Bulletin de méthodologie sociologique

Bulletin of sociological methodology

95 | 2007
July

Interviewer and Respondent Survey Quality Effects in a CATI Panel

Oliver Lipps

Electronic version

URL: <http://journals.openedition.org/bms/392>

ISSN: 2070-2779

Publisher

Association internationale de méthodologie sociologique

Printed version

Date of publication: 1 July 2007

Number of pages: 5-25

ISSN: 0759-1063

Electronic reference

Oliver Lipps, « Interviewer and Respondent Survey Quality Effects in a CATI Panel », *Bulletin de méthodologie sociologique* [Online], 95 | 2007, Online since 01 July 2010, connection on 01 May 2019.
URL : <http://journals.openedition.org/bms/392>

This text was automatically generated on 1 May 2019.

© BMS

Interviewer and Respondent Survey Quality Effects in a CATI Panel

Oliver Lipps

Introduction

- 1 Quite recently the *Bulletin de Méthodologie Sociologique* published an extensive literature overview on survey interviewers, with the aim to encourage research and contributions on the subject (van Meter 2005). Our paper extends existing work in at least two dimensions:

It is based on data which uses pure random interviewer-respondent assignments

It analyses interviewer effects with respect to *time* as an additional dimension, i.e., it examines several panel data waves on respondents and interviewers in a longitudinal way.

- 2 Because interviewers are usually not randomly assigned to areas in face-to-face surveys, interviewer and area effects may easily be confused in CAPI panel surveys. E.g. Schräpler (2001), analysing income nonresponse using the German Socio Economic Panel (GSOEP), applies multilevel models to separate interviewer and respondent effects, but admits that possibly interviewer and/or area effects are confused (p. 150) with the design used by the GSOEP. A non randomised interviewer-respondent assignment makes the analysis both of respondent and interviewer effects and a clear separation of them impossible. Therefore, in the second wave of the British Household Panel Study (CAPI), a subsample was surveyed using a special “interpenetrated” design (Mahalanobis 1946), in order to be able to separate interviewer and area effects with respect to noncontact and refusal (O’Muircheartaigh and Campanelli 1999), and to interviewer continuity on nonresponse (Campanelli and O’Muircheartaigh 1999, 2002). A further difficulty arises in panel surveys like the GSOEP because the same interviewer is preferably assigned to the same respondents after waves. The reason is that building up confidence between respondent and interviewer may have positive effects on data quality and response. E.g. Spiess and Kroh (2005) confirm this for the GSOEP and Laurie et al. (1999) for the British Household

Panel Survey (BHPS), however in the Australian HILDA panel, using the same interviewer after one wave has no effects on attrition (Watson and Wooden 2004).

- 3 In centralised telephone surveys with a completely random interviewer-respondent assignment, an interpenetrated design is realised by definition. However also in centralised telephone surveys interviewers are often assigned to respondents in a non-random way. E.g. in the US Health and Retirement Study, in which most interviews are being conducted by telephone, “supervisors tend to assign the best interviewers ... to the most difficult respondents.” (Hill and Willis 2001, p. 426). In addition, often local dialects or knowledge of the area are tried to be matched by assigning interviewers, who live(d) in the district or are knowledgeable of the area, to the respective respondents.
- 4 In panel surveys, it is especially important to examine effects of experience and burden of interviewers on respondent behaviour, without the confusion of a non-random interviewer-respondent assignment. To our knowledge a separation and analysis of interviewer, respondent and wave¹ effects in a panel survey are hitherto lacking in the literature. One reason is the mentioned non randomness of the assignments, which makes appropriate (non experimental) panel surveys very rare. Another reason is the rather complex cross-classified data structure (see Figure 1) which necessitates sophisticated modelling tools.
- 5 This article seeks to add more insight in interview quality effects in a centralised CATI panel survey, over several waves. In a first step we analyse one panel cross section during which an interviewer survey has been conducted in addition to the CATI survey. In this wave, in addition to the longitudinal respondents, a representative refreshment sample has been drawn and interviewed for the first time. Thanks to the refreshment sample, we are able to control effects from attrition and aging. We like to analyse whether the respondents who are new to the panel exhibit different survey quality behaviours, controlled for possible interviewer effects. In the longitudinal analysis step, for which no interviewer characteristics are available, we are particularly interested in learning effects of respondents and experience and survey burden effects of interviewers, respectively. Also panel habituation effects are investigated. We also examine the effect of incidentally interviewing the same respondent after one wave, and effects of incidentally matching interviewers and respondents by age, sex, and education.

Respondent and Interviewer Survey Quality Effects

- 6 Reasons for survey respondents not to give valid answers might be extremely different (Shoemaker 2002). A simple reason may be that the correct answer is not known (Frick and Grabka 2005), or an inadequate comprehension of the intent of the question or failure to retrieve adequate information (Groves et al. 2004), but also the sensitivity of questions and privacy or confidentiality reasons might play a role (Mayer 2001). The latter might even be partially dependent on field conditions like familiarity with the interviewer: Mensch and Kandel (1988) e.g. find that the number of prior interviewing contacts depresses drug use reporting. The respondent’s cognitive trade off of giving or not a valid answer may lead to a socially desired, untrue answer or to complete nonresponse. If respondents provide acceptable answers without expending the cognitive effort necessary to give the *accurate* response, this results in “satisficing” behaviour (Krosnick 1991, Pickery and Loosveldt 1998, 2001).

- 7 Individual behaviour can be explained by means of the rational choice theory, according to which respondents of a survey behave depending on the perceived consequences. The choice depends on features of the situation with regard to own preferences (Esser 1993). E.g. an individual answers a survey if “the act of participation is expected to bring rewards that exceed the cost of participation” (Hill and Willis 2001, p. 418).² A critique is that for low cost decisions like answers to a survey question no complicated cost benefit calculations are made. Rather, habits are chosen, which are automatic dispositions adopted from former behaviour and experiences, and prove successful in everyday problems (Esser 1993). A slightly more reflected behaviour is to typify the situation and to specify a *frame* which contains a dominant goal characteristic of the specific situation, learnt from previous situations. The frame thus determines the actions to take in the specific situation (Esser 1993). In an interviewer survey for example, if there are effects from specific interviewer characteristics which cause special respondent reactions, specific frames are activated by the respondents.
- 8 Interviewers are a possible source of error, too, both in face-to-face and in telephone surveys. The extreme form of interviewer effect is to fabricate interviews (Schnell 1991, Diekmann 2005). More prevalent, however, are not closely followed scripts or coding procedures. Peneff (1988), analysing 45 survey interviewers for the INSEE national French statistics institute, finds discrepancies between researcher’s norms concerning standardisation neutrality, and identical wording, and how questions are asked and responses constructed.
- 9 In short, the interviewer influence may affect both variation and localisation measures of the responses (Japac 2005, paper I, p. 3):
The interviewer can influence the respondent to pay less attention to the questions
The interviewer can influence the respondent in a certain direction.
- 10 Van der Zouwen and Dijkstra (1988) show that there are basically four types of inadequate interviewer behaviour, which cause biases:
The adaptation of questions for the interviewee by the interviewer
The interviewer’s inattentiveness
The interviewer’s choice of a response
The hinting by the interviewer to clarify the interviewee’s response (see also Smit 1993, Smit, Dijkstra and van der Zouwen 1997).
- 11 On the other hand, van der Zouwen (2006) pleads for not to “over-standardize” the interviewer behaviour, leaving the interviewer some room for “repairing” respondent inadequate answers or misunderstandings. This behaviour may cause different answers depending on the interviewer’s knowledge of concepts and definitions, which can be “in one of the following four cognitive states:
Available: the concepts and definitions can be retrieved with minimal effort
Accessible: the concepts and definitions can be retrieved with some effort
Generatable: the concepts and definitions are not exactly known, but may be guessed using other information in memory from previous surveys
Not available: the requested concepts and definitions are not known” (Japac 2006, p. 34). For example the interviewer might recall a similar problem from an earlier interview and adopt the definition taken there; others might choose to record a “don’t know” answer (Japac 2006).

- 12 Van der Zouwen and de Leeuw (1989) find in addition small but significant effects of the method of data collection on the quality of survey data. They report three “mode characteristics” or intervening variables which explain these mode effects:
The “persuasion power” of a mode
The complexity of the task for the respondent
The degree of the control over the question-answer process
- 13 Although the number and magnitude of possible stimuli in telephone surveys are smaller than in face-to-face interviews, interviewer effects still have to be expected. E.g. 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, Snijders et al. 1999), or face different kinds and magnitudes of (perceived) interview burden (Japec 2005). The latter stems from different workloads, a poorly planned survey administration, a lack of positive feedback and of clearly defined expectations, of getting many refusals, of trying to persuade reluctant respondents, of poorly designed questions, or of feeling that the survey topic is not important to society (Japec 2005, paper II, p. 15 f.). A “direct effect [on response quality] is if an interviewer finds it difficult to ask a respondent a sensitive question and chooses to tell the respondent that he or she does not have to answer the question” (p. 27).
- 14 Groves and Magilvy (1986) find that interviewer effects largely vary between surveys and try to reduce this instability by cumulating results over many surveys. For the central telephone surveys conducted by the University of Michigan Survey Research Center (SRC) they report a rather small interviewer intraclass correlation coefficient of around .01, which is smaller than those typically found in CAPI surveys (also Groves 1989, cited in Groves et al. 2004, p. 277). With respect to relevant covariates identified to vary with interviewer effects, Singer et al. (1983) report from an early study on the basis of a CATI survey that interviewer expectations, age and their assignment size have effects on the cooperation rates. Hox et al. (1991) are among the first to analyse interviewer effects using (adequate) multilevel models. From a literature review they confirm the interviewer effect magnitude reported earlier by Groves and Magilvy (1986): while “in well-conducted face to face interviews the intraclass correlation typically clusters around .02; in controlled telephone surveys it averages below .01”. Hox et al. report an intraclass correlation of .01 for item nonresponse and .06 for response bias (acquiescence) in a small experimental controlled mixed mode study. They find “very few interviewer variables that explain significant interviewer variance” (p. 457) and conclude that while the interviewer effects were generally small, they were derived from data “employing a thoroughly pilot tested questionnaire, and interviewers who were well trained ... , closely supervised, and provided with scripts for difficult situations. In large-scale surveys the field conditions may be less optimal, and differences between ... interviewers may be larger” (p. 458).
- 15 In the more recent literature, in order to analyse interviewer effects on item-nonresponse using the multilevel modelling technique, mostly CAPI data has been used. Pickery and Loosveldt (1998, 2001 and 2004), analysing CAPI data from Belgium, find “interviewer effects [on data quality measures], but were not able to explain them in terms of [available] interviewer variables” (1998, p. 43). As to the extent of the effects, they report comparatively high intraclass correlation coefficients of .21 for income (non)responses, .16 for “no opinion” responses, .18 for “don’t know” responses, and .04 for extreme response categories (2004, p. 83). Concerning dependent survey quality

variables susceptible to interviewer effects, Pickery and Loosveldt conclude that “question difficulty and scope of the interviewer task might explain the size of the interviewer effects on item nonresponse” (2001, p. 337, see also Brick et al. 1995, for CATI surveys). Schröpler (2001) analyses the German SOEP which also mostly uses CAPI techniques, separately for several waves. He reports comparatively high interviewer intraclass correlations with respect to stereotypical response styles amounting to between .29 and .41, and to even between .66 and .71 with respect to item nonresponse to gross income questions. Furthermore, he finds respondent correlations of stereotypical response styles which amount to around .15 between waves. Schröpler (2001) concludes that these ‘small’ values suggest “that this respondent behaviour is not a stable personality trait over time, but a temporary habit caused by a motivational deficit” (p. 10). As to income nonresponse, he reports higher nonresponse for both female interviewers and female respondents, and that the income refusal rate is a quadratic function of the respondent’s age in the SOEP. The between wave income nonresponse correlation is higher than .4 (p. 15).

- 16 Because in CAPI surveys the interviewer has a much stronger – physical - presence to the interviewee than in CATI, which works only through acoustic stimuli, telephone survey data should produce more conservative interviewer effects than face to face surveys (Pannekoek 1988). On the other hand, CATI survey based estimates are especially sensitive to interviewer effects because each interviewer usually performs many more interviews than in CAPI. This might dramatically increase the design effect (deff), which depends linearly on the sample assignment size, see Groves et al. (2004, p. 276). In addition, CATI interviewers tend to be a more homogenous population group as to their socio-demography and attitudes (Groves and Couper 1998, and Scherpenzeel 2005 for the Swiss Household Panel).
- 17 Heeb and Gmel 2001, using a CATI interviewed sample of 2,746 individuals conducted by 39 interviewers, analyse interviewer effects on alcohol consumption in Switzerland. They report an interviewer intraclass correlation coefficient of .023. This leads to a design effect $deff=1.89$. Andersen and Olsen (2002) analyse the Danish National Birth Cohort Study 1997-1999 which contains 12,910 CATI interviews carried out by 34 interviewers. They find little evidence for interviewer effects on answers to questions concerning smoking and alcohol consumption in the first trimester of pregnancy. Neither the interviewers' personal habits, nor their attitudes toward smoking, their alcohol consumption during pregnancy or their education, age, or parity correlate with the answers obtained. The authors admit that “training of the interviewers and continuous supervision may have contributed to this finding.” (p. 95, see also Billiet and Loosveldt (1988) as to the importance of interviewer training to difficult to administered questions). Lipps (2005) analyses various item nonresponse measures simultaneously using the fifth (2003) wave of the CATI Swiss Household Panel data. Although he finds comparatively high interviewer effects, only few interviewer fixed characteristics are significant. As a consequence no clear picture of favourable (easily measurable) interviewer characteristics can be drawn.
- 18 To summarise there seem to be interviewer effects on survey quality and item nonresponse aspects, especially in face to face surveys, in terms of difficult to ask or to code questions, and in surveys with less controlled interviewers and a laxer script handling by interviewers. In CAPI panel surveys like the GSOEP, we encounter rather high correlations of response quality measures between adjacent waves. The question remains

on the magnitude of interviewer and respondent variation in a longitudinal context in a CATI panel survey.

Data and Dependent Variables

- 19 The Swiss Household Panel (SHP) is a yearly conducted centralised CATI panel survey which started in 1999 with slightly more than 5,000 households, representative for the Swiss residential population. Questions are about household composition and socio-demographics, health, well being and attitudes, politics, social networks, and economics. Because of the survey design with a randomised interviewer-respondent assignment, we are able to disentangle interviewer, respondent and wave effects. For the longitudinal analysis, we use those respondents from the 2000 (2nd) through the 2005 (7th) wave, who are successfully interviewed throughout all six waves. These 2,733 respondents are interviewed by 237 interviewers over the six waves. The majority of the 237 interviewers involved stay in the panel for only one wave, as can be seen from Table 1.

Table 1: Swiss Household Panel Interviewer stay number of waves, between 2000 and 2005

Duration: number of waves	Number of Interviewers	Percent
1	132	55.7
2	64	27.0
3	23	9.7
4	13	5.5
5	3	1.3
6	2	.8

- 20 The interviewer-wave-respondent association can be schematised as follows in Figure 1.

Figure 1: Pattern of interviewer-wave-respondent association in the Swiss Household Panel

- 21 All 2,733 longitudinal respondents are interviewed in all six waves 2000 through 2005. In the figure, interviewer 1 interviews both in wave 2 and wave 3 (i.e. he/she belongs to the 64 interviewers who interview in two waves, see Table 1), interviewer 2 interviews only in wave 2, interviewer 3 only in wave 4, etc.
- 22 In 2004 (wave 6), the SHP interviewed a refreshment sample in addition to the “old” panel households, equally representative of the Swiss residential population. There are only slightly more “old” panel households, then in their sixth wave, than “new” panel households in 2004. After 2000 and 2003, the SHP conducted an additional interviewer survey in 2004, and - anticipating interviewer specific differences for the “old” and the “new” sample - asked the interviewers in addition about their difficulties to convince new and old respondents to participate separately. At large, the 2004 SHP contains 8067 completed individual CATI interviews, conducted by 114 interviewers. The 2004 interviewer questionnaire was filled out and sent back by 87 interviewers. After excluding the not usable questionnaires and the interviewers who conducted 5 or less individual interviews, a sample of 81 interviewers having completed 6796 individual CATI interviews remains for the cross-sectional analysis.
- 23 Of the 81 interviewers analysed, only 19 (23 %) interviewers are male; 59 (73%) have German as their mother tongue; 22 interviewers (23%) French. Three interviewers (4 %) are also able to conduct interviews in Italian. These proportions are about in line with the languages spoken by the Swiss population. 34 interviewers (41%) are still in education; accordingly, 61 (75%) are younger than 25 years. Most interviewers find it easy to ask the questions. Slightly more interviewers state higher difficulties to convince the new rather than the old households to participate. 67 (83%) of the interviewers claim that they would report their income, similarly, 45 (56%) would take part at a comparable survey.
- 24 In order to operationalise the dependent variables to examine data quality measures like giving a social desirable answer or to satisfice, we partly follow Jäckle et al. (2006), in that they also did not test the social desirability connotations of the items and use similar quality indicators of satisficing and giving socially desirable answers available in the SHP data, taking account of the Swiss context.³
- 25 In particular we use for the social desirability index four questions from the politics module of the SHP individual questionnaire. These comprise the extent to which a person

is interested in politics, participates in federal polls, whether one agrees in that women are generally penalised in society, and whether one is in favour of measures against these conditions. On a scale from 0 to 10, where a higher value means higher accordance, in order to quantify our social desirability indicator, we calculate the number of categories between 8 and 10. Because these measures can be assumed to correlate with social status, the respondent covariates shown in the modelling results in are control variables with respect to the social desirability variable, rather than substantial *explanations* of this measure. This applies particularly for the education variable.

- 26 For the satisficing index to be analysed, we use 22 satisfaction and attitude questions from the modules health, work, family and social networks, and politics. According to Krosnick (1991), we consider answering “don’t know” or “no answer” a form of satisficing, because the respondent is not motivated to expend the mental effort necessary to generate a substantive answer (see also Pickery and Loosveldt 2001). We also consider the extreme category response propensities (0 or 10) to answer to the 22 questions.⁴ The latter are known to be a strong predictor for unit-nonresponse in the next wave in the SHP (Lipps 2007). Lastly, we include income nonresponse, as a binary survey quality variable. Income nonresponse equals 1 if based on the information given by the respondent the total yearly personal income cannot be calculated (Gabadinho and Budowski 2002).
- 27 In the following Table 2, we depict the percentage of occurrence of the indicators under consideration, in the cross-sectional sample 2004 (including the refreshment sample) and the longitudinal sample, respectively.

Table 2: Probability of Occurrence of the Quality Measures Indicators considered for the cross sectional Sample in 2004 (N=6796) and the longitudinal sample in 2000-2005 (N=16,398 observations, 2,733 individuals)

Variable	Statistic	Mean			Standard Deviation		
		2004	2000-2005	2000-2005	2004	2000-2005	2000-2005
		“old”	refresh	longitud	“old”	refresh	longitud
Social Desirability		.290	.279	.386	.221	.220	.274
Noanswer		.008	.014	.009	.030	.043	.033
Extreme Categories		.210	.281	.204	.189	.206	.185
Income Nonresponse		.048	.075	.064	.213	.263	.245

- 28 It becomes clear that the members of the “old” sample in 2004 and especially the longitudinal sample answer in a more socially desired way, give fewer noanswers and extreme answer categories, and do more often report their income. In the following, we do not further consider the noanswer index, because its occurrence does not seem to be a problem.⁵

Models

- 29 As we deal here with clustered hierarchical data, multilevel methods are chosen for modelling. They are now often applied to survey data, in which respondents (1st level) are clustered within interviewers (2nd level). Unlike the purely hierarchical three level analysis e.g. applied by Schröpler (2001) with wave as lowest level, in our longitudinal

models, a cross-classified structure is necessary. This is because though respondents are clustered in interviewers and waves are clustered in respondents, respondents over waves are not clustered in interviewers. The schematic interviewer-wave-respondent association can be seen in Figure 1 above.

- 30 According to the distribution of the variables, we model social desirability and extreme category use as poisson⁶ distributed variables with a log link, and the income nonresponse variable as a logistically distributed binary variable. Due to underdispersion, we relax the standard binomial variance assumption of all indexes in all but the income nonresponse in the cross-sectional full (see below) model.
- 31 In order to construct the cross-classified longitudinal models, we first build two-level models with second level the respondents and first level the wave. In order to find the covariates of the full (intermediate) two level models, we proceed similar to Hox et al. (1991, p. 445): based on a theoretically meaningful choice (Hox et al. 1991, Pickery and Loosveldt 2001, Japac 2005), we include the significant wave dummies and select promising respondent level variables by backward OLS regression (.05 as criterion). Then, the interviewer specific variables are added and tested, again using a backward regression. The 2004 wave hierarchical cross-sectional model is built in the same way. Then these variables are entered in the two-level (interviewer-respondent in the case of the cross-sectional, respondent-wave in the case of the longitudinal) models. Unlike Hox et al., we do not test random effects of variables other than the intercept, in the two-level models, in order not to overburden the longitudinal models.
- 32 Based on the converged two-level longitudinal models, a (pseudo) third level with 1s for all observations is constructed in the longitudinal models. We define dummies for each of the interviewers, which have random coefficients on the third level. All variances on the third level are constrained to have the same value. Because of the small interviewer effects in the longitudinal variance components models, and because some of the three-level models did not converge with covariates other than the intercept, we decide not to include the interviewer level in the respective *full* longitudinal models, which contain the significant covariates.
- 33 The estimated fixed and random effects for each quality indicator for the cross-sectional and the longitudinal models are depicted in the right half of and are divided in two parts: first the variance components models, i.e. the fixed intercept and the random intercept coefficients on all levels considered (“VC”=Variance Components”) are listed. The random coefficients allow for the calculation of the variance share for each level considered. The interviewer variance proportion is depicted in the third last line of the variance components models. Secondly, the results of the “full” models which contain the fixed coefficients, together with the random parts are shown.

Table 3: Fixed and Random Effects, Respondent – 2nd level, Wave – 1st level, Interviewer – (pseudo) 3rd level

Wave	2004 cross sectional 6,796 Resp. - 31 matched Inter.						2000-2005 longitudinal 2,733 Resp. 237 Inter.					
	Social Desirability (Poisson Reg.)		Extreme Categories (Poisson Reg.)		Income Non-Response* (Logit Reg.)		Social Desirability (Poisson Reg.)		Extreme Categories (Poisson Reg.)		Income Non-Response** (Logit Reg.)	
Dep. Variable	V.C.	Full	V.C.	Full	V.C.	Full	V.C.	Full	V.C.	Full	V.C.	Full
Intercept	340	0	-1.421	-1.378	-2.822	-5.350	428	0	-1.602	-1.590	-2.553	-3.687
Wave 2	-	-	-	-	-	-	-	-	-	-	-	-
Wave 3	-	-	-	-	-	-	-	-	-	-	-	-
Wave 4	-	-	-	-	-	-	0.10	-	-	-	-	-
Wave 5	-	-	-	-	-	-	-	-	-	-	-	-
Wave 6	-	-	-	-	-	-	0.69	-	-	-	-	-
Wave 7	-	-	-	-	-	-	0.43	-	-	-	-	-
Intercept Date	-	-	-	-	-	-	-	-	-	-	-	-
INTERVIEWER	-	-	-	-	-	-	-	-	-	-	-	-
CHL Sample (5 years)	-	-	-	-	-	-	-	-	-	-	-	-
First Contact	-	-	-	-	-	-	-	-	-	-	-	-
Not in Labour Force	-	-	-	-	-	-	-	-	-	-	-	-
Full time employed	-	-	-	-	-	-	-	-	-	-	-	-
Higher education	-	-	-	-	-	-	-	-	-	-	-	-
Age (continuous)	-	-	-	-	-	-	-	-	-	-	-	-
Age 40-49 years	-	-	-	-	-	-	-	-	-	-	-	-
Age 50-59 years	-	-	-	-	-	-	-	-	-	-	-	-
Age 60-69 years	-	-	-	-	-	-	-	-	-	-	-	-
Age 70+ years	-	-	-	-	-	-	-	-	-	-	-	-
Male	-	-	-	-	-	-	-	-	-	-	-	-
INTERVIEWER	-	-	-	-	-	-	-	-	-	-	-	-
Experience ***	-	-	-	-	-	-	-	-	-	-	-	-
Same interviewer reference	-	-	-	-	-	-	-	-	-	-	-	-
Number interviewer waves	-	-	-	-	-	-	-	-	-	-	-	-
Reference Person for 0-1	-	-	-	-	-	-	-	-	-	-	-	-
Higher education	-	-	-	-	-	-	-	-	-	-	-	-
Male	-	-	-	-	-	-	-	-	-	-	-	-
Age	-	-	-	-	-	-	-	-	-	-	-	-
Will give income	-	-	-	-	-	-	-	-	-	-	-	-
Interaction with the SHP ***	-	-	-	-	-	-	-	-	-	-	-	-
Interaction with the SHP ****	-	-	-	-	-	-	-	-	-	-	-	-
New Response to questions ***	-	-	-	-	-	-	-	-	-	-	-	-
Old Response to questions ****	-	-	-	-	-	-	-	-	-	-	-	-
One time easy to ask ****	-	-	-	-	-	-	-	-	-	-	-	-
Random Effect Intercept	0.23	0.09	0.30	0.22	0.33	0.29	0.04	-	0.31	-	0.47	-
Random Effect Region	2.678	2.497	5.23	4.97	3.445	3.290	2.93	2.79	4.06	4.47	7.632	7.794
Random Effect Wave	-	-	-	-	-	-	0.97	1.075	2.07	2.34	1.316	1.714
Intercept variance share	0.05	0.4	0.4	0.97	0.05	0.45	0.45	0.45	0.45	0.45	0.45	0.45
Under Dispersion Factor	0.14	0.59	1.59	1.51	1.047	0.83	0.24	0.63	0.71	0.40	0.54	0.54
-2 * LL (ML) Deviance	20,399	19,945	-2,735	-5,274	-2,40	-2,120	40,434	40,715	-15,425	-17,430	-4,375	-5,531

All listed Coefficients "significant": at least twice their standard error (s.e.). Bold: at least 10x their s.e., "-": not applicable/ not considered. Deviance statistics for poisson or binary models are approximative.

* N (applicable) = 5,949

** N (applicable)=14,516

*** For the 2004 model: years at MIS plus 1 if SHP is not the first project, according to interviewer questionnaire. For the 2000-2005 model: years working in the panel – 1.

**** 0=very difficult, ..., 10=very easy

***** In binomial models the variance at the lowest level is constraint 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 $2.678 = .814$ (dispersion factor) * 3.29 (constraint)

Modelling Results

Cross Sectional 2-Level Models

- 34 For the 2004 cross-sectional models, after having identified the set of significant covariates, we first test whether the structure of the models for the “new” and the “old” panel members is different. We thus build two series of OLS regression models for each quality indicator: one with “new” as a dummy variable, and one using the interactions of “new” with all other significant covariates. As it turns out, the fit of the interaction models is only moderately better, but at the expense of a number of degrees of freedom (results not listed). A loglikelihood test shows that the models do not improve significantly. For the sake of parsimony we decide not to include interaction terms.
- 35 In the left of the estimated fixed and random parts of the three cross sectional variance components models are depicted. We find a very small interviewer specific proportion of

the total variance of less than 1% for the social desirability model, a medium proportion of around 5% for the extreme categories model, and a high proportion of 10% for the income nonresponse model. All three models improve with the covariates included, as can be seen from the deviance statistics.⁷ Most interestingly, the sample dummy (“old” sample) is never significant. That is, controlled for the significant respondent covariates (most important age and education), the “mature” panel members do not tend to give more socially desired answers, less extreme category responses, or less income nonresponse answers than those from the refreshment sample. This means that although there is attrition selectivity due to which predominantly the politically uninterested and socially excluded respondents dropped out during the waves 2000 through 2004 (Lipps 2007), survey quality measured by a combination of these variables is not different between the two samples. As to income nonresponse at least this is surprising because Schröpler (2001) reports that income nonresponse measures seem to be a predictor for unit nonresponse in the next wave, and Loosveldt, Pickery and Billiet (2002) find this for more general item nonresponse measures.

- 36 Concerning the other significant respondent variables, the German speaking Swiss give fewer socially desired and extreme category answers. With respect to education, it could be expected that the higher educated people give more socially desired answers, according to the definition of this indicator. Similarly they report fewer extreme category answers, as do male respondents. Regarding age, we depict the effect of age categories in the case of a nonlinear effect, and the effect of the continuously measured variable age in case the effect is linear. Socially desired answers and not reporting the income increase linearly with age, while the use of extreme category answers is especially low for people aged 40-49, and increases thereafter.
- 37 More interesting is the general nonsignificance of the interviewer effects: only the interviewer specific progress has negative effects on income response, i.e., the higher the proportion of already conducted interviews in an interviewer’s workload, the higher the income nonresponse. This may be a habituation or burden effect of the interviewer, in that she is increasingly less motivated to push the respondents to report their income. However the size of total interviewer workload has no effect. This points to a “late case” respondent effect, described later. The positive effect of the self reported impression to be able to convince new respondents (refreshment sample members) to participate at the SHP survey on the respondent use of extreme categories may be due to the fact that respondents on whom lots of efforts have to be spent to convince them to participate, give more extreme category answers (“satisfice” more). Similarly interesting, there is a time effect on socially desired answers and on income reporting: the later in the fieldwork period, the fewer socially desired answers and the higher the income nonresponse. The latter also co-varies with the relative progression of the interviews within interviewers. These effects are probably late cases effects: because the interviewer-respondent assignment is random, it takes longer to convince “difficult” cases who are consequently interviewed later. Presumably these respondents give fewer socially desired answers and refuse to report their income more often. Again, an interviewer habituation effect with respect to a decreasing motivation to push the respondents to give their income may also play a role.

Longitudinal Models

- 38 We like to mention up front that the significance of the coefficients in the longitudinal models cannot be compared with those of the cross-sectional models due to the much higher sample size (16,398 person waves). Overall, the signs and the magnitudes of the coefficients are basically in line with those from the cross-sectional models, if applicable. Some specifics and variations must however be discussed.
- 39 We first note that the interviewer random effects are generally smaller than those in the 2004 cross sectional models. This is also a consequence of the cross-classified data structure: due to the inclusion of the wave, interviewer variance does not contain between-wave variation. Following this argumentation those who work during several waves have a reduced variation, and presumably a tendency to the mean. Moreover, given there is no between-wave variation included in the interviewer variance, there are fewer interviewers per wave (237/6~40) compared to the 2004 wave (81), probably causing a reduced variation.
- 40 The respondent specific random effect is comparatively small in the social desirability model, and high in the extreme category use model and particularly in the income nonresponse model. This shows that there is considerable between wave variation of the social desirability indicator, amounting to around 77% of the total variation. Similarly to the findings in Schröpler (2001) about stereotypical answers, this means that giving social desirable answers is a stable respondent trait only to a minor extent, but does rather depend on situational motivation and moods. In contrast the interviewer specific variation is negligible for the social desirability indicator. For extreme category answers and income nonresponse the interviewer variance proportion amounts to about 5%, the between wave variance proportion to 32% and 14%, respectively.
- 41 The latter shows that – unlike for the social desirability measure – income nonresponse can be considered a stable personality trait, because the lion's share of the total variation is due to the respondent. However, we wonder about the strong decreasing income nonresponse after wave 5, from 6.4% in 2003 to 4% in 2004. This may be due to the fact that in the first panel wave, it was communicated to the respondents that the panel is initially funded for five years. The expectation of a panel termination after five waves led to a high and selective attrition after wave five (Lipps 2006, 2007). However, the fact that also the longitudinal panel participants increase their income response shows that the longitudinal “loyal” subsample consists of persons who may have got a guilty conscience seeing that many of their cohabiters drop out in this year (Lipps 2006) and give better income reports.
- 42 With regard to the respondent coefficients, as in the cross-sectional models, they reflect correlates with respect to the indicator analysed. Only the higher prevalence of giving more extreme category answers by the middle age group is not in line with the findings from the cross-sectional models. The only significant interviewer variable is panel experience, which has negative effects on income nonresponse.

Summary and Conclusion

- 43 In the article, we explore interviewer and respondent interview quality effects in the non-experimental central CATI Swiss Household Panel (SHP) survey which uses a random

interviewer-respondent and interviewer-wave assignment. Using data from the 2nd through the 7th wave, cross-classified multilevel models are chosen in order to separate respondent, interviewer and wave effects in longitudinal variance components models. For the cross-sectional models, we use standard hierarchical multilevel models.

- 44 In the longitudinal models, we find an interviewer variance share of less than 1% for the social desirability measure, and around 5% for extreme category use and income nonresponse. While the respondent variance is less than a third (23%) of the wave variance (77%) for social desirability, it doubles (63%) those of the wave variance (32%) for extreme category use and is even six times as high (82%) as the wave variance (14%) for income nonresponse. The wave specific variations basically remain the same even after including the significant wave dummies. Socially desired answering thus appears to be very variable in individuals, with a much higher intrapersonal than interpersonal variation component. Interviewer effects are negligible. It can thus be concluded that giving social desired answers is not a fixed personality trait, but rather dependent on situational factors. A much smaller intrapersonal variation is calculated for giving extreme category answers and especially income nonresponse, compared with the interpersonal variation. Therefore, intrapersonal stability applies rather with respect to giving extreme value answers (“satisfice”) and particularly not reporting income. As to the interviewer specific share of the total variance, both amount to a substantial rate of around 5%.
- 45 With respect to covariates, we mainly find the expected respondent fixed effects like age, education, and culture, i.e. the difference between the Swiss German and the rest of Switzerland. Surprisingly few interviewer covariates are significant: Interviewer experience has only some positive effects on reporting income. Incidental interviewing the same respondent by an interviewer after one wave has no effect. Also interviewer attributes like trust with respect to data protection, and various satisfaction indicators do not play any role. Similarly, the interviewer-respondent matching variables sex, age, and education had no effect, once the (respondent) main effects are controlled.
- 46 In the cross-sectional two-level models we use data from the SHP 2004 wave and the 2004 interviewer survey. Here, we find slightly higher interviewer shares of the total effects, especially with respect to income nonresponse. We were especially curious about the differences between the original (1999) sample, then in their 6th wave, and the refreshment sample, first asked in 2004, and equally representative of the Swiss population. The sample discriminating dummy is significant in none of the three models. This means that, although attrition in the SHP was considerable, the survey quality indicators considered are not significantly different for the mature and the refreshment sample, *if the respondent socio demographic characteristics are controlled*. In case of social desirability, the sample dummy is not even significant in the unconditional model.
- 47 We can identify a within wave seasonal effect on socially desired answers and income nonresponse. We encounter fewer socially desired answers and more income nonresponse the later the interview is conducted in the field. An explanation is that the more difficult cases must have first convinced to participate, before the interview starts. This time effect is also in line with known effects from the so called late cases (Stoop 2005, see also Kennickell 2000).

BIBLIOGRAPHY

- Andersen, A.-M. N. and Olsen, J. (2002). Do Interviewers' Health Beliefs and Habits Modify Responses to Sensitive Questions? A Study using Data Collected from Pregnant Women by Means of Computer-assisted Telephone Interviews. *American Journal of Epidemiology* **155** (1), 95-100
- Billiet, J. and Loosveldt, G. (1988) Improvement of the Quality of Responses to Factual Survey Questions by Interviewer Training. *The Public Opinion Quarterly* **52** (2), pp. 190-211
- Brick, J.M., McGuinness, R., Lapham, S.J., Cahalan, M., Owends, D. and Gray, L. (1995). *Interviewer Variance in two Telephone Surveys*. SESTAT Methodological Research Compendium, SESTAT (<http://srsstats.sbe.nsf.gov/>)
- Campanelli, P. and O'Muircheartaigh C. (1999) Interviewers, Interviewer Continuity, and Panel Survey Nonresponse. *Quality & Quantity* **33**, pp. 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**, pp. 129-144
- Diekmann, A. (2005) „Betrug und Täuschung in der Wissenschaft“. Datenfälschung, Diagnoseverfahren, Konsequenzen. *Swiss Journal of Sociology* **31** (1), pp. 7-29
- Esser, H. (193) *Response Set: Habit, Frame or Rational Choice?* In: Krebs, D. and Schmidt, P. (eds.): *New Directions in Attitude Measurement*, pp. 293-314. Berlin: de Gruyter
- Frick, J.R. and Grabka, M.M. (2005). Item nonresponse on income questions in panel surveys: Incidence, imputation and the impact on inequality and mobility. *Allgemeines Statistisches Archiv* **89** (1), pp. 49-61
- Gabadinho, A. and Budowski, M. (2002) Construction of Income Variables for Wave 2. *Swiss Household Panel Discussion Paper* **3-02**, Neuchâtel
- Groves, R.M. and Magilavy, L.J. (1986) Measuring and explaining Interviewer Effects in centralised Telephone Surveys. *Public Opinion Quarterly* **50**, pp. 251-266
- Groves, R.M., Fowler, F.J., Couper, M.P., Lepkowski, J.M., Singer, E. and Tourangeau, R. (2004). *Survey Methodology*. Wiley Series in Survey Methodology.
- Heeb, J. and Gmel, G. (2001). Interviewers' and respondents' effects on self-reported alcohol consumption in a Swiss health survey. *Journal of Studies on Alcohol*, July, pp. 434-442
- Hill, D. and Willis, R. (2001) Reducing Panel Attrition. A Search for effective Policy Instruments. *J. Human Resources* **36** (3), pp. 416-438
- 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: Biemer, Groves, Lyberg, Mathiowetz and Sudman (eds.): *Measurement Errors in Surveys*. New York: Wiley
- Jäckle, A., Roberts, C. and Lynn, P. (2006) Telephone versus Face-to-Face Interviewing: Mode Effects on Data Quality and likely Causes. Report on Phase II of the ESS-Gallup Mixed Mode Methodology Project. *ISER Working Paper*, **2006-41**, Colchester, University of Essex

- Japiec, L. (2005) *Quality issues in interviewer surveys: some contributions*. Department of Statistics, Stockholm University. PhD Thesis
- Japiec, L. (2006) Quality issues in interviewer surveys: some contributions. *Bulletin de Méthodologie Sociologique* **90**, pp. 26-42
- Kennickell, A. (2000) *What do the "late" Cases tell us? Evidence from the 1998 Survey of Consumer Finances*. Paper prepared for the 1999 International Conference on Survey Nonresponse, Portland, Oregon
- Krosnick, J.A. (1991) Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys. *Applied Cognitive Psychology* **5**, pp. 213-236
- Laurie, H., Smith, R. and Scott, L. (1999) Strategies for Reducing Nonresponse in a Longitudinal Panel Survey, *Journal of Official Statistics* **15**, pp. 269-282
- Lipps, O. (2005). Interviewer and Respondent Effects in CATI Data – a multi-criterion, multilevel analysis. *Swiss Household Panel Discussion Paper* **2-05**, Neuchâtel
- 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. *Swiss Household Panel Working Paper* **6-06**, Neuchâtel
- Lipps, O. (2007) Attrition in the Swiss Household Panel. To be published in *Methoden, Daten, Analysen* **1**.
- Loosveldt, G. and Carlton, A. (1997) Evaluation of Nonresponse in the Belgian Election Panel Study '91-'95. *Proceedings of the Survey Research Methods Section*, American Statistical Association
- 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** (4), pp. 545-557
- Mahalanobis, P.C. (1946) Recent Experiments in Statistical Sampling in the Indian Statistical Institute. *J Royal Statistical Society* **109**, pp. 325-378
- Mayer, T.S. (2001). Interviewer Attitudes about Privacy and Confidentiality. *Research Report Series*, Statistical Research Division, U.S. Bureau of the Census
- Mensch, B.S. and Kandel, D.B. (1988) Underreporting of Substance Use in a National Longitudinal Youth Cohort: Individual and Interviewer Effects. *Public Opinion Quarterly* **52** (1), pp. 100-124
- Oksenberg, L. and Cannell, C. (1988) *Effects of Interviewer vocal Characteristics on Nonresponse*. In: Groves, R., Biemer, P., Lyberg, L., Massey, J., Nicholls II, W. and Waksberg, J.: *Telephone Survey Methodology*. New York: Wiley
- O'Muircheartaigh C. and Campanelli, P. (1999) A Multilevel Exploration of the Role of Interviewers in Survey Non-Response. *J. Royal Statistical Society A* **162** (3), pp. 437-446
- Pannekoek, J. (1988). Interviewer Variance in a Telephone Survey. *Journal of Official Statistics* **4** (4), pp. 375-384
- Peneff, J. (1988) The Observers observed: French Survey Researchers at Work. *Social Problems* **35** (5), pp. 520-535.
- Pickery, J. and Loosveldt, G. (1998) The impact of Respondent and Interviewer Characteristics on the Number of "No Opinion" answers. *Quality & Quantity* **32**, pp. 31-45
- Pickery, J. and Loosveldt, G. (2001) An Exploration of Question Characteristics that Mediate Interviewer Effects on Item Nonresponse. *Journal of Official Statistics* **17** (3), pp. 337-350

- Pickery, J. and Loosveldt, G. (2002). A Multilevel Multinomial Analysis of Interviewer Effects on Various Components of Unit Nonresponse. *Quality & Quantity* **36**, pp. 427-37
- Pickery, J. and Loosveldt, G. (2004). A Simultaneous Analysis of Interviewer Effects on Various Data Quality Indicators with Identification on Exceptional Interviewers. *Journal of Official Statistics* **20** (1), pp. 77-89
- Rasbash, J., F. Steele, W. Browne, and B. Prosser (2004). *A users guide to MLwiN version 2.0*. London, Institute of Education
- Scherpenzeel, A. (2005). Interviewer Study 2004. Profiles of the MIS interviewers working for the SHP and effects on the data. *Swiss Household Panel Discussion Paper* **3-05**, Neuchâtel.
- Schnell, R. (1991) Der Einfluss gefälschter Interviews auf Survey-Ergebnisse. *Zeitschrift für Soziologie* **20** (1) pp. 25-35
- Schräpler, J.-P. (2001). Respondent Behaviour in Panel Studies. A case Study of the German Socio-Economic Panel (SOEP). *DIW Discussion Paper* **244**, pp. 118-156
- Shoemaker, P.J., Eichholz, M. and Skewes, E.A. (2002). Item Nonresponse: Distinguishing between don't Know and Refuse. *Int. J. Public Opinion Research* **14**, pp. 193-201
- Singer, E., Frankel, M.R. and Glassmann, M.B. (1983) The Effect of Interviewer Characteristics and Expectations on Response. *Public Opinion Quarterly* **47** (1), pp. 68-83
- Smit, J.H. (1993) *Suggestive Questioning and Probing in Interviews - Occurrence, Causes and Effects*. Faculty of Social Cultural Sciences VU, Department of Social Research Methodology, University of Amsterdam. PhD Thesis.
- Smit, J.H., Dijkstra, W. and van der Zouwen, J. (1997) Suggestive Interviewer Behaviour in Surveys: An experimental Study. *J. Official Statistics* **13** (1), pp. 19-28
- Snijders, T.A.B. and Bosker, R.J. (1999). *Multilevel analysis*. Newbury Park, California: Sage
- Snijders, G., Hox, J. and de Leeuw, E. (1999) Interviewer's tactics for Fighting Survey Nonresponse. *J. Official Statistics* **15** (2), pp. 185-198
- Spiess, M. and Kroh, M. (2005) Documentation of Sample Sizes and Panel Attrition in the German Socio Economic Panel (SOEP) 1984-2004. *Data Documentation* **6**, DIW Berlin
- Stokes, L. and Yeh, M.-Y. (1988) *Searching for Causes of Interviewer Effects in Telephone Surveys*. In: Groves, R., Biemer, P., Lyberg, L., Massey, J., Nicholls II, W. and Waksberg, J.: *Telephone Survey Methodology*. New York: Wiley
- Stoop, I. (2005) *The Hunt for the last Respondent: nonresponse in Sample Surveys*. PhD Thesis, Utrecht University
- van der Zouwen, J. (2006) Interviewer and Survey Researcher: Mutual Dependencies. *Bulletin de Méthodologie Sociologique* **89**, pp. 49-64
- van der Zouwen, J. and Dijkstra, W. (1988) Types of inadequate Interviewer Behaviour in Survey Interviews: Their Causes and Effects. *Bulletin de Méthodologie Sociologique* **18**, pp. 5-20
- van der Zouwen, J. and de Leeuw, E. (1989) The Relationship Between Mode of Administration and Quality of Data in Survey Research. *Bulletin de Méthodologie Sociologique* **19**, pp. 3-14
- van Meter, K. (2005) Ongoing Research: Studying Survey Interviewers: A Call for Research and an Overview. *Bulletin de Méthodologie Sociologique* **87**, pp. 51-71
- Watson, N. and Wooden, M. (2004) Sample Attrition in the HILDA Survey. *Australian Journal of Labour Economics* **7** (2), pp. 293-308.

Winkler, N., Kroh, M. and Spiess, M. (2006) Entwicklung einer deutschen Kurzskaala zur zweidimensionalen Messung von sozialer Erwünschtheit. *DIW Discussion Paper* 579, DIW Berlin

NOTES

1. The number of waves in a panel survey can be considered as survey experience of (longitudinal) respondents and (longitudinal) interviewers. In reality, however, most interviewers stay with the panel only for few waves, at least in centralised telephone panel surveys.
2. A nice illustration of the response decision of individuals who value different aspects with different weights is given in Groves et al. (2004, p. 177).
3. See for an example of the development of an appropriate scale for measuring social desirability Winkler, Kroh and Spiess (2006).
4. E.g. Pickery and Loosveldt (2004) view the non-occurrence of at least one extreme answer category as proving a low interview quality. However we have the impression that a heavy use of extremes documents a certain form of satisficing. This assessment is the result of listening to a number of individual CATI interviews.
5. With 1% of all 22 attitude questions not answered, each respondent has a mean number of .2 not answered questions.
6. See for poisson models in a multilevel context with interviewers as second level Pickery and Loosveldt (1998).
7. The differences of the deviance statistics is approximately χ^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.

ABSTRACTS

Especially in panel surveys, respondent attrition, respondent learning, and interviewer experience effects play a crucial role with respect to data quality. We examine three interview survey quality indicators in the same survey in a cross sectional as well as in a longitudinal way. In the cross sectional analysis we compare data quality in the mature original sample with that in a refreshment sample, surveyed in the same wave. Because in the same wave an interviewer survey was conducted, collecting attitudes on their socio demography, survey attitudes and burden measures, we are able to consider interviewer fixed effects as well. The longitudinal analysis gives more insight in the respondent learning effects with respect to the quality indicators considered by considering the very same respondents across waves. The Swiss Household Panel, a CATI survey representative of the Swiss residential population, forms an ideal modelling database: the interviewer – respondent assignment is random, both within and across waves. This design avoids possible confusion with other effects stemming from a non-random assignment of interviewers, e.g. area effects or effects from assigning the best interviewers to the hard cases. In order to separate interviewer, respondent and wave effects, we build cross-classified multilevel models.

Effets synchrones et longitudinaux sur la qualité des panels CATI des intervieweurs et des répondants : C'est en particulier dans les enquêtes par panel que les effets de la perte de répondants, de l'apprentissage par les répondants, et de l'expérience des intervieweurs jouent un rôle crucial en ce qui concerne la qualité des données. Nous examinons de manière synchrone et longitudinale pour la même enquête trois indicateurs de qualité. Dans l'analyse synchrone, nous comparons la qualité des données dans l'échantillon d'origine non-renouvelé avec celle d'un échantillon qui l'a été, les deux étant sondé dans la même vague. Pour la même vague, une enquête des intervieweurs a été faite, fournissant des données sur leur démographie sociale, leurs attitudes envers des enquêtes, et leur estimation de la charge de travail, nous permettant ainsi d'estimer l'effet des intervieweurs. L'analyse longitudinale permet l'analyse des effets de l'apprentissage par les répondants par rapport aux indicateurs de qualité en considérant les mêmes répondants dans plusieurs vagues. Le Panel Suisse des Ménages – une enquête CATI représentative de la population résidentielle suisse – fournit une base de données idéale pour modéliser puisque l'affectation des intervieweurs est au hasard pour toutes les vagues et dans chacune. Ce plan évite la confusion possible avec d'autres effets dus aux affectations non aléatoires telles que les affectations par quartiers où les meilleurs enquêteurs traitant les cas les plus difficiles. Pour séparer les effets des intervieweurs, des répondants et des vagues, nous utilisons des modèles multi-niveaux croisés.

INDEX

Mots-clés: Affectation aléatoire, Effets des intervieweurs, Effets des répondants, Effets longitudinaux, Enquêtes CATI centralisées, Modèles multi-niveaux, Qualité des données d'enquête

Keywords: Centralised CATI, Cross-Classified Multilevel, Interviewer Effects, Longitudinal Effects, Random Assignment, Respondent Effects, Survey Data Quality

AUTHOR

OLIVER LIPPS

Swiss Household Panel; oliver.lipps@unine.ch