Journal of Official Statistics, Vol. 25, No. 3, 2009, pp. 323–338

Cooperation in Centralised CATI Household Panel Surveys – A Contact-based Multilevel Analysis to Examine Interviewer, Respondent, and Fieldwork Process Effects

Oliver Lipps[1]

In this research, we analyse the contact-specific mean of the final cooperation probability, distinguishing on the one hand between contacts with household reference persons and with other eligible household members, and on the other hand between first and later contacts. Data comes from two Swiss Household Panel surveys.

The interviewer-specific variance is higher for first contacts, especially in the case of the reference person. For later contacts with the reference person, the contact-specific variance dominates. This means that interaction effects and situational factors are decisive. The contact number has negative effects on the performance of contacts with the reference person, positive in the case of other persons. Also time elapsed since the previous contact has negative effects in the case of reference persons. The result of the previous contact has strong effects, especially in the case of the reference person. These findings call for a quick completion of the household grid questionnaire, assigning the best interviewers to conducting the first contact.

While obtaining refusals has negative effects, obtaining other contact results has only weak effects on the interviewer's next contact outcome. Using the same interviewer for subsequent contacts has no positive effects.

Key words: Cross-classified; call data; random interviewer-respondent assignment.

1. Introduction

In centralised telephone surveys, it is usually difficult to measure cooperation effects of respondents and interviewers because the same interviewer typically does not conduct all calls with a sample member (Lipps 2008). In the Swiss Household Panel (SHP), for instance, the interviewer might call the telephone number of a household that other interviewers have already contacted. It may be the case that a fixed appointment was agreed on, or the reference person showed some reluctance during a former contact but agreed to be called later. Thus many interviewers are possibly involved in the completion of a single household grid or a single individual questionnaire. The choice of the telephone number is performed completely at random from the pool of still uncompleted numbers at a given time (interviewer shift). This assignment allows for the separation of the effects of interviewers, respondents, and contacts, on contact outcomes in a randomised setting, thus effectively achieving an interpenetrated design.

A schematic relationship of this random assignment is depicted in Figure 1.

[1] Swiss Foundation for Research in Social Sciences (FORS), c/o University of Lausanne, Vidy, CH – 1015 Lausanne, Switzerland. Email: oliver.lipps@fors.unil.ch

Acknowledgments: Support by the Swiss National Science Foundation is gratefully acknowledged. I thank four anonymous reviewers for valuable comments on an earlier version of the article.

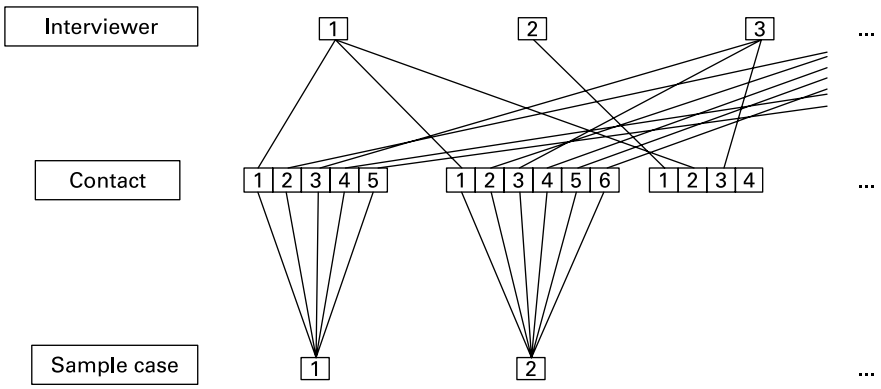

Fig. 1. Interviewer-Sample Case Assignments via Contacts in Surveys with Random Assignment

The primary aim of the current article is to promote a better understanding of different effects of the actors[2] involved in the response cooperation process in a centralised CATI panel survey. The actors are the interviewers conducting the contacts, the household reference persons asked to complete the household grid questionnaire, and "other persons"[3] asked to complete their individual questionnaires. To analyse and disentangle interviewer and respondent effects, we have to investigate the contact level. Furthermore, we seek to identify covariates which are able to explain the variance on the appropriate level. This might shed light on appropriate measures to be taken in order to improve the calling procedure.

The article is organised as follows. First, we present an appropriate model of survey cooperation, and previous findings of interviewer, respondent and fieldwork characteristic effects. Next, we introduce the data and the modelling approach used, before discussing the model results. We conclude with recommendations for fieldwork organisation, namely how to assign interviewers to contacts in an efficient way.

2. Models of Cooperation

While the respondent socio-demography is significant for the outcome of the first contact with an interviewer, Groves and Couper show that it loses its predictive power for those requiring more than one contact to obtain a final disposition. The reasons are twofold: first, socio-demographic variables with higher cooperation propensities fall out of the sample due to the omission of first-contact respondents; second, for later contacts it is rather the attributes of the prior contacts which are important indicators of the cooperation likelihood (1996, p. 74). Important to note is that the socio-demographics are "fallible: they are correlates, not causes of the survey participatory behaviour" (p. 81). This is also emphasised by Stoop (2005), who specifies causes of (non)cooperation: "social isolation, social participation, . . . , interest in societal well-being, doing voluntary work, political interest and knowledge, . . . , electoral participation, the type of sponsor, and attitudes towards surveys" (p. 126).

[2] In accordance with the language of multilevel modelling we will subsequently talk about levels when appropriate.
[3] In the sequel, "other persons" are taken to be interview-eligible household members other than reference persons.

For interviewers, as it is likely that "most of the acculturation process of producing effective interviewers occurs during training and on the job" (Groves and Couper 1998, p. 195), training, and experience seem to be important characteristics (Snijkers, Hox, and de Leeuw 1999; Hox and de Leeuw 2002). Inteviewers' attitudes towards the importance of their work and their expectations regarding difficulty in gaining cooperation seem to be significantly related to response rates (Singer, Frankel, and Glassman 1983; Lehtonen 1996). In an interviewer survey, Groves and Couper (1998, p. 209) find that positive interviewer expectations are associated with higher response rates. Also greater perceived authority and legitimacy of the sponsor might play a role (p. 206 and Cialdini 1984; Groves et al. 1992).

Groves and Couper built a theory of cooperation behaviour which heavily relies on the characteristics of the interaction between respondent and interviewer (1996, 1998). It is not so much fixed interviewer characteristics which determine the outcome of contacts. Stoop, reviewing the determinants, states that the interaction "depends on survey characteristics and fieldwork design, the social environment, characteristics of the interviewer and individual and household characteristics" (2005, pp. 55 f.). Groves and Couper's concept of "maintaining interaction" (1998, pp. 37 ff.) is based on the strategy of avoiding a termination of the interaction during initial contacts. This concept includes the ability of "stepping back" (e.g., Hox, De Leeuw, and Snijkers 1998, p. 174) as one possible interviewer tactic to adequately react to initially reluctant individuals.

3. Previous Findings

Respondent effects on cooperation in panel surveys are analysed mostly for CAPI surveys. For example, while experiencing a "pleasant" interview during the first panel wave seems to be an important factor with regard to continuation (Loosveldt, Pickery, and Billiet 2002), respondents who complain about survey burden or express an intention to quit the survey are, in fact, more likely to drop out (Martin, Abreau, and Winters 2001). Campanelli and O'Muircheartaigh (2002) find that the interviewer's subjective rating of the respondent's cooperation in the previous wave is a good predictor for nonresponse in a future wave. Nonresponse in the following wave can be predicted with the aid of standard socio-demographic variables collected in a former wave, plus political interest and social participation measures in order to include motivational factors (e.g., Pickery, Loosveldt, and Carton 2001).

Also interviewers are an important determinant of cooperation. In face-to-face panel surveys, in order to build up confidence and trust to reduce attrition, often the same interviewer is used for the same household over many years (Schräpler 2001; Campanelli and O'Muircheartaigh 1999). Positive interviewer continuity effects are reported by Buck et al. (2003), using the German Socio-Economic Panel (SOEP) and the British Household Panel Survey (BHPS), both conducted face-to-face. However, it is not clear whether confidence between interviewer and respondent is improved because the interviewer visits the same households year after year, or because the interviewer revisits just those households with whom confidence could be successfully established. The latter seems to be the case in the U.S. Health and Retirement Survey (HRS[4]), where the field

[4] The HRS samples only individuals 50 or more years old. In the HRS the first interview is done face-to-face; in the biennial follow-ups, mostly the telephone is used.

supervisors reassign interviewers to successfully interviewed households (Hill and Willis 2001). In order to analyse the trust hypothesis and to disentangle area from interviewer effects in a CAPI survey, an interpenetrated[5] design was used in a subsample of the second wave of the British Household Panel Survey by Campanelli and O'Muircheartaigh (1999). They find considerable area and interviewer random effects, without significance of the easily measurable interviewer characteristics of the conventional sex-age type. More interestingly, "there was significant variation in the effectiveness of an interviewer continuity strategy among individual level refusals. This variability, however, could not be explained by the measurable characteristics of individuals, households, or areas [or interviewers]" (p. 73). They conclude that "interviewer continuity per se does not affect response rates directly" (Campanelli and O'Muircheartaigh 2002, p. 143). The remainder of the BHPS nevertheless shows interviewer continuity effects, which suggests that without experimental control one could come to the wrong conclusion.

As to the magnitude of interviewer effects in cross-sectional surveys, Hox, de Leeuw and Kreft (1991) calculate $\rho = .02$ with an insignificant variance component for interviewers in a mixed mode (telephone and face-to-face) small controlled field experiment. None of the interviewer variables are significant. The authors admit that the interviewer sample was perhaps simply too homogeneous, at least after receiving thorough training and using a detailed script to persuade respondents to participate. Pickery, Loosveldt, and Carton (2001), using the second wave of the face-to-face Belgian Election panel survey, find an interviewer intraclass correlation coefficient ρ of $\rho = .044$ (p. 517, Table 3). Surprisingly, the effect of the interviewer from the first wave on the refusals in the second wave is stronger than the effect of the interviewer who had to actually convince the respondent to cooperate. None of the interviewer socio-demographic or experience variables are significant. Pickery, Loosveldt, and Carton (2001) wonder whether the interviewer variability is in reality geographical variability. However, similar to O'Muircheartaigh and Campanelli (1999), they find only a small geographical variability. Japec (2005) reports an interviewer intraclass correlation coefficient of $\rho = .027$ in the Swedish part of the 2002 face-to-face European Social Survey (ESS) for the response rate, and $\rho = .048$ for the refusal rate. She does not find a positive relationship between interviewer experience and response rates.

Although the number of possible stimuli in telephone surveys is smaller than in face-to-face interviews, interviewer effects can still be expected. For example, voice characteristics and speech patterns seem to play a role (Oksenberg and Cannell 1988). Interviewers may not follow directions or have different argumentation skills (Stokes and Yeh 1988; Snijkers, Hox, and de Leeuw 1999), or face different kinds and magnitudes of (perceived) burden (Japec 2005). Also existing research using data from the telephone SHP confirms this: although there are considerable interviewer effects on survey cooperation, it is not possible to substantially reduce them using available interviewer socio-demographic, attitudes, or satisfaction variables (Lipps 2006).

[5] Random interviewer-respondent assignment, see Mahalanobis (1946).

4. Data

We use call[6] data from the SHP, a nationwide, yearly, centralised CATI panel survey. The SHP started in 1999 with slightly more than 5,000 randomly selected households. Every year, the household reference person is required to first complete the household roster in the grid questionnaire. After the completion of the grid questionnaire, a household-related questionnaire is to be completed. Once all individuals in the household are enumerated, each household member from the age of 14 on has to complete his/her own individual questionnaire. The SHP recruited a refreshment sample in 2004, also representative of the Swiss residential population. The same year, the first wave of the Swiss pilot of the Statistics on Income and Living Conditions in Switzerland (CH-SILC) was conducted in parallel to the SHP, by the same survey agency. The fieldwork design for both surveys was the same, the questionnaires almost identical. Half of the pilot SILC households were surveyed a second time in the subsequent year 2005. In the wave analysed here (2005) three samples are therefore available:

- the original panel members then in their seventh wave (SHP Wave 7)
- the refreshment sample members, then in their second wave (SHP Wave 2)
- the Swiss SILC sample members, then in their second wave (SILC Wave 2)

Attrition analyses regarding the SHP confirm that, similar to those with other surveys (Groves and Couper 1998; Stoop 2005), the socially isolated drop out to a greater extent (Lipps 2006, 2007). In addition there is evidence of problems in tracking households which have moved. Attrition in the SHP occurs predominantly at two stages within the household survey process:

- when the household reference person is asked to complete the household grid
- once the household grid is completed, when eligible individuals *other than* household reference persons are asked to complete their individual questionnaires

In the models we included only individuals who completed their individual questionnaires in the preceding wave (2004).

5. Modelling Approach and Variables

From the considerations above, it becomes clear that first contacts with households are different from later contacts. First and later contacts therefore need separate analyses. We suspect different effects from the survey, socio-demography and attitudes, and the previous call history for the reference person and other persons. As a result, we distinguish between contacts with the household reference person and other interview-eligible persons.

As dependent variable, we use the mean final cooperation probability, distinguished by reference persons and others, and for the first and the later contacts, respectively. We distinguish the following contact results (Table 1):

[6] The term call will be used for any contact attempt, whether someone was contacted or not (Stoop 2005, p. 139).

Table 1. Considered outcomes of the contacts

Refusal (incl. broken appointment)
Vague appointment
Fixed appointment made by another person
Fixed appointment
Completed Interview

In the lower right part of Table 2 in the appendix, the mean probabilities of a final cooperation depending on the contact result are depicted. For example, any "vague appointment" with a reference person made in the first contact leads to a final grid questionnaire completion of the current case with a probability of 71%. For other persons, also contacted the first time, the same contact outcome has an overall individual questionnaire completion probability of 81%. Because the dependent variables are probabilities we use poisson (count) models, with a log link.

Lipps (2008) uses cross-classified[7] models with a similarly defined dependent variable, but only considers contacts with reference persons and does not distinguish between first and later contacts. With respect to the fieldwork effects, he finds negative effects from later fieldwork times in the SHP. This is a typical "late case" effect.

Similar to the analysis by Pickery , Loosveldt, and Carton (2001), we first examine whether the interviewer from the 2004 Wave has an effect on the first contact result in 2005 (Table 3). Both models include first the former interviewer of the individual questionnaire (2004; upper part), then the current first contact interviewer as second level (lower part). Only the current interviewer has effects on both grid (intraclass correlation coefficient 4.7%) and other person first contact cooperation (intraclass correlation coefficient 2.0%). Contrary to the findings of Pickery, Loosveldt, and Carton (2001), the former interviewer does not have an effect on current cooperation in the telephone SHP. We thus decide to only include the current interviewer in the final first contact models. We also test the interviewer random effects against an interviewer fixed effects model, using the two interviewer experience covariates which are significant in the (respondent random effect only) model (Table 4). A Hausman test results in Prob $>$ chi$^2 = .27$; it is therefore safe to use the random effects model. Note that in Table 4 only variables which are significant in at least one model are listed. To estimate the models, we use the default setting implemented in the MLwiN software: the first-order Taylor approximated MQL method. Departing from this default caused nonconvergence of many models. Due to severe underdispersion, however, we relax the assumption of a poisson distribution by allowing for an extra-distributional parameter.

We build up the final models step by step, including covariates from different categories:

1. Variance components model: this model includes only the intercept. In order to separate interviewer and respondent and contact effects in the later contacts models, we build cross-classified multilevel models first, with the first level the contact, and the second levels the crossed respondents and interviewer (Rasbash et al. 2004; also

[7] See for an instructive introduction Fielding and Goldstein (2006).

Figure 1). We include as many covariates as interviewers, with random coefficients for each interviewer. All variances are constrained to have the same value.

In the models with the inclusion of substantive covariates, we drop the interviewer random effects and end up with standard underdispersed poisson models in the case of the first contact and with hierarchical two-level models in the case of later contacts. This simplification is due to convergence problems and to a desire not to overburden the models.

2. + Survey/Survey Phase: here we consider whether the contact is a refusal conversion[8] attempt and we add the two survey dummies (SHP Wave 7 and SILC Wave 2). The SHP Wave 7 individuals can be expected to show a stronger panel commitment, because uninterested individuals of this (original) sample may have already refused during the previous waves. The difference between the SILC and the SHP sample is that the latter know that they are subject to a longer survey duration. Moreover the sponsors of the SHP and the SILC surveys are different: the SHP is mainly funded by the Swiss National Science Foundation, the SILC is a Eurostat project, run by the Swiss Statistical Federal Office in Switzerland. Here a scientific institution is contrasted with a federal authority, with the latter supposed to exert a higher authority.
3. + Socio-Demography: here we add all relevant individual or household characteristics already shown to be significant for attrition in the SHP (Lipps 2007). The political interest score is a combination of satisfaction with various life domains, standardised with mean 0 and standard deviation 1. We include the 2004 interviewer assessment of the likelihood that the respondent will participate in future waves, from 0 = "most probably not" to 3 = "definitely"[9].
4. + Call history: here we add the contact number and the result of the previous contact (dropped in the case of a first contact). The latter variable is a dummy for "appointment fixed" (vs. vague appointment). We include information on whether the contact is (incidentally) made by the interviewer who interviewed the respondent in the previous year (in the case of the first contact) or did the previous contact (in the case of a later contact). We also include the number of calls already done on the household, and, in the case of a later contact, the number of days elapsed since the previous contact. The interviewer within-wave learning experience variables are the number of vague or fixed appointments thus far and the number of completed interviews and refusals. These are controlled for the total number of contacts of the interviewer.

6. Modelling Results

6.1. First Contact

In the first contact variance components models, around 4.7%[10] (grid) and 2.0% (other persons) of the total variance of the contact-specific completion probability stems from

[8] Apart from a selection of experienced interviewers who obtained additional training for the refusal conversion phase, no special design changes were made for the refusal conversion.
[9] This question is not asked in the SILC and was therefore imputed by the mean value from the SHP.
[10] = .020/(.020 + .408).

the interviewers. This figure is in line with the literature. The deviance statistics[11] for both the grid and the other person models strongly increase once we drop the interviewer random intercept. Even the last grid model (+ call history) has a much higher deviance statistic than the variance components model. This shows the relevance of the interviewer variation to grid cooperation.

The survey/survey phase variables have rather strong fixed effects. Not surprisingly, all respondents in the refusal conversion phase show much lower cooperation. Also as expected, seventh-wave SHP respondents exhibit a substantially better cooperation than second-wave respondents both of the SHP and especially the SILC. This is probably mainly due to a distinct scepticism towards the European Union in Switzerland; further half of the SILC households received a written questionnaire asking income details in between their first and second survey wave. Dropping these households, the first contact outcome of the SILC sample is the same as that of the SHP wave two sample (not shown).

The previous year's within-household response rate has an expected strong positive effect on the contact performance, as well as some socio-demographic respondent characteristics. Neither age, nor sex, nor language (German/French-speaking part of Switzerland) play a role. In line with the social exclusion concept, multi-adult households, those individuals with a greater political interest and higher education cooperate better. This holds, however, only for reference persons. It seems that other persons' cooperation is to some degree determined by the household reference person rather than their own characteristics. In fact, only other persons whose household grid is completed are asked to participate. A positive assessment of future participation by the 2004 interviewer has positive effects on the first contact results.

The call history only weakly affects the first contact results. Using the same interviewer as in the wave before has no effect. This could be expected as there are inconsistent effects already in face-to-face panel surveys, which offer more interviewer stimuli. More interestingly, what the interviewer has experienced before this first contact has some effects: the number of refusals already experienced has a proportionately negative influence; this effect is greater for reference persons. While the number of completed interviews has no effect for grid respondents, the number of fixed appointments seems to slightly positively affect first contact results.

6.2. *Later Contacts*

Roughly 26% (32%) of other persons (reference persons) considered here are only contacted twice; 17% (18%) have three contacts, 12% (12%) have four contacts, 8% (8%) five contacts, etc. The maximum number of contacts amounts to 178 (58). Other persons thus are contacted more often, with a mean number of 9.2 (5.5) contacts.

For the variance components, we find that in the grid model only 1% of the total variance is due to the interviewer, 21% is due to the respondent. The rest is contact-specific variance, *within* respondent and interviewer. For the other person contact

[11] The difference of the deviances of two nested models is approximately chi^2 distributed with the number of additional variables as degree of freedom. Note that the likelihood estimate is only approximative for binomial and poisson models.

performance, the corresponding model does not converge. As for the random intercepts in the subsequent models, the grid contact effect is three to four times as high as the grid respondent effect. Regarding the other persons' random effect, respondents and contacts have about the same magnitude. This shows that the interviewer-respondent interaction quality is of particular importance for the reference person in the grid completion.

Interestingly, after inclusion of the survey and the survey phase, grid respondents in the refusal conversion phase perform only slightly worse. SHP Wave 7 respondents again show better cooperation. The effects of the socio-demographic variables are similar to those of the first contact models. While political interest is now significant for other persons, the effect of the previous wave interviewer assessment is reduced. Education is no more significant for reference persons.

After inclusion of the call history, the reference person level variance decreases to almost zero. This is mostly caused by the inclusion of the previous contact result. Other persons' contact performance is only weakly affected. Using the interviewer from the previous contact has a negative effect on the current contact with other persons. We can confirm findings from Groves and Heeringa (2006), who report a negative effect from the number of prior calls, and – with respect to the grid response – especially from the time elapsed since the previous contact. Similar to the effects on the first contact result, the interviewer's negative learning experiences also apply to the later contact models: we find similar effects, though with smaller magnitudes, with respect to the grid respondent.

7. Conclusion

In order to learn more about respondent, interviewer, and fieldwork process effects on cooperation in centralised CATI household surveys, we examine the degree of cooperation on the contact level. We use data from the Swiss Household Panel (SHP) and the second wave of the Swiss part of the Statistics on Income and Living Conditions (CH-SILC) pilot study, both conducted by the same fieldwork agency. The interviewer-respondent interaction is completely at random in the surveys considered, thus achieving an effective interpenetrated design. We consider those response stages, which are most "critical" with respect to attrition: the household reference person when asked to complete the household grid questionnaire, and eligible individuals in the household "other" than the reference person when asked to complete their individual questionnaire. We distinguish between first and later contacts.

Using a multilevel modelling approach, we find that the interviewer effects are highest in the first contact models, especially when contacting the reference person. In later contacts, the interviewer share of the total variation is almost negligible. Generally, contact performance in the refusal conversion phase is much worse; however, only slightly negative for reference persons in later contacts.

Socio-demography and the last wave within-household response rate are more important for the reference persons' than for other persons' contact performance. Contrary to previous research, these variables are still important in later contacts with the result of the previous contact controlled for. In these later contacts, reference persons' performance decreases with the number of the contact, while it increases for other persons. Interestingly, the result of the previous contact is much more important for reference persons' than for other persons' cooperation.

We conclude that especially first contacts with reference persons should preferably have a favourable result (with a completion of the grid questionnaire or at least a fixed interview appointment) while with other persons the principle of maintaining interaction appears more important. It would probably be a good idea to let only the best interviewers do the first household contacts.

Using the same interviewer from previous contacts has no positive effects. The interviewer experience within the survey plays a role: while positive experiences like obtaining an interview do not improve the performance of future contacts, it worsens with negative experiences like obtaining a refusal. This "frustration" effect is especially pronounced in first contacts with reference persons. Also this speaks in favour of an assignment of the best interviewers to conduct the first contact with a household.

8. Appendix: Modelling Results

Table 2. Descriptive Statistics 2005: Households, Calls, and Contacts

Samples: SHP Wave 2, SHP Wave 7, SILC Pilot Wave 2				
Respondents who have been validly interviewed in 2004				
Number of households *called*	6,422			
Number of households *contacted*	6,343			
Number of *calls*	144,093			
Thereof:				
Phone not answered	76.0%			
Fixed appointment with person concerned	3.1%			
Vague appointment with person concerned	9.8%			
Fixed appointment made by another person	2.2%			
Refusal (incl. broken appointment)	3.0%			
Completed interview	5.5%			
Other (nonsample calls, etc.)	.4%			
First contacts	Other person	Grid		
Number	3,318	7,279	Final cooperation of respondent	
Thereof:			Other person	Grid
Refusal (incl. broken appointment)	9.9%	14.4%	0%	0%
Vague appointment	40.1%	31.2%	81%	71%
Fixed appointment made by another person	20.8%	7.3%	88%	72%
Fixed appointment	14.0%	22.0%	95%	90%
Completed interview	14.4%	25.2%	100%	100%
Later contacts:				
Number	13,475	9,876	Final cooperation of respondent	
Thereof:			Other person	Grid
Refusal (incl. broken appointment)	6.3%	21.6%	0%	0%
Vague appointment	47.3%	41.1%	74%	59%
Fixed appointment made by another person	7.9%	8.3%	81%	64%
Fixed appointment	13.0%	7.2%	92%	84%
Completed interview	25.6%	21.8%	100%	100%

Table 3. Completion Probability, Poisson Regressions (log link) with Underdispersion, first contact with Respondent (Grid or other Individual), Interviewer: second level, Respondent: first level

Grid questionnaire	Coefficient	Other persons	Coefficient
N (first contacts)	7,279	N (first contacts)	3,318
N (Interviewers 2004)	162	N (Interviewers 2004)	155
Fixed effects		Fixed effects	
Intercept	**−.320**	Intercept	**−.234**
Interviewer 2004 intraclass corr. Coeff. (rho)	.011	Interviewer 2004 intraclass corr. Coeff. (rho)	.000
N (Interviewers 2005)	152	N (Interviewers 2005)	153
Fixed effects		Fixed effects	
Intercept	**−.321**	Intercept	**−.234**
Interviewer 2005 intraclass corr. Coeff. (rho)	.047	Interviewer 2005 intraclass corr. Coeff. (rho)	.020

Table 4. Completion Probability, Poisson Regressions (log link) with Underdispersion. All Coefficients "significant" (at least twice their Standard Error). Bold: at least 10x their s.e., –: not applicable/not considered. Not listed: not significant. In brackets: not significant. Deviance statistics for poisson models are approximative

1st contact	Variance components model[b]		+Survey/survey phase		+Socio – Demography		+Call history	
N (first contacts)	Grid:7,279/Other person: 3,318							
	Grid	Other	Grid	Other	Grid	Other	Grid	Other
Intercept	**−.321**	**−.230**	**−.324**	**−.234**	**−.268**	**−.231**	**−.577**	**−.594**	**−.575**	−.513
+Survey										
Refusal conversion					**−.833**	**−.839**	**−.805**	**−.806**	**−.677**	**−.818**
SHP Wave 7					.075	.064	.069	.061	.060	.050
SHP Wave 2					Base	Base	Base	Base	Base	Base
SILC Wave 2					−.048	(−.029)	−.072	−.048	(−.028)	−.051
+Response propensity/Socio-demography										
Within HH response rate previous year							**.239**	.317	**.240**	.278
Political interest score							.013		.013	
Number of adults in household							.018		.018	
Higher education							.023		.023	
"Respondent takes part in next wave" (Iwer)							.027	.026	.027	.024
+Call history										
Same interviewer as in previous interview										
Number of calls already to household										−.001
Interviewer number of fixed appointments									.003	
Interviewer number of refusals									−.007	−.001
Random intercept Interviewer σ^2	.020	.006	–	–	–	–	–	–	–	–
Random intercept Respondent σ^2	**.408**[a]	**.290**	**.461**	**.306**	**.457**	**.303**	**.454**	**.303**	**.451**	**.299**
(Under) Dispersion factor	.124	.088	.140	.093	.139	.092	.138	.092	.137	.091
Deviance statistic	3,372	685	4,006	744	3,686	671	3,604	663	3,549	646

Table 4. Continued

Later contacts	Variance components model[b]		+Survey/survey phase		+Socio – demography		+Call history	
N (later contacts)	Grid: 13,475/Other person: 9,876							
	Grid	Other	Grid	Other	Grid	Other	Grid	Other	Grid	Other
Intercept	**−.291**	(n.c.)	**−.287**	**−.235**	**−.389**	**−.274**	**−.583**	**−.631**	**−.481**	**−.562**
+ Survey										
Refusal conversion					−.138	**−.668**	−.133	**−.613**	**−.205**	**−.606**
SHP Wave 7					**.120**	.095	**.116**	.094	.063	.078
SHP Wave 2					Base	Base	Base	Base	Base	Base
SILC Wave 2					(−.008)	.036	(−.027)	(.031)	−.022	(.031)
+Response propensity/Socio-demography										
Within HH response rate previous year							**.226**	.378	**.126**	.323
Political interest score							.034	.029	.019	.027
Higher education										
Number of adults in household							.014		.008	
"Respondent takes part in next wave" (Iwer)							.016		(.008)	
+ Call history										
Contact number 2									Base	Base
Contact number 3									(−.003)	.024
Contact number 4									−.022	.040
Contact number 5 +									−.043	.045
Status of last Contact (fixed vs. vague apptmt)									**.210**	.025
Same Interviewer as in previous contact										−.025
Number of calls already to household									−.001	−.001
Days since last contact									**−.003**	
Interviewer Number of fixed appointments										.001
Interviewer Number of refusals									−.001	−.001
Random intercept Interviewer σ^2	.003	(n.c.)	–	–	–	–	–	–	–	–

Table 4. Continued

Later contacts	Variance components model[b]		+Survey/survey phase		+Socio – demography		+Call history			
N (later contacts)	Grid: 13,475/Other person: 9,876									
	Grid	Other	Grid	Other	Grid	Other	Grid	Other	Grid	Other
Random intercept Respondent σ^2	**.051**	(n.c.)	**.054**	**.104**	**.045**	**.107**	**.041**	**.107**	**.000**	**.105**
Random intercept Contact σ^2	**.181**	(n.c.)	**.181**	**.095**	**.191**	**.099**	**.197**	**.099**	**.253**	**.105**
(Under) Dispersion factor	.055	(n.c.)	.055	.029	.058	.030	.060	.030	.077	.032
Deviance statistic	−529	(n.c.)	−383	–	−380	−3,631	−478	–	−1,330	–

[a] In binomial or poisson models the variance at the lowest level is constrained to the area under the logistic curve ($\pi^2/3 \sim 3.29$); see Snijders and Bosker (1999). Due to underdispersion, we have a variance of .408 = .124 (dispersion factor) * 3.29 (constraint).

[b] Cross-classified multilevel for first two models (not converging for other persons), multilevel hierarchical clustering structure (omitting interviewer random intercept) in subsequent models.

9. References

Buck, N., Nicoletti, C., McCulloch, A., and Burton, J. (2003). Report on Attrition Analysis and Item-Nonresponse. CHINTEX Working Paper 16, ISER, University of Essex.

Campanelli, P. and O'Muircheartaigh, C. (1999). Interviewers, Interviewer Continuity, and Panel Survey Nonresponse. Quality & Quantity, 33, 59–76.

Campanelli, P. and O'Muircheartaigh, C. (2002). The Importance of Experimental Control in Testing the Impact of Interviewer Continuity on Panel Survey Nonresponse. Quality & Quantity, 36, 129–144.

Cialdini, R.B. (1984). Influence: The New Psychology of Modern Persuasion. New York: Quill.

Fielding, A. and Goldstein, H. (2006). Cross-classified and Multiple Membership Structures in Multilevel Models: An Introduction and Review. Research Report RR791, Department for Education and Skills, University of Birmingham.

Groves, R., Cialdini, R., and Couper, M. (1992). Understanding the Decision to Participate in a Survey. Public Opinion Quarterly, 56, 475–493.

Groves, R. and Couper, M. (1996). Contact-Level Influences on Cooperation in Face-to-Face Surveys. Journal of Official Statistics, 12, 63–83.

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

Groves, R. and Heeringa, S. (2006). Responsive Design for Household Surveys: Tools for Actively Controlling Survey Errors and Costs. Journal of the Royal Statistical Society, Series A, 169, 439–458.

Hill, D. and Willis, R. (2001). Reducing Panel Attrition. A Search for Effective Policy Instruments. Journal of Human Resources, 36, 416–438.

Hox, J. and de Leeuw, E. (2002). The Influence of the Interviewer's Attitude and Behavior on Household Survey Nonresponse: An International Comparison. In Survey Nonresponse, R. Groves, D. Dillman, J. Eltinge, and R. Little (eds). New York: Wiley, 103–120.

Hox, J., de Leeuw, E., and Kreft, I. (1991). The Effect of Interviewer and Respondent Characteristics on the Quality of Survey Data: A Multilevel Model. In Measurement Errors in Surveys, P. Biemer, R. Groves, L. Lyberg, N. Mathiowetz, and S. Sudman (eds). New York: Wiley.

Hox, J., De Leeuw, E., and Snijkers, G. (1998). Fighting Nonresponse in Telephone Interviews; Successful Interviewer Tactics. ZUMA Nachrichten Spezial (August), 173–185.

Japec, L. (2005). Quality Issues in Interviewer Surveys: Some Contributions. Department of Statistics, Stockholm University. PhD Thesis.

Lehtonen, R. (1996). Interviewer Attitudes and Unit Nonresponse in Two Different Interviewing Schemes. In International Perspectives on Nonresponse. In Proceeding of the Sixth International Workshop on Household Survey Nonresponse, S. Laakonen (ed.). Helsinki: Statistics Finland, 130–140.

Lipps, O. (2006). Attrition of and within Households in the Swiss Household Panel – Roles of the Reference Person, Other Persons, and the Interviewer in the Course of Time. Discussion Paper 6-06, Swiss Household Panel, Neuchâtel.

Lipps, O. (2007). Attrition in the Swiss Household Panel. Methoden – Daten – Analysen, 1, 45–68.

Lipps, O. (2008). A Note on Interviewer Performance Measures in Centralised CATI Surveys. Survey Research Methods, 2, 61–73.

Loosveldt, G., Pickery, J., and Billiet, J. (2002). Item Nonresponse as a Predictor of Unit Nonresponse in a Panel Survey. Journal of Official Statistics, 18, 545–557.

Mahalanobis, P. (1946). Recent Experiments in Statistical Sampling in the Indian Statistical Institute. Journal of the Royal Statistical Society, 109, 325–370.

Martin, E., Abreau, D., and Winters, F. (2001). Money and Motive: Effects of Incentives on Panel Attrition in the Survey of Income and Program Participation. Journal of Official Statistics, 17, 249–266.

Oksenberg, L. and Cannell, C. (1988). Effects of Interviewer Vocal Characteristics on Nonresponse. In Telephone Survey Methodology, R. Groves, P. Biemer, L. Lyberg, J. Massey, W. Nicholls, II, and J. Waksberg (eds). New York: Wiley.

O'Muircheartaigh, C. and Campanelli, P. (1999). A Multilevel Exploration of the Role of Interviewers in Survey Non-Response. Journal of the Royal Statistical Society, Series A, 162, 437–446.

Pickery, J., Loosveldt, G., and Carton, A. (2001). The Effects of Interviewer and Respondent Characteristics on Response Behaviour in Panel Surveys. Sociological Methods & Research, 29, 509–523.

Rasbash, J., Steele, F., Browne, W., and Prosser, B. (2004). A User's Guide to MLwiN. Version 2.0. Centre for Multilevel Modelling, University of Bristol.

Schräpler J.-P. (2001). Respondent Behavior in Panel Studies. A Case Study of the German Socio-Economic Panel (GSOEP). DIW Discussion Papers 244, DIW Berlin.

Singer, E., Frankel, M., and Glassman, M. (1983). The Effect of Interviewer Characteristics and Expectations on Response. Public Opinion Quarterly, 47, 68–83.

Snijders, T. and Bosker, R. (1999). Multilevel Analysis. Newbury Park, California: Sage.

Snijkers, G., Hox, J., and de Leeuw, E. (1999). Interviewer's Tactics for Fighting Survey Nonresponse. Journal of Official Statistics, 15, 185–198.

Stokes, L. and Yeh, M.-Y. (1988). Searching for Causes of Interviewer Effects in Telephone Surveys. In Telephone Survey Methodology, R. Groves, P. Biemer, L. Lyberg, J. Massey, W. Nicholls II, and J. Waksberg (eds). New York: Wiley.

Stoop, I. (2005). The Hunt for the Last Respondent: Nonresponse in Sample Surveys. PhD Thesis, Utrecht University.

Received March 2007
Revised July 2009