"Determinants of contact procedure duration and data richness: empirical analysis based on ESS Round 4 contact files"

Hideko Matsuo and Jaak Billiet
Center for Sociological Research, Katholieke Universiteit Leuven, Belgium

Contact address:

K.U. Leuven, Centrum voor Sociologisch Onderzoek (CeSO)
Parkstraat 45 Bus 3601
3000 Leuven
tel: +32 16 32 31 61
fax: +32 16 32 33 65

email: Hideko.Matsuo@soc.kuleuven.be

Abstract

Survey research has focused on achieving high response rates as low response rates are more likely to lead to non-response bias. This is why the European Social Survey has paid much attention to the development and implementation of targeted fieldwork specifications. This paper assesses the interplay between the number of contact attempts and other contact procedure and observable variables, the duration of the contact procedure, and data richness. The analysis in this paper consists of two parts. In the first part, the determinants of the contact procedure duration is studied. The data used come from ESS Round 4 contact files for Belgium and Poland. In this part, a proportional hazard model (competing risk model) is used including as co-variate the number of contact attempts, other contact procedure variables and neighbourhood characteristics. In the second part of this paper, it is assessed through test statistics, correlation and multiple regression analysis whether respondents collected through additional contact attempts enrich the responses in the dataset. Implications for the further specification of fieldwork efforts from the viewpoint of reducing non-response bias in cross-national research are discussed.

Key words
Contact procedure, survey cooperation, proportional hazard model, data quality, contact files

1. Background

Achieving high survey response rates is important because it reduces the likelihood of non-response bias. In many countries, survey response rates have decreased over time. For instance, compared to ESS Round 3, response rates decreased in half of the 24 participating countries in ESS Round 4.

The literature on non-response, and on efficient and effective strategies for reducing it, is wide-ranging. This is because maximizing response rates and minimizing response bias is costly – requiring, for instance, a higher number of interviewers, additional contact attempts, a longer fieldwork period – and requires carefully considered contact procedure efforts.

Empirical studies have identified fieldwork strategies amenable to improving survey outcomes. Some findings concern the number of contact attempts for selected sample units. Research based on ESS contact files for several rounds has shown that, generally speaking, increasing the number of contact attempts is effective. Each additional contact attempt results in an increase in the response rate (Matsuo et al., 2010, Figure 1). A negative linear relationship can usually be expected between the mean number of contact attempts and the non-contact rate. Similarly, Purdon et al. (1999) found for telephone surveys that in order not to exceed a 4 percent non-contact rate, a minimum of 4 telephone calls is required.

The timing of contact attempts also matters. Purdon et al. (1999) found that the most effective time to call is a weekday evening. An analysis of ESS Round 4 data also showed that first contact attempts on weekday evenings are more successful in many countries, though country variation was not negligible. Second and subsequent contact attempts are best made in weekends rather than on weekdays during working hours or in the evening.

The nature of the contact procedure (e.g. number, timing and mode of contact attempts) is strongly determined by the interviewer, who plays a key role in the sample unit's decision on whether to cooperate with the survey. Groves and Couper (1996) found that interviewers who change the content of their interactions with householders over successive contacts are likely to achieve somewhat higher cooperation rates. Snijkers et

al. (1999) found that the most effective interviewers were those who combined basic rules (e.g. identifying oneself clearly to the potential respondent), tactics (tailoring and adapting introduction), social skills and self-confidence. Another important factor is the early identification by the interviewer of soft or hard refusers. Depending on the characteristics of the refusal (soft versus hard refusal), interviewers can visit the sample unit later on, allowing for a cooling down period after the initial refusal, or request other interviewers to visit the sample unit.

The above literature does not specifically address which conditions favour a short contact procedure that encourages survey cooperation and achieves high data richness. In order to start filling this gap, the focus of this paper is on two research questions:

(1) What are the determinants (contact procedure and observable variables) of the duration of the contact procedure?;

(2) What is the effect of the number of contact attempts on response estimates (data richness)?.

The duration of the contact procedure is in this paper understood as the duration between the initial contact attempt and the final contact attempt. The final contact attempt usually generates the final survey outcome: interview, non-contact, other type of non-response and refusal. Only refusal gets priority coding if the last contact attempt resulted in non-contact or any other non-response outcome but had refusal during the contact procedure. .

Use is made in this paper of ESS contact files, which record the complete history (occurrence, timing, mode, and outcome) of the contact procedure and also the reasons for non-response. Observable information is also available from the contact forms: interviewers make their own assessments of the type of housing, the physical condition of the housing, the presence of litter and vandalism.

The next section discusses in more detail data contained in ESS Round 4 contact forms and the method of analysis. This is followed by a section on results from the two kinds of analyses undertaken. The paper ends with some conclusions and a discussion on future research.

2. Data and method

2.1. Data

ESS documents long-term attitudinal and behavioural change concerning a wide number of socio-political issues in European and neighbouring countries. It is also designed to capture the interaction between the changing institutions, attitudes, beliefs and behavioural patterns of Europe's diverse population. The survey covers over 30 countries. To achieve the highest methodological standards in cross-national and cross-cultural research, ESS is demanding with regard to sampling, question testing, event recording, translation and response rate enhancement procedures. These measures are stipulated in protocols.

This paper makes use of the main and contact files for ESS Round 4, which was implemented in 2008/2009 and in which 31 countries participated. The main file consists of respondents' answers to the questions. The contact files contain other kinds of unique information. They collect for all sample units, including respondents and non-respondents, information on the calling schedule and the contact outcomes, profile information on refusers, and dwelling and neighbourhood characteristics. This information is collected during the entire fieldwork period by the interviewer either on paper (PAPI) or by computer (CAPI).

In this paper, data are used for **two countries** participating in ESS Round 4: Belgium and Poland. These two countries were selected because:

- Their contact files were reasonably complete;
- Both countries used the same kind of sample frame (individual based);
- Both countries had different (contrasting) response rates – for Belgium 59 percent, for Poland 70.9 percent - which helps to test the effects of the contact procedure on the final response outcome;
- Both countries implemented their fieldwork in more or less the same period (November 2008 – January/February 2009).

Detailed information on the contact files for each of these two countries is presented in Table 1. For Belgium, there were 1760 respondents out of a total of 3060

sample units. For Poland, there were 1614 respondents out of a total of 2428 sample units. Non-contact rates were low in both countries: for Belgium 2.25 percent, for Poland 1.58 percent. However, refusal rates differed: 25.95 percent for Belgium versus 18 percent for Poland (based on Matsuo et al., 2010).

2.2. Method

In the first part of the analysis, use is made of a proportional hazard model, which models the survival time in the contact procedure specifying the dependence on the covariates of exit time. In the second part of the analysis, test statistics (for studying differences between response distributions for initial respondents and for later respondents) and correlation and multiple regression (for studying the effect of the number of contact attempts on respondents' answers to selected questions) are used.

In the first part of the analysis, a proportional hazard model, more specifically a competing risk model, is applied. This means that several possibilities exist for the mode of exit and censoring: (1) interview; (2) refusal; (3) other type of non-response. We look at whether the effects of different covariates differ by type of event. The covariates include the number of contact attempts (the main variable we are interested in), a number of other contact procedure variables (number of face-to-face contacts, number of evening and weekend contacts, number of refusals), and 4 types of observable data (type of housing; physical condition of the housing; presence of litter; presence of vandalism). In order to test country-specific effects, a country dummy is included.

The dependent (time) variable is operationalized as the duration of the complete contact procedure until the final outcome measured in weeks, so the number of completed weeks between the initial and the last visit/contact attempt of the interviewer to the sample unit. For each country, the beginning of the observation period coincides with the beginning of the fieldwork period (Belgium 13 November 2008, Poland 3 November 2008). The observation period ends 60 days later. This means that for some cases, the event occurs within the observation period and for some case it does not. In the case of non-contact, the observation is extended until the end of the observation period: it is assumed that non-contact cases had no contact with the interviewer during the

observation period though in reality the non-contact status is determined at a much earlier stage.

In the second part of analysis, the effect of the number of contact attempts on data richness is tested. We focus on the difference between the sample obtained through the first contact attempt and the sample obtained through the second and later contact attempts (can expand here comparing 1-3 contact attempts versus 4+ contact attempts). For this purpose, we use test statistics (X^2 or differences between *t-tests*). In addition, correlation and multiple regression are used. Following established approaches (Billiet et al., 2007 and Vehovar, 2007), 'sensitive' questions (primary suspects) are selected to test for differences. These questions are: immigrants make country worse or better place to live (*imwbcnt: 11 scale*); take part in social activities compared to other of same age (*sclact: 5-scale*); subjective general health (*health: 5 scale*) and how interested in politics (*polintr: 4-scale*). The model contains the number of contact attempts, the main variable we are interested in plus; key socio-demographic variables (age; gender; education; partnership status; job status; household composition), a number of other contact procedure variables (number of weeks and number of interviewers) and the 4 aforementioned observable variables. These variables are studied in two steps: (1) the full model (model 1); (2) restricted socio-demographic model (model 2) and (3) restricted contact procedure model (model 3).

3. Results

3.1. Mean number of weeks and mean number of contact attempts

Before discussing the main results of the analyses, some additional background is provided for both countries. Table 2 presents univariate statistics on number of contact attempts and duration (weeks/days) for all final status events (interview, non-contact, refusal, other type of non-response). These values are calculated on the basis of a 60-day observation period as discussed above concerning 2140 sample units (total: N=3060) for Belgium and 2384 sample units (total: N=2428) for Poland.

In general, overall mean duration are longer and the overall mean number of contact attempts higher for Belgium than for Poland. As for overall durations, the means are 1.9 weeks/15.44 days for Belgium, and 1.34 weeks/10.98 days for Poland. The overall

mean number of contact attempts is 2.8 times for Belgium and 2.09 times for Poland. Mean values differ by final outcome statuses. For interview cases, mean durations are 1.08 weeks (9.65 days) for Belgium and 0.91 weeks (7.77 days) for Poland. The mean number of contact attempts for interview cases is 2.77 times for Belgium and 2 times for Poland. Mean values for non-contact cases were 4.24 weeks (32.65 days) and 5.19 contact attempts for Belgium and 5.78 weeks (42.42 days) and 3.74 contact attempts for Poland.

3.2. Determinants of contact procedure duration

As explained above, the dependent (time) variable in the competing risk model is the number of weeks between the initial and the last contact attempt for each mode of exit during the observation period (60 days). Hazard ratios are obtained through the exponential value of the *beta* coefficient (Exp beta). The hazard ratio describes the effect of an independent variable on the risk of being in each outcome category. A hazard ratio above 1 indicates that the risk of being in the outcome category is *increased* by the independent variable. When the hazard ratio is below 1, a *decrease* in the risk of being in the outcome category is indicated.

In the first model presented in Table 3 (which treats all events in the same manner), the number of contact attempts has significant effect: it has a strong (p<.0001), risk decreasing effect: one additional contact attempt decreases the risk of experiencing any kind of event by 14 percent. Four other variables are also significant: living in apartment (type of housing); country variable; number of face-to-face contacts; and number of refusals. Two out of these four variables have strong effects (p<.0001): living in apartment decreases the risk of experiencing any kind of event by 15 percent; living in Belgium compared to living in Poland decreases this risk by 10 percent; one additional refusal decreases this risk by 54 percent. Additional moderate effect (p<.05) is found for additional face-to-face contacts decreasing by 5 percent the risk.

The results of the restricted model including only the covariates for the contact procedure (model 2) are broadly consistent with the results of model 1, although the variable number of face-to-face contacts is no longer significant.

A similar trend is found for the model when the hazard is event of interview treating 'other types of non-response' and 'refusal' as censored cases. Additional refusals decrease the risk substantially compared to the first model. Additional face-to-face contacts increase the ratio beyond 1 implying a higher risk of interview, but this result is not significant.

When we focus on the hazard for refusal, a strongly decreased risk is found for additional contact attempts (HR: 0.790; p<.001), as well as for additional face-to-face contacts (HR: 0.885; p<.01). At the same time, a strongly increased risk is found for an additional number of refusals (HR: 2.233; p<.0001) and for additional evening and weekend contacts (HR: 1.156, p<.0001). Neither living in an apartment, nor the country one living in is important. Additional evening and weekend contacts are important in both models. None of the other models had such a significant effect.

When we focus on the hazard for other type of non-response[1] (treating interview and refusal as censored cases), a significant, strong, risk decreasing effect is found for additional contact attempts. Results are significant for other variables as well: type of housing; presence of litter; country; and refusals. Strong effects (p<.0001) are found for country and refusals; and moderate effects for living in apartment (p<.01) and living in littered neighbourhood (p<.05). Those who live in Belgium are 3.4 times as likely as those living in Poland to exit by this outcome category. Those living in an apartment and those living in littered neighbourhoods have 40 percent and 63 percent greater risk of ending up in this outcome category. Those with additional refusals have a strongly decreased risk for ending up in this outcome category.

Overall, the results obtained via the competing risk model show that the number of contact attempts is an important variable associated with overall contact procedure duration, i.e. with survival time to any of the possible outcome categories, and with the hazard for interview (strongly decreased risk). Other important variables for the overall contact procedure duration are the type of housing, the country, and additional refusals. Focusing on the hazard for interview, those living in apartments and those living in

---

[1] This includes the following final contact status: contact form missing; partial interview: break-off; invalid interview; broken appointment; respondent not available/away; respondent mentally or physically not available; language barrier; contact but no interview, other; respondent moved to unknown destination; respondent has moved, still in country; address not traceable; and not attempted.

Belgium also have strongly decreased risks. Results are different for refusals and other types of non-response. Additional refusals and additional evening and weekend contacts are associated with a higher risk for refusal. Living in an apartment, in a littered neighbourhood and in Belgium results in a strongly increased risk of experiencing another type of non-response.

3.3. The effect of the number of contact attempts on data richness (response estimates)

In the second part of the analysis, differences in responses between different sample groups (first contact attempt vs. 2+ contact attempts) are analysed through test statistics (X^2 or differences between means t-test in the first place). The effect of the number of contact attempts (including socio-demographic variables of the respondent) on the respondent's responses is analyzed via bi-variate correlations and multiple regression analysis. Table 4 illustrates the differences between the sample obtained through the 1st contact attempt and the sample obtained through the second and subsequent contact attempts. Means, standard deviations, Pearson correlations and Cronbach's alpha estimates are presented in Table 5.

Comparing response distributions between initial respondents and later respondents based on the number of contact attempts, differences are found for a number of variables for both countries: age, job status and type of housing. In addition, differences are found for the presence of litter and vandalism for Belgium and for education and subjective health for Poland.

Table 6 presents the results of the multiple regression. Several so-called 'sensitive' attitudinal variables (immigration, social participation, subjective health and political interest) (primary suspects) were regressed on the number of contact attempts including several socio-demographic and observable variables. The equation containing these variables accounted for a wide range of the variance observed for each of these attitudinal variables implying that fitting the model differs by variable of interest. For example, R^2 for immigration variable was 0.05 and 0.04 for Belgium and Poland respectively, while R^2 for political interest was 0.14 and 0.36 respectively. Bèta weights

(standardized multiple regression coefficients) were studied to assess the relative importance of the number of contact attempts for predicting these variables.

Starting with Belgium, we first look at the immigration variable. No effects are found for the number of contact attempts. Effects are found, however, for the following predictors: educational level, presence of litter, and presence of vandalism. For educational level, the bèta weight is 0.154 (p<.0001), while for the presence of litter and the presence of vandalism, the beta weight is -0.085 (p<.05) and 0.072 (p<.05) respectively. Even in the restricted model, the number of contact attempts has no significant effects, implying that there is no added value in undertaking additional contact attempts in order to enrich the dataset. Overall, it is found that higher educated people have more positive views on immigration, while respondents living in areas with litter have more negative views on immigration.

For social participation, a significant, positive effect is found for the number of contact attempts. Significant bèta weights are found also for age, education, job status, presence of litter, and presence of vandalism. The effects of both education and presence of vandalism are positive, the former more than the latter. The effects for age, job status and presence of litter are negative, the effect of the last one being the largest. The restricted 'contact procedure' model confirms the importance of the number of contact attempts: respondents recruited after a higher number of contact attempts are socially more active than those recruited after just one contact attempt.

For subjective health, no effect is found for the number of contact attempts. Effects are found, however, for a number of socio-demographic variables. While a higher age and living in a neighbourhood with litter is negatively associated with the subjective health status, having a job and higher education are positively associated.

For political interest, the number of contact attempts has no effects. When socio-demographic variables are included in the model, Table 6 shows that results are significant and positive for education and living with a partner and significant and negative for being female.

Turning to Poland, generally speaking the number of contact attempts matters more than in Belgium. Starting with the immigration variable, significant, positive effects are found for the number of contact attempts in both the full and the restricted model,

pointing to enrichment of the response estimates in the data set. In the full model, significant effects are also found for other variables: age, education, living with partner, and living in an apartment. These effects are positive for education, living with partner, and living in an apartment, and negative for age.

For social activities, only a partial effect is found for the number of contact attempts: only in the full model, a significant, negative effect is found. In the full model, significant, positive effects are found for education and living in an apartment, and significant, negative effects for age and gender.

The (partial) importance of number of contact attempt is clear from the comparison of responses on subjective health. A significant, positive effect is found in the restricted model only. In addition, negative effects are found for age (strong), gender (female) and living in a neighbourhood with vandalism, and positive effects for education (strong) and being employed.

The results for political interest are similar to those for subjective health: a significant, positive effect is found in the restricted model only. In the full model, significant, positive effects are found for age, education (stronger than for other variables), and living in an apartment. Significant, negative effects are found for living with partner and gender (female).

Generally speaking, the above results show that number of contact attempts is important on responses for quite a number of variables, pointing to the fact that additional contact attempts can enrich the responses (dataset)[2]. The effects are stronger for Poland than for Belgium. For other variables, in sum, out of 13 predictors, education (4 variables), age, job status, presence of litter (3 variables) are important predictors in Belgium; and age and education (4 variables) and number of contact attempts (3 variables) in Poland.

---

[2] It should be also noted that it is not automatically enriching because this effect differs by variables.

4. Conclusion and Discussion

The objective of this paper was to contribute to the literature on the favourable conditions to a short contact procedure that encourages survey cooperation and achieves high data richness. Within this context, the focus of this paper was on two research questions: (1) What are the determinants of the duration of the contact procedure; (2) What is the effect of the number of contact attempts on response estimates (data richness). Use was made of ESS Round 4 contact files and to answer the first question, use was made of a competing risk model, while to answer the second question, use was made of test statistics, correlation and multiple regression analysis.

Overall, the results obtained through the competing risk model showed that the number of contact attempts is an important variable associated with overall contact procedure duration, i.e. with survival time to any of the possible outcome categories, and with the hazard for interview (strongly decreased risk). Generally speaking, the results from the correlation and multiple regression analysis showed that responses for quite a number of selected 'sensitive' variables (primary suspects) are influenced by the number of contact attempts, pointing to the fact that additional contact attempts enriched the responses (dataset). The effects were stronger for Poland than for Belgium.

Our results point to the important effect of number of contact attempts on the overall duration of the contact procedure. At the same time, it demonstrates the added value of undertaking additional contact attempts in terms of enriching the dataset.

These results need to be confirmed by further research applied to many other ESS participating countries. Such research could focus on improved operationalisations of the proportional hazard model, which is not completely free of problems. Strong assumptions are made. It is assumed, for instance, that there is no automatic link between the number of contact attempts and the duration of the fieldwork. However, in practice, there may be a 'standard' duration associated with each contact attempt as the timing of contact procedure events depends on the country-level organisation of the fieldwork and the individual interviewer, differences across countries and interviewers that deserve further exploration. Also, specific assumptions are imposed for non-contacts. For these cases, a

maximum duration of 60 days is assumed even through the non-contact status is usually determined much earlier.

6. References

Billiet, J., M. Philippens, R. Fitzgerald & I. Stoop (2007) 'Estimation of Nonresponse Bias in the European Social Survey: Using Information from Reluctant Respondents'. In: *Journal of Official Statistics*. 23, 2, 135-62.

Groves, R. M. and M. P. Couper (1996) Contact-level influences on cooperation in face-to-face surveys. In: *Journal of Official Statistics*. 12, 1, 63-83.

Matsuo, H., J. Billiet & G. Loosveldt (2010) Response-based Quality Assessment of ESS Round 4: Results for 24 countries based on Contact Files. CeSO: Working paper CeSO/SM/2010-1, 71 pp.

Pratesi, LM., K. L. Manfreda, S. Biffignandi and V. Vehovar (2004) List based web surveys: quality, timeliness, and nonresponse in the steps of the participation flow. In: *Journal of official statistics*. 20, 3, 451-65.

Purdon, S., P. Campanelli and P. Sturgis (1999) Interviewers' calling strategies on face-to-face interview surveys. In: *Journal of Official Statistics*. 15, 2, 199-216.

Snijkers, G. J. Hox and E. de Leeuw (1999) "Interviewers' tactics for fighting survey nonresponse". In: *Journal of Official Statistics*. 15, 2, 185-98.

Vehovar, V (2007) 'Non-response bias in the European Social Survey'. In: G. Loosveldt, M. Swyngedouw and B. Cambré. Eds., *Measuring meaningful data in social research*. Acco publisher, 335-56.

Figure 1: Obtained response rates (%) after 1st, 2nd, 3rd, 4th and more contact attempts
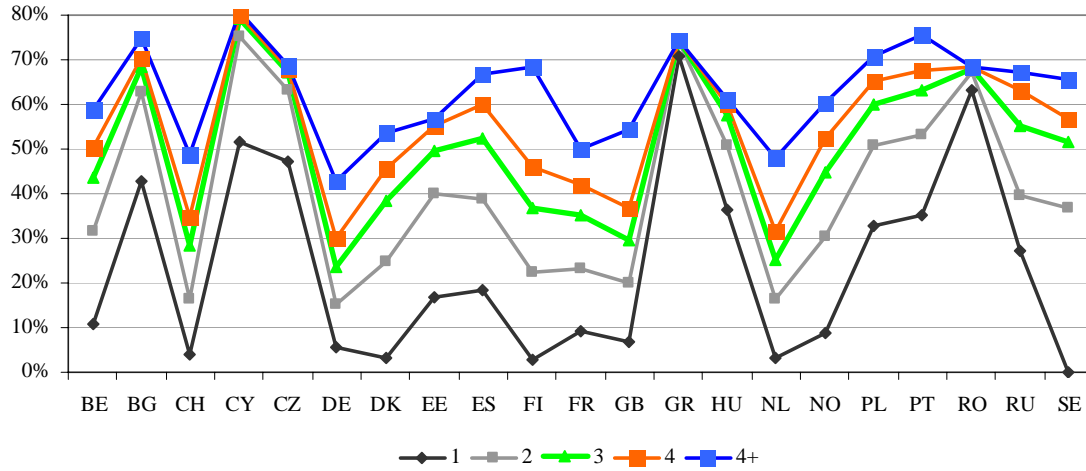
(Source: Matsuo, Billiet and Loosveldt 2010)

Figure 2: Scatter plot between non-contact rates and mean number of contact attempts
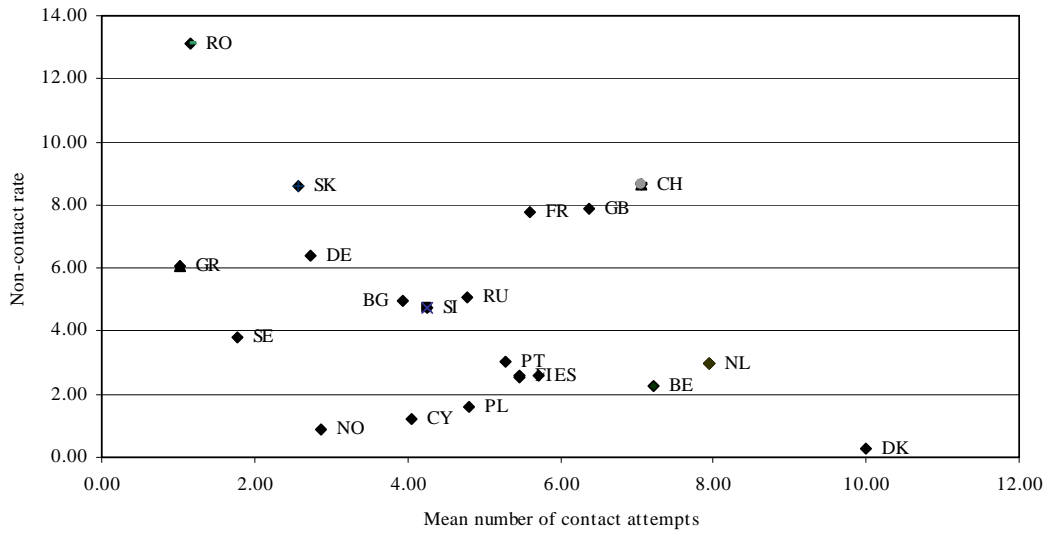
(Source: Matsuo, Billiet and Loosveldt 2010)

Table 1: Response patterns and characteristics of contact procedure in Belgium and Poland

	Belgium	Poland
Fieldwork period	13/11/2008 -20/3/2009	03/11/08 – 15/02/09
Type of sample	Individual named	Individual named
Interviewed & total samples	3060/1760	2428/1614
Ineligible cases (%)	2.52%	6.18%
Response patterns		
Completed interview (N/%)	1760/59	1614/70.85
Non-contact (N/%)	67/2.25	36/1.58
Refusal (N/%)	774/25.95	410/18
Other type of non-responses (N/%)	382/12.81	218/9.57

Table 2: Univariate statistics on durations and contact attempts during the 60 day observation period

	Belgium			
All types (N=2140)	*mean/median*	*SD*	*minimum*	*maximum*
Days	15.44/5	18.94	0	60
Weeks	1.9/0	2.61	0	8
Contact attempts	2.8/2.0	1.95	1	13
Interviewed (n=1278)				
Days	9.65/4.0	13.68	0	59
Weeks	1.08/1.0	1.88	0	8
Contact attempts	2.77/2.0	1.78	1	13
Non-contact (n=37)				
Days	32.65/39.0	17.8	0	60
Weeks	4.24/5.0	2.44	0	8
Contact attempts	5.19/6.0	2.65	1	10
Refusal (n=528)				
Days	30.2/36.0	22.2	0	60
Weeks	3.95/5.0	3.07	0	8
Contact attempts	2.77/2.0	1.93	1	11
Other non-response (n=251)				
Days	12.66/2.0	17.38	0	58
Weeks	1.56/0	2.32	0	8
Contact attempts	2.73/2.0	2.35	1	10
All types (N=2381)	Poland			
Days	10.98/1.0	16.01	0	59
Weeks	1.34/0	2.16	0	8
Contact attempts	2.09/2.0	1.35	1	7
Interviewed (n=1595)				
Days	7.77/0	13.32	0	59
Weeks	0.91/0	1.78	0	8
Contact attempts	2.0/2.0	1.3	1	7
Non-contact (N=34)				
Days	42.42/44.0	9.77	13	57
Weeks	5.76/6.0	1.48	1	8
Contact attempts	3.74/3.0	1.814	1	7
Refusal (N=393)				
Days	20.73/18.0	18.42	0	56
Weeks	2.6/2.0	2.51	0	8
Contact attempts	2.45/2.0	1.38	1	7
Other non-response (N=211)				
Days	14.06/2.0	17.97	0	57

| Weeks | 1.81/0 | 2.44 | 0 | 8 |
| Contact attempts | 2.21/2.0 | 1.48 | 1 | 7 |

Table 3: Effects of contact procedure and observable variables on the survival time (completed weeks) of different events via competing risk model

	Events on all types		Events on interview		Events on refusal		Events on other NR types	
Belgium (N=2140)/PL (N=2381)	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Type of house *0: other; 1: apartment*	0.858***		0.779***		1.115		1.405**	
Presence of litter *0: not common; 1: common*	1.057		0.919		0.742		1.631*	
Presence of vandalism *0: not common 1: common*	0.995		1.022		0.702		1.481	
Physical condition of house *0: Very good/good; 1: Satis., bad & very bad*	0.910		0.893		1.098		1.259	
Country *0; Poland; 1=Belgium*	0.906**	0.963	0.847**	0.970	0.956	0.928	3.437***	1.847***
# attempts	0.763***	0.751***	0.721***	0.711***	0.790***	0.791***	0.728***	0.722***
# face to face contacts	0.958*	0.968	1.010	1.042	0.885**	0.871**	0.927	0.892*
# evening & weekend contacts	0.999	1.004	0.959	0.951*	1.156**	1.149**	1.062	1.128*
# refusals	0.461***	0.464***	0.065***	0.067***	2.233***	2.249***	0.008***	0.007***
Likelihood ratio; df; p-value	*140.79; 9; p<.0001*	*1433.67; 5; p<.0001*	*2575.32; 9; p<.0001*	*2426.55; 5; p<.0001*	*580.02; 9; p<.0001*	*605.11; 5; p<.0001*	*427.46; 9; p<.0001*	*503.14; 5; p<.0001*

***p-value<.0001; **p-value<.01; p < .05

Table 4: Responses to key questions in main survey comparing initial obtained sample (contact attempts=1) and later obtained sample (contact attempts 2+) in Belgium and Poland

	Belgium (n=1278)	Poland (n=1595)
Socio-demographic variables		
Age X^2 ; *df = 5*; *prob.*	16.2039; **p=0.0063**	14.82; **p=0.0112**
Gender X^2 ; *df =1*; *prob.*	1.3448; p=0.2462	0.0334; p=0.855
Education X^2 ; *df = 2*; *prob.*	2.8461; p=0.2410	23.4796; **p<.0001**
Living partner X^2 ; *df = 1*; *prob.*	0.007; p=0.9333	0.5991; p=0.4389
Job status X^2 ; *df = 1*; *prob.*	8.5085; **p=0.0035**	5.9165; **p=0.015**
HH composition X^2 ; *df =1*; *prob.*	0.0744; p=0.7851	0.2335, p=0.6289
Neighborhood variables		
Type of house X^2 ; *df = 1*; *prob.*	1.4630; **p=0.0413**	25.9675; **p<.0001**
Litter X^2 ; *df = 1*; *prob.*	11.6369; **p=0.0006**	0.3642; p=0.5462
Vandalism X^2 ; *df = 1*; *prob.*	5.1732; **p=0.0229**	1.9271; p=0.1651
Physical condition X^2 ; *df=1;prob.*	1.8971; p=0.1684	0.3507; p=0.5537
Attitudinal variables		
Immigration X^2 ; *df = 10*; *prob; T-value; df; prob.*	9.7934; p=0.4588 T=0.04; df=1259; p=0.9706	8.5005; p=0.5801 T=1.62; df=1468; p=0.1056
Social Activities X^2 ; *df =4*; *prob.; T-value; df; prob.*	0.8777; p=0.9277 T=0.77; df=1269; p=0.4394	4.3471; p=0.3611 T=0.70; df=1535; p=0.4869
Subjective health X^2 ; *df=4; prob.; T-value; df; prob.*	4.9708; p=0.2903 T=1.52; df=1275; p=0.1289	8.1829; p=0.0851 T=2.58; df=1589; **p=0.0099**
Political interest X^2 ; *df=3; prob.; T-value; df; prob.*	1.7339; p=0.6294 T=-1.15; df=1276; p=0.2509	3.3545; p=0.3401 T=1.47; df=1588; p=0.1407

Table 5: Bivariate correlation between predictors and 4 variables of interest

Belgium	immigration	Social activities	Subjective health	Political interest
Age	-0.09**	-0.102**	-0.312***	0.064*
Education	0.179***	0.154***	0.199***	0.326***
Living partner	0.049	-0.041	-0.006	0.140***
Job status	0.091**	0.007	0.251***	0.132***
Gender	-0.011	-0.019	-0.032	-0.147***
HH composition	0.075**	0.025	0.104**	0.009
#attempts	0.030	0.052	0.031	0.002
#weeks	0.001	-0.013	-0.015	-0.031
#interviewer	-0.06*	-0.040	-0.045	-0.049
Type of house	0.031	-0.0121	0.008	0.019
Litter	-0.041	-0.076**	-0.042	-0.095**
Vandalism	0.029	-0.004	0.015	-0.061*
Physical condition	-0.011	-0.0508	-0.009	-0.048
Mean (SD)	4.98 (2.05)	2.67 (1.07)	2.93(0.78)	1.37(0.89)
Poland				
Age	-0.110***	-0.181***	-0.57***	0.153***
Education	0.106***	0.157***	0.266***	0.273***
Living partner	-0.005	-0.019	-0.080**	0.151***
Job status	0.098**	0.118***	0.253***	0.112***
Gender	-0.026	-0.064**	-0.08**	-0.204***
HH composition	0.021	0.084***	0.173***	-0.027
#attempts	0.051***	-0.0005	0.11***	0.041
#weeks	-0.05*	0.021	0.08**	0.003
#interviewer	-0.032	0.012	0.04**	0.047
Type of house	0.096**	0.076**	0.002	0.103***
Litter	-0.02	-0.019	-0.025	-0.075**
Vandalism	0.01	-0.045	-0.077**	-0.051*
Physical condition	-0.003	-0.06*	-0.081**	-0.058*
Mean (SD)	5.98 (1.98)	2.57 (0.911)	2.62 (0.92)	1.30 (0.85)

p-value<.0001; **p-value<.01; p < .05
Socio-demographic variables are coded as follows.
- **Age and education is numeric;**
- **Gender (0=male; 1=female);**
- **Living with partner (0=no; 1=yes);**
- **Job status (0=unemployed;1=employed);**
- **Household composition (0=living alone; 1=living with more than 1)**

Contact procedure variables are coded as follows.
- **Number of contact attempts and weeks are both numeric;**
- **Number of interviewers (0=one interviewer; 1=more than one)**

Observable variables are coded as follows.
- *Type of house (0: other; 1; apartment)*
- **Presence of litter (*0: not common; 1: common*)**
- **Presence of vandalism (*0: not common 1: common*)**
- **Physical condition of house (*0: Very good/good; 1: Satis., bad & very bad*)**

Table 6: Effects of number of contact attempts and other variables on 4 variables of interest

	Belgium			Poland		
(immigration)	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Age	-0.061	-0.065		-0.087**	-0.09**	
Education	0.154***	0.155***		0.059*	0.054	
Living partner	0.042	0.032		0.005*	0.009	
Job status	0.004	0.004***		0.054	0.053	
Gender	-0.028	-0.007		-0.017	-0.018	
HH composition	0.049	0.039		-0.009	-0.006	
# attempts	0.017		0.030	0.063*		0.107***
# weeks	0.008		0.020	-0.124**		-0.121**
# interviewer	-0.039		-0.058	0.002		0.009
Type of house	0.028	0.028		0.087**	0.088**	
Litter	-0.085*	-0.08*		-0.021	-0.022	
Vandalism	0.073*	0.075*		0.011	0.009	
Physical condition	0.003	0.007		0.006	0.008	
R²	**0.051**	**0.045**	**0.003**	**0.04**	**0.029**	**0.04**
(soc. activities)						
Age	-0.072*	-0.09*		-0.134***	-0.12***	
Education	0.165***	0.17***		0.108**	0.101**	
Living partner	-0.05	-0.044***		-0.034	-0.032	
Job status	-0.086**	-0.083*		0.042	0.037	
Gender	-0.035	-0.03		-0.055*	-0.054*	
HH composition	0.010	-0.009		0.05	0.05	
# attempts	0.099**		0.088*	-0.083**		-0.016
# weeks	-0.077		-0.028	0.029		0.031
# interviewer	0.004		-0.028	-0.015		-0.005
Type of house	-0.023	-0.019		0.089**	0.075**	
Litter	-0.109**	-0.103**		0.027	0.029	
Vandalism	0.077**	0.063		-0.045	-0.047	
Physical condition	-0.031	-0.024		-0.038	-0.04	
R²	**0.052**	**0.047**	**0.006**	**0.067**	**0.067**	**0.0006**
(subj. health)						
Age	-0.275***	-0.264***		-0.541***	-0.538***	
Education	0.095**	0.104**		0.124***	0.122***	
Living partner	-0.001	0.004		0.019	0.02	
Job status	0.1274***	0.124**		0.058*	0.056**	
Gender	-0.005	-0.004		-0.053*	-0.053**	
HH composition	0.002	0.006		-0.021	-0.021	
# attempts	-0.027		0.023	-0.04		0.088*
# weeks	0.010		0.034	0.033		0.043
# interviewer	-0.022		-0.055	-0.006		-0.010
Type of house	0.005	0.012		0.006	0.00006	
Litter	-0.088*	-0.08**		0.048	0.048	

Vandalism	0.047	0.04		-0.058*	-0.058*	
Physical condition	0.005	0.006		-0.034	-0.034	
R²	**0.144**	**0.140**	**0.003**	**0.364**	**0.362**	**0.013**
(pol. interest						
Age	0.055	0.069*		0.202***	0.202***	
Education	0.319***	0.327***		0.311***	0.31***	
Living partner	0.112**	0.105**		0.077**	0.078**	
Job status	-0.001	-0.059**		-0.0002	-0.0004	
Gender	-0.14***	-0.143***		-0.232***	-0.233***	
HH composition	-0.06	-0.059		-0.024	-0.024	
# attempts	0.019		0.027	0.023		0.059**
# weeks	-0.067		-0.032	-0.060		-0.070**
# interviewer	0.012		-0.026	0.023		0.069**
Type of house	0.003	0.008		0.050*	0.051*	
Litter	-0.047	-0.05		-0.05	-0.048	
Vandalism	-0.017	-0.015		-0.004	-0.006	
Physical condition	0.021	0.013		-0.03	-0.03	
R²	**0.145**	**0.150**	**0.002**	**0.188**	**0.186**	**0.005**

***p-value<.0001; **p-value<.01; p < .05