Differential survey strategies based on R-indicators[1].

Annemieke Luiten and Willem Wetzels
Statistics Netherlands
CBS-weg 11
NL-6412 EX Heerlen
a.luiten@cbs.nl

Keywords : Differential fieldwork strategy, responsive design, response, contact, cooperation, representativeness, R-indicators, mixed-mode design.

1. Introduction

In the recent survey literature a lot of attention has been devoted to the level of effort invested in a survey and the so-called continuum-of-resistance model, see e.g. Lin and Schaeffer (1995). In the continuum-of-resistance model, households and enterprises are thought to behave along two dimensions, ease-of-contact and ease-of-participation. Attached to those dimensions are individual contact and response probabilities, and, when combined, overall individual response probabilities which form the basis of our perception of representativeness.

Associated with the continuum-of-resistance model is the level of effort invested by the survey organisation. The more effort the survey researcher invests in contacting households and converting reluctant respondents, the higher the response rate. The level of effort invested has increased during the past decades in many countries in order to maintain acceptable response rates. As a consequence, the costs of surveys per sampled unit have also increased. It is, therefore, of great importance how the additional efforts are allocated, i.e. the efficiency of these efforts becomes of a growing importance.

In the literature the implications of increased efforts are often debated. Apart from an increased risk of measurement errors, it is also questioned whether the additional efforts lead to more quality and a more representative set of respondents. Clearly, if difficult-to-contact or difficult-to-convert individuals are different from other individuals, then a focus on easy-to-contact and easy-to-convert units will increase the contrast between respondents and non-respondents. As a result, the response rate may have increased but the non-response error may not have changed or even may be increased. For instance, a follow-up using telephone interviewing may help raising response rates, but can only be applied to households with a listed phone number. Hence, a single-minded increase of the level of effort may not help improving the quality of the response.

One may, therefore, differentiate the level of effort between households and enterprises to get a balanced, representative composition of the response, see Groves and Heeringa (2005) and Van der Grijn, Schouten and Cobben (2006). In determining the level of effort needed for a certain household or enterprise, indicators of representativeness may serve as useful tools. The differentiation of the level of effort may be directed at increasing the response rate while maintaining or even enhancing the representativeness of the response.

Groves and Heeringa (2005) propose so-called responsive designs, which are designs that are dynamic with respect to the composition of the response, i.e. they aim at controlling the response to a survey during the data collection. One may also decide to differentiate beforehand, so that different sampled units are assigned different fieldwork protocols based on historic fieldwork paradata.

[1] *This research was part of the RISQ Project, financed by the 7th Framework Programme (FP7) of the European Union. Cooperation Programme, Socio-economic Sciences and the Humanities, Provision for Underlying Statistics*

Representativity Indicators and partial Representativity Indicators (Shlomo et.al., 2009; Loosveldt and Beullens, 2009) may serve as tools to facilitate differentiated fieldwork strategies before and during the data collection phase. This paper describes a pilot in which the ascription of differential strategies to different groups in the sample is aided by these indicators. In this pilot, the differentiation occurred before the commencement of fieldwork, based on prior knowledge of comparable sample units' behaviour in similar surveys. In a parallel pilot, described by Kleven et.al., (2010), the indicators are used dynamically, during data collection. Aim of the pilot was to augment representativeness of sample realisation, against minimally equal, but ideally less, costs and with minimally equal, but ideally higher, response rates.

2. Method

As a vehicle for the pilot the monthly Survey of Consumer Confidence (SCC) was used. This is a CATI survey, conducted among 1500 households of whom a listed telephone number can be found. Questions are asked of any person in the household core (head of household or partner). The length of the questionnaire is about eight minutes. Questions are asked related to sentiments about the household's economic situation and expenditure. Fieldwork is conducted in the first ten workdays of each month.

Because the SCC is conducted monthly, a wealth of information is available about contact and cooperation characteristics of former sample units. This accumulated knowledge was used to determine fieldwork strategy prior to the start of the fieldwork.

The fieldwork of the pilot was conducted during the months of October, November and December 2009. It was conducted alongside the regular SCC, during the same 10 day fieldwork period, with a similar sampling method, a similar sample size and, as far as possible, the same interviewers. The SCC served as control for the response and representativeness measures.

In order to achieve the aim of better representativeness with lower costs, a mixed mode design was chosen, in which a mail and/or web first round was followed by a CATI follow-up of nonrespondents. Mail and web questionnaires not only cost less to administer than CATI questionnaires, they can also reach respondents that are otherwise hard to contact and/or to convince to cooperate. Calculation of consumer confidence occurs on data collected within the first ten days of each month. As it is not feasible to conduct a mixed mode design with CATI follow-up within ten days, the design of the pilot was adapted. Figure 1 illustrates the design of the pilot.

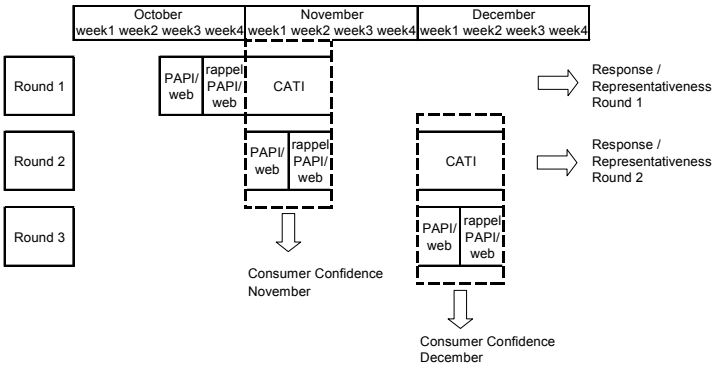

Figure 1. Design of the pilot.

The first mail/web round was conducted during the last fortnight of October[2]. One week after sending the advance letter, a reminder was sent. Ten days later, the CATI follow-up of nonresponse

[2] Because of time constraint issues, and because the data of this first round were not used to calculate consumer confidence, it was decided to do the first round in the last to weeks of October, in stead of the first two weeks. Analyses showed that this had no response implications.

started, which was conducted in the first two weeks of November. Three days before the first of November, advance letters were sent for the second round of fieldwork, again to be followed by a reminder one week later. CATI follow-up of non-respondents of round 2 started on the first of December. As in November, three days prior to the first of December, advance letters and questionnaires were sent to the third sample. This sample received the advance letter and one reminder, but no CATI follow-up. As is shown in figure 1, Consumer Confidence is calculated across two different samples: mail/web response of month T and the CATI follow-up response of month T-1. Response rates and representativeness of the response, are however calculated within one sample (i.e., within each round).

Fieldwork strategy of the pilot was determined based on what could be learned of the response propensities of sample units in two existing datasets. The SCC 2004 (available at www.r-indicator.eu) was used to estimate contact and cooperation propensities for the telephone survey. The dataset of the SCC 2004 contains cooperation and contact information of about 18.000 sample units, as well as auxiliary information, made available from CBS registries. The CBS Safety Monitor 2007 was used to estimate cooperation propensities for the web/mail survey.

2.1 Linked data

The samples of both SCC and the experimental SCC (the pilot) were linked to the Social Statistical Database of Statistics Netherlands. This database consists of administrative information on persons, households, jobs, benefits and pensions.

The variables used for the analysis are displayed in Table 1. There is geographical, demographic and socio-economic information on different levels. The lowest level in the registries is the person. In this analysis, however, the level is the household. All person variables are therefore aggregated to a household level, based on information about the household core (head of household and partner). Because of this aggregation, the variables ethnic group and gender have a category to indicate a mixture of the categories on the personal level (e.g., mixed native-foreign). The next level comprises information at the postal code level.

Table 1 Linked data to the Survey of Consumer Confidence

Variable	*Categories*
Household level	
Ethnic Group	Native, Morrocan, Turkish, Suriname / Netherlands Antilles, other non-western, other western, mixed and unknown. For the present analyses aggregated to native, foreign, mixed and unknown
Gender	all male, all female, mixed, unknown
Average age of household core	15-30;31-44;45-65; over 65, unknown
Type of Household	Single, partners without children, partners with children, single parents, unknown
Postal code area level	
Degree of urbanization	very strong, strong, moderate, low, not urban, unknown
percentage non-western non-natives	very high, high, average, low, very low, unknown
average monthly income	quartiles

Each variable has a category ‘information not available’. This has to do with linking sample units to registries. As registries are never entirely up to date, people moving, building or demolishing dwellings, and unregistered people may lead to unavailable information both at the level of the individual or household, or the level of the postal code. Rather than treating these absent data as missing values, they are incorporated as meaningful values.

2.2 Over- and under-represented groups

Loosveldt and Beullens (2009) describe how partial R-indicators can be calculated to determine which groups are over- or under-represented in sample realisation. This technique was used to determine groups within SCC 2004 with a high, medium, or low contact propensity and groups with

a high, medium, or low cooperation propensity. This propensity was then projected upon the new samples for the pilot.

A simple sum score was used to determine the expected contact and cooperation propensity in the samples of the pilot and control group. For example, the partial R-indicators showed that elderly households, households with low incomes, households of non-Dutch origin, households living in a neighbourhood with a high percentage of people of non-Dutch origin, and single persons were less likely to participate than other households. The more of these elements present in a single household, the lower the chance of cooperation. I.e., an elderly household with a low income would have a lower cooperation propensity than an elderly household with a high income. A similar exercise was done for chance of contact, where it was shown that young households, living alone or in a partnership without children, households living in highly urban areas, households of non-Dutch origin and households living in neighbourhoods with a high percentage of non-Dutch, have a low contact propensity. Again, the propensity is lower, the more elements present. Based on these analyses, each sample unit was classified as having a high, medium or low contact propensity and having a high, medium or low cooperation propensity. Results of the first round showed that the medium cooperation group should be split in two. In the second round, therefore, four groups were differentiated.

The propensity analysis for SCC 2004 was repeated for a Statistics Netherlands' survey with a mixed mode design (the Safety monitor), to investigate how high, medium or low response propensity in a CATI survey related to response behaviour in a web / mail first round. For the Safety monitor, people were invited to participate in a web survey. They could however request to receive a mail questionnaire.

It was shown that cooperation propensity, as calculated for the CATI SCC data was highly predictive of web response as well. Web response of the people predicted to be relatively 'easy', i.e., having a high cooperation probability, was 31,3% in the first wave web round, whereas the 'hard' group had a response of 4,8%. However, the group with a low web response, had a relatively high mail response. Mail response in the group we defined as 'easy' on cooperation was 6,4%, but 13,5% in the group with the lowest cooperation propensity.

These findings led to the conclusion that both a mail and a web version of the pilot questionnaire were necessary in order to gain cooperation in the hardest group.

2.3 Differential fieldwork strategy

Web/mail wave.

On the basis of the predicted web and mail response of the three cooperation groups, the following design was decided upon for the first web/mail wave:

- households with a high chance of cooperation would receive an invitation for the web survey
- households with a medium chance of cooperation would receive an invitation for the web survey and a mail questionnaire. Either could be filled in.
- households with a low chance of cooperation received only a mail questionnaire. This simplified the advance letter to a great extent, and it was expected that that would be beneficial to response.

All households received one reminder. The reminder mentioned that an interviewer would call, if the questionnaire was not received within shortly. No new mail questionnaire was sent along with the reminder.

Telephone wave.

In the second wave, the nonresponse was followed up by CATI. In this wave it was attempted to

1. stimulate chance of contact for sample units with a low contact propensity
2. dampen the number of contact attempts for units with a high contact propensity
3. stimulate cooperation for sample units with a low cooperation propensity, and
4. dampen cooperation for sample units with a high cooperation propensity.

For different groups, different approach strategies were defined in the CATI management system, by means of the definition of different time slices. The CBS CATI management system is a Blaise application. Defining time slices enables the CATI management system to allocate telephone numbers according to criteria that can be different for different time slices. By defining multiple time slices per day, an address can be called more than once a day. Defining different time slices for groups of addresses makes differential fieldwork strategy possible.

One time slice was defined for elderly Dutch households (65 years and older). This group has a high contact propensity, but a low cooperation propensity. To make interviewer capacity available for groups who needed a higher number of contact attempts, the CATI fieldwork for this group was postponed to the second week of fieldwork period. The households were called primarily during daytime. One evening only was reserved for hitherto uncalled numbers in this group: the last night of the fieldwork period. The definition of this time slice not only freed valuable capacity for evening calls, but was also cost effective, as daytime shifts are remunerated 40% less than evening shifts. In the second month or the pilot, the definition of this time slice was slightly adapted however, because the dampening effect was too strong. In the second month, the fieldwork for this group started in the first week of the fieldwork period, and numbers were called on two weekday evenings each week. In the last week a further adaptation was made, to make numbers of this group available on two additional evenings.

The second time slice consisted of single households, households of non-Dutch origin, households in highly urban areas and households consisting of young people (30 years or under). The time slice was to be called in every shift (morning, afternoon and evening), every day of the fieldwork period.

A third time slice consisted of people of 31 to 45 years of age, not belonging to the second time slice. This group was to be called during the evening for the first two contact attempts. Subsequent attempts could be made during the day also. The last time slice was the miscellaneous ‘other’ group. They received the default treatment that the control group, the regular SCC also received.

Although the definition of time slices determines when numbers can be called, whether they are actually called is dependent on the available interviewer capacity in a shift. To assure that if limited capacity was available, numbers of households with the lowest contact propensity would be called with preference, these numbers were prioritized in each day batch by using an algorithm that used the predicted contact probability.

Definition of time slices and prioritizing numbers in a day batch, were measures taken to influence contact probability. In order to influence cooperation probability, the assignment of numbers to specific interviewers was manipulated. Based on their SCC work in 2008 and the first half of 2009, interviewers were classified in three categories, according to the cooperation rates achieved. A top quartile of the best interviewers (mean cooperation rate in 2008-2009: 82,1%), a middle group of the second and third quartile (cooperation rate 74%) and a third group in the lowest quartile (65,6%). The best interviewers called households with the lowest cooperation propensity. The interviewers with the lowest response rated called households with the highest probability of cooperation. The group in between called the middle group. On top of that, if appointments were made

for a certain date or time, the appointment would be followed up by an interviewer of the lowest quartile. On the other hand, if a 'soft' appointment was made '*call me back some other time'*, this would be followed up by an interviewer in the best quartile. In the second month of the pilot, the middle group was split in two, to be able to make a finer distinction in the households with a medium cooperation propensity. The assignment of groups of addresses to groups of interviewers was handled by the CATI management system. To prevent planning problems, interviewers of a 'better' quartile would always be allowed to call numbers meant for a 'lower' quartile. In practice, this possibility was seldom used, however. See the Blaise CATI guide (2004) for details of how definition of time slices and allocation of interviewers to addresses may be attained.

2.4 Fieldwork in the control group

The regular SCC is a one mode - telephone only- survey. No information is available beforehand of the characteristics of the households. In practice, this means that all households have an equal probability to be selected in the day batch, although households with whom appointments are made are prioritized. 80% of the fieldwork is performed during evening shifts. During daytime shifts, an interviewer is present to call appointments made for daytime, and s/he may use spare time to work other numbers. Supervisors determine daily whether the work advances satisfactory and whether it would make sense to call an address one or more additional times. The basis for this decision is overall response rate. As in the experimental group, an advance letter is sent some days prior to commencing fieldwork. In neither pilot nor SCC incentives were given or promised, and no refusal conversion was attempted.

3. Results

3.1 Response

Table 2 shows response results for the regular Survey of Consumer Confidence (the control group) and the pilot. Despite the slight changes in the design in the second month, results were highly comparable and are collapsed.

Table 2 Response results of the SCC and the pilot

	SCC		Pilot	
Results	*N*	*Percent*	*N*	*Percent*
Ineligible	225	7,5	144	4,8
Non-contact	196	6,5	183	6,1
Not present during fieldwork period	73	2,4	62	2,1
Not able (ill, dementia)	115	3,8	122	4,1
Language problems	40	1,3	26	0,9
Refusal	467	15,6	548	18,3
Response	1884	62,8	1915	63,8
Response WEB-PAPI			1081	36,0
Response CATI			834	27,8

In both pilot months, the number of response cases was higher in the experimental group. Because of the substantial number of ineligible cases in the SCC, the response rate RR01 (AAPOR, 2006), i.e., the response of eligible cases, was slightly higher in the SCC, however. Ineligible cases in this kind of CATI research consist mostly of disconnected telephone numbers. Disconnected numbers are correlated with a predicted low chance of noncontact and non-cooperation however. It will be shown that sending a mail questionnaire to high-risk addresses contributed substantially to the response of these households, and to a better representative response.

3.2 Representativeness

The response results show comparable response rates for pilot and control group. Table 3 shows, by means of R-indicators, the representativeness of this response, as well as that of each

steps in the fieldwork process: the representativeness of the eligible part of the sample, of those contacted, of those able, and of those cooperating. The R-indicator ranges from 0 (no representativeness) to 1 (complete representativeness). As can be seen in the table, the R-indicator of each subsequent step is higher in the pilot than in the control group. Only for the R-indicator of response confidence intervals do not overlap, however ($p < .05$).

Table 3 R-indicators and 95% confidence interval for eligible, contacted, able, cooperating and responding cases in the SCC and the pilot.

	SCC		Pilot	
	R	CI	R	CI
Eligible	0,84	(0,809 - 0,870)	0,88	(0,851 - 0,910)
Contacted	0,83	(0,796 - 0,862)	0,87	(0,836 - 0,900)
Able	0,86	(0,830 - 0,886)	0,85	(0,827 - 0,882)
Cooperation	0,87	(0,837 - 0,901)	0,89	(0,857 - 0,916)
Response	0,77	(0,738 - 0,804)	0,85	(0,816 - 0,877)*

** p<.05*

Analysis of the partial R-indicators shows how the experimental manipulations affected sample composition. Partial indicators show for each auxiliary variable, as well as their respective categories, how much the variable deviates from representativeness (partial R = 0). The larger the value of partial R, the higher the contribution to un-representativeness. Partial R indicators can be calculated unconditionally or bivariately, but also conditionally. In the latter case, the partial R indicators are corrected for the other auxiliary variables in the model. Unconditional indicators can be either positive (the category is overrepresented) or negative (underrepresented). Conditional indicators can only be positive. See Shlomo, Skinner, Schouten, Carolina and Morren (2009) for a detailed description of the characteristics of partial R-indicators.

Table 4 shows the results of the analysis of unconditional partial indicators for some of the auxiliary variables used, for each step in the fieldwork process. See Luiten and Wetzels (2010) for an extensive discussion of all results. The SCC starts with determining eligibility, whereas the pilot starts with the response on the web/mail round. For each auxiliary variable, the italic value is the composite contribution to representativeness; the other values describe the positive or negative contribution of the categories of the variable. Table 4 shows for the variable 'Age', that better representativeness is attained in the pilot in all columns, with the exception of the column 'able'. The better representativeness of Age in the pilot is attained for all age groups concerning eligibility, although the difference is only in the extent to which groups are over- or under-represented. Concerning contact, representation of the under 30 years old and the elderly is better in the pilot, signifying a higher contact rate for the young households, and a lower contact rate for the elderly. The findings for cooperation show that representativeness of the age groups is the same for pilot and control group, with the exception of the elderly, who are better represented in the pilot. In the final column, this is reflected in a better overall representation of Age in the pilot, especially through a better result for the young households and the households of which no information is available. The results for 'Gender' are comparable, with a notable exception: representativeness of cooperation is not better in the pilot, as a result of an over-representation of households consisting of one or more males. Together with under-representation in contact, this translates in a virtually perfect representativeness of male households in the response, however. Female households are less under-represented, households of unknown gender composition are less under-represented and households of mixed gender are less over-represented in the pilot. Similar findings are found for the variable 'Household composition', in which all categories appear to be better represented in the response of the pilot.

Table 4 Unconditional partial indicators for SCC and pilot

	SCC					Pilot					
	eligible	contact	able	coop-eration	response	response web/mail	eligible	contact	able	coop-eration	response
Age	*59*[(1)]	*52*	*38*	*35*	*58*	*62*	*43*	*33*	*45*	*21*	*36*
< 30	-26	-41	8	17	-25	-10	-18	-12	1	14	-10
30-44	-17	-13	15	14	1	-34	-11	-23	20	12	0
45-64	13	8	14	7	29	3	9	15	18	-8	22
65>	26	22	-29	-26	-10	43	20	10	-35	-7	-10
no information available	-42	-17	-10	-2	-43	-28	-30	-8	-6	3	-25
Gender	*93*	*88*	*67*	*30*	*156*	*81*	*69*	*65*	*64*	*49*	*91*
Male(s)	-18	-37	-9	-1	-43	13	-5	-30	3	25	-3
Mixed	21	27	27	7	54	2	15	20	20	-9	30
Female(s)	-2	-8	-30	-13	-38	0	-1	-5	-29	-4	-28
no information available	-51	-27	-19	5	-53	-28	-37	-9	-8	0	-34
Household composition	*51*	*49*	*35*	*18*	*88*	*53*	*40*	*38*	*37*	*29*	*52*
Single	1	-28	-23	0	-37	4	0	-27	-31	10	-33
Partners, with children	15	20	13	2	32	37	18	14	12	-16	17
Partners, no children	11	20	16	7	37	-32	3	18	17	-3	25
Single parent	-16	-5	-3	-16	-28	3	-13	-7	0	22	4
no information available	-45	-29	-16	-5	-57	-20	-33	-11	-3	4	-26

*(1) Conditional R-indicators * 1000*

To illustrate how unconditional R-indicators relate to traditional analyses of (non)response, table 5 shows bivariate analyses of the relation between auxiliary variables and eligibility, contactability, etc., in the pilot and the regular SCC. The values in table 5 should be compared to the italic values in table 4, showing the overall contribution to representativeness of the variable. It can be seen that results of the two analyses are highly comparable. Whenever the partial R-indicators show a larger deviation from representativeness, Cramèrs V is larger. The bivariate analysis does not show, however, what the contribution of the respective subgroups to the deviation is, as the partial R-indicators do.

Table 5 Bivariate analyses (Cramèr's V) of eligible, contacted, able, cooperating, and responding cases in SCC and Pilot

	Eligible		*Contacted*		*Able*		*Cooperation*		*Response*		*Response Web/mail*
Variable	*SCC*	*Pilot*	*SCC*	*Pilot*	*SCC*	*Pilot*	*SCC*	*Pilot*	*SCC*	*Pilot*	*Pilot*
Age	0.25	0.22	0.18	0.13	0.15	0.17	0.08*	0.05 ns	0.13	0.09	0.13
Gender	0.22	0.19	0.19	0.15	0.16	0.14	0.04 ns	0.06*	0.20	0.11	0.06*
Household composition	0.20	0.19	0.18	0.15	0.12	0.15	0.05 ns	0.07*	0.18	0.11	0.11

** p < .05; ns No significant relation; all other values p<.01*

The results of table 4, as described above, show that in all probability, variables are interconnected. The better representativeness of singles, males and elderly, for example, may very well be a better representativeness is the group of elderly single males. To analyse if this is the case, conditional R-indicators were calculated, that correct for the other auxiliary variables in the model. Table 6 shows conditional R-indicators for the same selection of variables.

Correction for the other variables, does not change the interpretation of the variables 'Age' and 'Gender'. Age is better represented in the pilot than in the SCC in all steps. Gender is better represented in the response, but not in cooperation. When corrected for the other variables, singles and single parents are no longer better represented in the pilot than in the SCC.

Table 6 Conditional partial R- indicators for SCC and pilot

	SCC					Pilot					
	eligible	contact	able	coop-eration	response	response web/mail	eligible	contact	able	coop-eration	response
Age	*31*[(1)]	*40*	*26*	*29*	*24*	*37*	*21*	*31*	*29*	*16*	*13*
< 30	32	67	4	17	25	7	16	5	1	9	3
30-44	27	30	12	14	5	51	12	56	21	5	8
45-64	14	20	18	17	21	24	6	18	27	9	6
65>	23	41	31	35	6	52	10	15	36	1	1
no information available	1	1	2	10	2	0	0	1	1	4	1
Gender	*14*	*13*	*17*	*10*	*31*	*15*	*3*	*14*	*11*	*21*	*12*
Male(s)	9	9	8	4	40	15	0	11	8	21	6
Mixed	7	4	9	3	27	1	0	1	0	2	1
Female(s)	5	5	13	4	26	6	0	7	3	21	8
no information available	0	0	0	0	0	0	0	0	0	0	0
Household composition	*9*	*16*	*14*	*24*	*22*	*25*	*8*	*21*	*15*	*22*	*18*
Single	2	10	3	14	6	2	1	17	6	6	14
Partners, with children	1	5	4	7	4	26	2	7	11	11	5
Partners, no children	1	4	5	4	4	25	1	16	2	5	4
Single parent	2	3	2	16	4	2	3	1	3	21	11
no information available	2	3	4	19	31	7	0	2	1	3	1

*(1) Conditional R-indicators * 1000*

The equivalent of the conditional R-indicators would be a multivariate logistic regression. To illustrate the R-indices in table 6, table 7 shows the logistic model for the regression on response of all auxiliary variables for the SCC and the pilot. As representativeness in the RISQ project is defined as 'absence of predictable contribution', non-representative groups would show up in a standard (logistic) regression, as well as in the analysis of the conditional R-indicators. In the multivariate logistic regression on response 'Gender', 'Household' and 'Age' are selected in the model for the SCC, while 'Gender' and 'Income' are selected for the pilot. Again, conclusions of the two analyses are comparable: the larger the multivariate deviation from representativeness in the conditional R-indicators, the more the log odds in the logistic regression deviate from 1.

Table 7 Logistic model for the response propensity in SCC and pilot

SCC					Pilot				
Variable	*Category*	*β*	*Exp(B)*		*Variable*	*Category*	*β*	*Exp(B)*	
Gender	male(s)	-0,91	0,40	***	Gender	male(s)	-0,17	0,85	ns
(reference = mixed)	female(s)	-0,73	0,48	***	(reference = mixed)	female(s)	-0,36	0,7	***
	no info	-0,92	0,40	***		no info	-0,41	0,67	ns
Household	Partners, no children	0,02	1,02	ns	Income	2nd	0,14	1,15	ns
(reference = partners with children)	Single	0,16	1,18	ns	(reference = lowest quartile)	3rd	0,23	1,26	ns
	Single parent	-0,13	0,88	ns		highest quartile	0,44	1,55	ns
	No info	0,72	0,49	**		no info	0,55	1,74	ns
Age	30-45 y.o.a.	-0,22	0,80	*					
(Reference = 45-65 years of age)	less than 30 y.o.a.	-0,59	0,56	**					
	over 65 y.o.a.	-0,21	0,81	*					
	no info	-							

**** p < .001; ** p < .01; * < .05; ns not significant*

Income, incorporated in the logistic model of the pilot, also showed a somewhat larger deviation of representativeness in the conditional partial R-indicators than this variable in the SCC. This was caused by over-representation of the higher income group in the web / mail first round.

3.4 Costs

One of the aims of this pilot was to augment data quality while ideally diminishing costs. To compare the costs of the pilot to that of the control group, only the actual costs of observation and subsequent data processing (for the mail questionnaire) are considered. Two measures were taken that would introduce a substantial amount of costs saving: the use of a web round, and the larger share of day-time interviewing, as interviewers at Statistics Netherlands receive a 40% higher remuneration for evening work. Whether the use of a mail questionnaire would diminish costs, compared to a CATI version, was not clear beforehand, because of uncertainty concerning the number of respondents that would choose mail over web, and the subsequent amount of data handling necessary.

Without counting the allowance for evening work, the pilot turned out to be 18% cheaper than the regular SCC. Counting the 40% rise for evening work, the difference was 22%.

4. Discussion

In this paper an experiment is described, aimed at obtaining a better representative response, against lower costs. In the control group, a CATI Survey of Consumer Confidence (SCC) was held, using an uniform fieldwork strategy. In the experimental group, a mixed mode differential fieldwork strategy was deployed. Previous rounds of the SCC were used to calculate partial R-indicators, identifying groups that are over- or under-represented in contact and / or cooperation. A fieldwork strategy was designed to either stimulate or discourage contact and / or cooperation.

Results show that the differential fieldwork strategy was successful in maintaining the level of response, while significantly augmenting representativeness and at the same time substantially reducing costs. The R-indicators showed that representativeness was especially augmented as a result

of better representative eligible- and contacted cases in the pilot. The manipulation of cooperation had less impact.

Analysis of unconditional and conditional partial R-indicators made detailed consideration of the impact of the manipulations on different groups possible. For example, it was shown that the experimental manipulations had a large effect on the groups of whom no information was available. Also, young households were better represented as a result of measures to stimulate contact. The net representation in the response of the elderly households did not differ in the pilot and the control group, but the manipulations to dampen contact in this group, while at the same time stimulating cooperation, were visible in the partial R-indicators.

In this paper some attention is given to the conditional and un-conditional partial R-indicators versus traditional measures to express the relation between response (or contact, or cooperation) and (auxiliary) variables. While the conclusions reached by the R-analyses can be obtained by traditional response analysis, the R-indicators excel in ease of computation, wealth of information and concise presentation.

More detailed analyses will become available shortly on the website of the R-indicator, both of the R-indicator results and of the comparison with traditional non-response analysis. In addition, more attention will be given to the effects of the experimental manipulation and the mixed-mode aspects of the experiment.

REFERENCES

AAPOR (2006). Standard definitions: Final dispositions of case codes and outcome rates for surveys. 4th Edition. Discussion Paper Kansas: AAPOR, The American Association for Public Opinion Research.

Westat (2004). Blaise CATI guide. Rockville, MD.

Cobben, F., 2009, Nonresponse in Sample Surveys. Methods for Analysis and Adjustment. PhD-thesis. The Hague: Statistics Netherlands.

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 (3), 439-457.

Kleven, Ø, Holmøy, A. Fosen,J., Lagerstrøm, B. and Zhang, L.C. (2010). The use of R-indicators in responsive survey design – Some Norwegian experiences. Paper presented at the Q2010 Conference, Helsinki, May 3rd -6th.

Lin, I. and Schaeffer, N. (1995). Using survey participants to estimate the impact of nonparticipation. Public Opinion Quarterly, 49 (2), 236-258.

Loosveldt, G. and Beullens, K. (2009). RISQ – Fieldwork Monitoring. Available at http://www.risq-project.eu/papers

Luiten, A. and Wetzels, W. (2010). The use of R-indicators in data collection control. Results of a mixed mode pilot. Available at http://www.risq-project.eu/papers

Shlomo, N., Skinner, C., Schouten, B., Carolina, T., and Morren. M. (2009). Partial Indicators for Representative Response. Available at http://www.risq-project.eu/papers

Van der Grijn, F., Schouten, B., and Cobben, F. (2006). Balancing representativity, costs and response rates in a call scheduling strategy, Paper presented at 17th International Workshop on Household Survey Nonresponse, August 28-30, Omaha, NE, USA.