ***	2.732	10.475 ***	0.968	17.214 ***	3.771 ***	3.744 ***
Chi^2	2189	307	203	283	441	242	902	501
Pseudo R^2	0.264	0.27	0.254	0.495	0.302	0.366	0.445	0.362
N	20319	2981	2276	3203	2994	2132	3178	3555

3.2 Decomposition analysis

The general aim of a decomposition is to separate out differences in outcome that are due to differences in sample characteristics from those that are due to differences in the coefficients. In the current case we are interested in explaining differences in fieldwork outcomes (e.g. contact) between countries.

A fieldwork outcome is typically measured in terms of a binary variables, where in the case at hand 0 is non-contact and 1 is contact. The difference in contact rate between countries A and B is

$$\overline{P}(y_A) - \overline{P}(y_B) = \overline{P}(\beta_A, X_A) - \overline{P}(\beta_B, X_B) \text{ , where} \quad (2)$$

$\overline{P}(\beta_i, X_i)$ is the contact rate in country i, which is equivalent to the probability of contact associated with the characteristics of country i evaluated at the coefficients of country i.

3.2.1 Aggregate decomposition

The aim of the aggregate decomposition is to find out what the contact rate in country A would have been if it had had the *characteristics* of country B; similarly what the contact rate in country A would have been if it had the *coefficients* of country B. This can be calculated by expanding equation (4) above and rearranging the addends (c.f. Even and Macpherson, 1993).

$$\overline{P}(y_A) - \overline{P}(y_B) = \underbrace{\sum \beta_B \left(\overline{X}_A - \overline{X}_B\right)}_{\Delta \text{ characteristics}} + \underbrace{\sum (\beta_A - \beta_B)\overline{X}_B}_{\Delta \text{ coefficients}} + \underbrace{\sum (\beta_A - \beta_B)\left(\overline{X}_A - \overline{X}_B\right)}_{\Delta \text{ `interaction'}} \tag{3}$$

The difference in the mean predicted contact rate between countries A and B is therefore composed of three parts. The first addend describes the difference in predicted probabilities arising from a change in sample characteristics assuming that the coefficients remain unchanged (characteristics effect). The second addend describes the difference in predicted probabilities arising from a change in coefficients assuming that the characteristics remain unchanged (coefficients effect). The third addend reflects the extent to which these two assumptions are simultaneously broken; i.e. that some variables have both different coefficients and different distributions. This is in a way similar to an interaction effect in regression analyses. A positive value in the third addend would mean that there are characteristics in country A that are more prevalent (i.e. they have a higher mean) and at the same time have a stronger positive effect on contact propensity compared to country B. The characteristics effect indicates how many percentage points the contact rate in country B would have been higher (lower) if country B had had the manipulable and non-manipulable characteristics of country A. Similarly the coefficients effect indicates how many percentage points the contact rate in country B would have been higher (lower) if country B had had the coefficients of country A.

3.1.2 Returns from a pooled model

The comparison of two subgroups as described in the aggregate decomposition above is the standard procedure. This standard decomposition makes sense when comparing subgroups that are inherently bi-modal, e.g. men and women. However, in the case of comparing contact rates across countries, the situation is not bi-modal. Though one could compare pairs of countries with each other, this is rather tedious and unintuitive when more than three countries are involved. Comparing seven ESS countries with each other (plus the comparisons with the opposite reference group) yields 42 comparisons, with each a characteristics effect, a coefficients effect and an 'interaction' effect to be examined. In addition to the sheer magnitude of results that need to be interpreted, one would also need to make sense of the possibly different results of comparing Belgium to Finland, to Greece, to Ireland, to Portugal, to Spain and to the UK. Therefore, instead of 42 comparisons of groups, I decided to compare each country with the pool of all seven countries. This means that I compared the characteristics and coefficients of the model of each country with the characteristics and coefficients of the pooled model, where in each comparison the pooled model was the reference.

3.1.3 Detailed decomposition

We can also further decompose the characteristics effect by looking at the effect of individual characteristics (as a part of the whole characteristics effect). In linear models this is unproblematic. However, in nonlinear models, like these logit models of contact, the contribution of each variable is not linear, i.e. their contribution depends on which point they are examined at and what the values of all the other variables in the model are. While there are different – equally imperfect – ways of looking at the contribution of separate characteristics, the by now most common and simplest is to evaluate the contribution of each characteristic at their mean and weight the characteristics effect by the contribution of each variable at their mean. The proportion of the characteristics effect due to the r_{th} variable then is (Yun 2005):

$$\sum \beta_{Country}\left(\overline{X} - \overline{X}_{Country}\right) = \sum \beta_{Country}\left(\overline{X} - \overline{X}_{Country}\right)\frac{\beta_{rCountry}\left(\overline{X}_r - \overline{X}_{rCountry}\right)}{\beta_{Country}\left(\overline{X} - \overline{X}_{Country}\right)} \tag{4}$$

The contribution of the dummies of a categorical variable needs to be examined together, because their contribution will depend on the choice of reference group. The sum of the characteristics effects of the categories (dummies), however, is invariant to the choice of reference group.

3.1.4 Interpreting the decompositions

The results from the decompositions can yield new insights into the correlates of differences in contact rates across countries. From a large characteristics effect and its detailed decomposition one can find out which characteristics contribute to the difference in contact rates in the two models that are decomposed. If these are manipulable characteristics they can be optimized to improve the contact rate.

However, large coefficients and interaction effects are more difficult to interpret. A coefficients effect indicates that the characteristics included in the model may have a different effect across countries on contact propensity. Alternatively, the coefficients effect can capture that the characteristics included in the model do not explain the difference in contact rate, i.e. at least one relevant variable was omitted from the model. In either case, the manipulable characteristics that were measured then do not explain the difference in contact rate and adjusting them to improve the contact rate is not likely to have much of an effect.

For the interpretation of the 'interaction' effect recall that this reflects the extent to which some variables have both different coefficients and different distributions. It thereby also reflects the extent to which the assumptions of the characteristics effect (i.e. the change in characteristics assuming that the coefficients remain unchanged) and the assumptions of the coefficients effect (i.e. the change in coefficients assuming that the characteristics remain unchanged) are simultaneously broken. This means that decompositions where there is a sizable interaction effect should be interpreted with caution. The main characteristics and coefficients effects can then only be interpreted when simultaneously considering the 'interaction' effect. This problem can only be solved with further analyses, for example by identifying the variables that cause the 'interaction' effect. Excluding them and re-running the model, the effects of the remaining variables could then be interpreted.

5. Results of the general model

The results from the decompositions between the pooled model and each of the seven countries are listed in Table 3 below. The first part of the table displays the predicted probabilities for the pooled model (row 1), the country models (row 2) and the predicted probabilities from evaluating the characteristics of the country at the coefficients of the pool (row 3) and from evaluating the characteristics of the pool at the coefficients of the country (row 4). The fifth row displays the difference in predicted probabilities (i.e. contact rates) between the pooled model (i.e. across the seven ESS countries) and the model in each country. This is the difference that the decompositions aim to disentangle. Below this, the aggregate decomposition is shown. The measurement unit here is the percentage points of the total difference in contact rate explained by the composite parts. Finally, the detailed decomposition displays the percentage points of the characteristics effect explained by each characteristic.

Table 3: Results from the aggregate and detailed decompositions $\overline{P}(y) - \overline{P}(y_{Country})$

	Belgium	**Finland**	**Greece**	**Ireland**	**Portugal**	**Spain**	**UK**
Predicted probabilities	%	%	%	%	%	%	%
Pr(X, β)	94.8	94.8	94.8	94.8	94.8	94.8	94.8
$Pr(X_{Cntry}, \beta_{Cntry})$	95.3	95.7	98.2	93.4	96.4	90.3	95.1
$Pr(X_{Cntry}, \beta)$	96.9	95.9	96.8	96.1	93.7	93.2	91.7
$Pr(X, \beta_{Cntry})$	88.9	95.1	92.0	89.4	98.7	91.2	95.7
$Pr(X, \beta) – Pr(X_{Cntry}, \beta_{Cntry})$	-0.45	-0.92	-3.40	1.46	-1.57	4.54	-0.32
Aggregate decomposition	% points	% points	% points	% points	% points	% points	% points
Characteristics	-6.37	-0.69	-6.18	-4.00	2.30	0.93	0.57
Coefficients	1.64	0.21	-1.45	2.71	-2.72	2.94	-3.49
'interaction'	4.29	-0.43	4.22	2.75	-1.15	0.67	2.60
Detailed decomposition (characteristics)	% points	% points	% points	% points	% points	% points	% points
timing of calls * number calls	-4.64	0.35	0.24	-0.21	0.75	0.09	0.70
nworked	-2.67	0.02	0.05	0.00	0.14	0.03	0.01
coopr	1.37	0.17	0.04	-0.14	0.15	0.04	0.17
calling strategy	-0.95	-0.05	0.27	0.21	0.20	0.01	0.27
intpctel	0.35	-1.10	-2.70	-4.32	2.01	0.82	-0.65
farmsingle	0.34	-0.01	0.03	0.03	-0.04	0.05	-0.06
inter	-0.17	-0.02	0.01	0.12	0.02	-0.04	-0.11
anysecurity	-0.21	-0.03	0.00	0.00	-0.01	0.00	-0.04
state of building	-0.18	0.02	-0.06	-0.10	-0.01	0.00	-0.04
comparative state of building	-0.08	0.00	0.00	-0.03	0.00	0.00	0.02
urbanicity	0.46	-0.04	-4.06	0.45	-0.91	-0.06	0.29

5.1 Finland

Understanding the interpretation of the decomposition is easiest when looking at the Finnish decomposition. The Finnish contact rate is 0.92 percentage points higher than the average across the seven ESS countries. Looking at the aggregate decomposition we see that 0.69 of these percentage points are explained by a difference in characteristics. (The coefficients and interaction effects are relatively low in the Finish decomposition.) This means that, if Finland had had average ESS characteristics instead of the Finnish characteristics, its contact rate would have been 0.69 percentage points lower. The detailed decomposition of this characteristics effect further shows that most of the characteristics effect is explained by the percentage of cases that an interviewer contacted by phone.

5.2 Belgium

Belgium is another case where the coefficients and 'interaction' effects are relatively small and the main effects can be interpreted. The contact rate in Belgium is only slightly (0.45 percentage points) higher than the average across the seven ESS countries. Nevertheless, the aggregate decomposition shows that there are quite sizable composite effects. The characteristics effect reveals that if Belgium had had average ESS characteristics, its contact rate would have been 6.37 percentage points lower. The detailed decomposition of this characteristics effect further shows that the distribution of two manipulable fieldwork factors, the interaction of timing of calls and number of calls that an interviewer made (4.64 percentage points) and the interviewer cooperation rate (2.67 percentage points), explain a large part of the characteristics effect.

5.3 Greece

In Greece the contact rate is 3.40 percentage points above the seven-country average and has a sizable characteristics effect of 6.18 percentage points. The detailed decomposition shows that this characteristics effect is mainly due to the effect of urbanicity (4.06 percentage points) and the percentage of cases that an interviewer attempted to contact by phone (2.70 percentage points).

5.4 Ireland, Portugal, Spain and UK

The decompositions of the Irish, Portuguese, Spanish and UK contact rates are unfortunately inconclusive. These four countries all have very large coefficients and/or 'interaction' effects.

Future research will look into how to best disentangle the 'interaction' and the main characteristics and coefficients effect. The aim of this research will be to find out exactly which variables cause the 'interaction' effect and to find a way to determine the contribution of separate variables to the main characteristics effect. This is necessary to identify manipulable factors that can be adjusted for fieldwork optimization and that drive differences in nonresponse across countries.

6. Models with interviewer characteristics

As mentioned in the introduction, three of the countries (Belgium, Finland and UK) that were looked at in the general models above, made additional data on interviewer characteristics (the interviewer's sex, level of education and year of birth) available for analysis. These variables were included in a second set of models predicting contact (results not shown). In the decomposition, each country was again compared to the pool of countries. However, since the variables in these models are only available for Belgium, Finland and the UK, the pool that these countries were compared to changed into a pool consisting of these three countries.

Including the interviewer variables in the logit models showed that two of the interviewer characteristics (education and age) have a significant effect on contact in Belgium but not in the other countries. The more educated a Belgian interviewer is, the more likely is he/she to make contact. And the older an interviewer is, the more likely is he/she to make contact.

There is also an effect of the interviewer variables on the decompositions. In all three countries the coefficients effect decreased. In a decomposition any effect of omitted variables and misspecification is captured in the coefficients effect (and as a consequence also in the 'interaction' effect). The fact that adding the interviewer characteristics decreases the coefficients effect is an indication that these interviewer characteristics capture, at least to some extend, some of the variation that we were previously missing.

On the whole, the pattern of the aggregate and detailed decomposition from the general model is repeated in the model with interviewer characteristics. There is still a large positive characteristics effect indicating that had the UK had average characteristics their contact rate would have been lower. Similarly, the interaction effect in Belgium and Finland remains too large for detailed interpretation.

7. Conclusion

The paper presented a method of analysing cross-national differences in non-response that thus far had not been applied to nonresponse research. It has demonstrated that these decompositions can be a valuable method, especially where the comparison of only a small number of countries is concerned. However, its limitations have also become apparent. The decompositions showed that all the countries analyzed had quite sizable coefficients and 'interaction' effects. These were somewhat reduced when more interviewer characteristics were taken into account. Nevertheless, the coefficient effects show that the effectiveness of similar fieldwork strategies regarding success in achieving contact differs across countries; at least considering the fieldwork strategies that were measured and I was able to include in the models.

Future research will need to improve the logit models to receive more informative decompositions of the correlates of differences in contact rates across countries. Furthermore, analyses will need to be expanded to explaining differences in cooperation rates across countries. Some fieldwork procedures that lead to high contact rates, might lead to low cooperation rates. Therefore, the two need to be examined in combination to draw further inference on the effect of manipulable factors on cross-national differences in nonresponse.

Acknowledgements

Part of this paper is based on work carried out during Annelies Blom’s visit to the European Centre for Analysis in the Social Sciences (ECASS) at the Institute for Social and Economic Research, University of Essex, supported by the Access to Research Infrastructure action under the EU Improving Human Potential Programme. She wishes to thank Annette Jäckle for her contribution to discovering non-linear decompositions for survey methodological research.

References

Bethlehem, Jelke G. (2002). *Weighting nonresponse adjustments based on auxiliary information* in Groves, R.M., D. Dillman, J.L. Eltinge, and R.J.A. Little (eds.). *Survey Nonresponse*. New York: Wiley & Sons.

Campanelli, P., P. Sturgis and S. Purdon (1997). *Can you hear me knocking? An investigation into the impact of interviewers on survey response rates*. London: National Centre for Social Research.

Lynn, P. and P. Clarke (2002). Separating refusal bias and non-contact bias: evidence from UK national surveys. *The Statistician*, 51(3), pp.319-33.

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

Even, W.E. and D.A. Macpherson (1993). The decline of private-sector unionism and the gender wage gap, *Journal of Human Resources*, 28, pp.279-296

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

Nicoletti, C. and N.H. Buck (2004). E*xplaining interviewee contact and co-operation in the British and German Household Panels*. In M. Ehling and U. Rendtel (eds.), *Harmonisation of panel surveys and data quality*. Wiesbaden: Federal Statistical Office.

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

Stoop, I.A.L. (2005). *The Hunt for the Last Respondent*. The Hague: SCP.

Yun, M.S. (2005). A Simple Solution to the Identification Problem in Detailed Wage Decompositions. *Economic Inquiry*, 43, pp.766–772