Pattern of contacts attempts in surveys

Alexandre Pollien and Dominique Joye, FORS, Lausanne

Introduction

Survey specialists are more and more concerned about an explosion of non-responses in the surveys; even if efforts to reach people have increased. In addition, changes in technology, transition from fixed to mobile telephony, Internet development, have greatly complicated the work of contacting. In this context, methodologists are considering alternative forms of surveys, as mixed mode, or new indicators of quality using for example paradigm of "Total Survey Error" which take into account the balance of costs and various errors related to the response.

The perspective developed here differs from this body of researches by proposing to rehabilitate a sociological approach to deal with survey participation. Three dimensions should be considered:

1. The surveys are conducted in a society at a given time. This involves to consider a "climate survey" which look at legitimacy of the various actors. Social representations of surveys and research institutions can encourage or discourage propensity to cooperate. In this respect, trust in the confidentiality of data use is crucial. The theme of the survey may be more or less recognized as worthy in public opinion. In other words, participation depends on the social representation of the survey.

2. Cooperation depends on the posture of the respondent on the interview's situation that is also related to his social position. The feeling of ability to express oneself in the quasi-public sphere of the survey is related to self-esteem and sense of legitimacy to give his opinion or willing to access to speech in a given context. Factors of cooperation are various inducements to cooperate, as saliency of the proposed theme and monetary incentives. There are also psychological motivations to cooperate, as the weight of social authority of the interviewer versus the propensity to respond favourably to requests.

3. Finally participation depends on the investigation process which consists in shorter or longer series of interactions between respondent and interviewer. We can consider the interactions, as situated in a specific social register. The series of contact attempts reflect the probability of being contacted and of having agreed to cooperate, depending on strategies of the interviewer. Our hypothesis is that the sequence of contact is related partly to lifestyle of the respondent. The appropriate time to be contacted and to cooperate is sociologically defined.

The conceptualization proposed here shows the importance of exploring a set of tracks to explain and prevent nonresponse and correct the data. Many studies have shown that inference of nonresponse with information restricted on respondents, remains doubted (Stoop et al., 2010). Imputation of non-response involves having information on the entire sample. In this perspective, we propose in this paper to explore a way of assessing the quality of the sample based on the sequences of contact attempt between interviewer and respondent.

Leverage-saliency theory vs. sociology of contacts process

The construction of representation of the population follows various steps including the way by which someone designed as a fraction of the population ends up face to an interviewer and will take part to the survey. The process, which leads from a target address to an interview, constitutes the contacts process. This phase of the survey is made up of a series of actions accomplished by the interviewer in order to find and to convince the sample unit to cooperate.

The various elements composing the procedure of contact constitutes processual chain: the interviewer gradually discovering social and spatial environment, the dwelling house, the type of household and finally the selected person. In the other side, the respondent receives a notification letter, then a personal visit at a more or less appropriate time, hears the request, and finally, after eventual refusals and postponements, agrees to participate... or not.

The literature on survey methodology distinguishes two dimensions, the propensity to be contacted and the propensity to cooperate in order to identify probable bias (Lynn et al. 2000). It is now established that the difficulties to be reached characterizes rather socio-demographic profiles (free time available to the activity of interviewing) (Keeter et al. 2000) whereas the reluctance to cooperate is related to attitudinal profiles and social integration (social interest or opening, social participation) (Groves et al. 2004).

Considering these problems, some authors focus on the paradata (Kreuter, 2009). The propensity to cooperate is studied in connection with environment variables, as type of housing, or characteristics of the interviewer. The sequence of contacts reveals the level of accessibility or cooperation of the respondent. This perspective consists in considering the whole sequence of successive states encountered. The sequences of contacts attempt is not a simple mean to lead to an interview or not, but an itinerary that express the accessibility and the propensity to cooperate across a work of contacting and lobbying.

The aim of this study is to understand and define the nature of these sequences. This perspective must go beyond the contact sequence as a pure reflection of the respondent: "Tell me how I have obtained your participation; I will tell you who you are". Our assumption is therefore that the

methodological characteristics of a survey and strategies of interviewer determine the structure of typical sequences of contact.

Handling the sequences of contacts attempts

Following the *European social survey*, standards of contact procedures have been set for all our face to face surveys: ESS and MOSAiCH.

Method

Each interviewer received a number of addresses. He sends them each a notification letter, and then attempts to contact them in person. If the respondent refuses to participate, the first phase of contact procedure will stop. The second phase is called "refusal conversion". The addresses are reallocated to interviewers. To simplify our analysis, we stopped at the first phase of contacts, so there is one interviewer by address. To build the sequences of contacts attempts, we used the ESS 2010 and MOSAiCH 2012.

The first problem is the high complexity and variance of the sequence. Various transformations and reductions have been made from the raw data in order to obtain synthetic and homogeneous codification of the outcome of each contact sequences.

The successive similar contacts in a same day have been reduced. A codification has been elaborate to simplify the outcome:

Types of outcome:	Participant:	Abbrev.
Non-contact		NC
Appointment made with…	household or proxy	AP
Refusal by…		RP
Appointment made with…		AR
Refusal by…	respondent	RR
Interview of…		IN
Respondent unavailable		UN

The situations, where the interviewer does not take part, does not give an appointment, does not refuse, were reduced to the non-available code. This code relates to the ineligible, invalid addresses, sick or invalid people and person who don't speak a language available in the survey. (French, German, Italian).The codes distinguish also the participant: undefined person of the household or relative of the respondent (proxy) versus the respondent himself, as well as the outcome of the contact attempt: interview, refusal, appointment or other.

The difficulty is to manage this huge heterogeneity. We perform an optimal matching analysis, with "TRATE" substitution-cost matrix and a cost of insertion/deletion of 1 (Gabadinho et al., 2010). The distances matrix between sampled cases has been clustered with the Ward Method.

The typology was tested on the registry variables, environmental and nonresponse survey.

Another experiment was conducted. We measured the distance to the "perfect sequence" which is the simple sequence that leads directly to an interview. The data have been interpreted as a score of propensity to participate. We have therefore tested the effect of such weighting on the control data.

Conclusion

This perspective is exploratory. Nevertheless, it promises to better understand the chance of success for an interview. It could also be used to compare interviewer's strategies and to show the effect of methodology in the comparison between countries. There is the possibility to view this information in relation with the other informations that we have about sampled units (environment and registers data). This method could help us to better understand the response process.

But it seems very difficult to rectify a sample using an indicator of these sequences, because the sequences of contact are related to both the respondent and interviewer. Furthermore, from the perspective of sequence analysis, the richest information are provided by the longest sequences that are, in fact, less representative of a normal procedure, without misunderstanding, error or problems.

The sequence reflects "social relationships" with target which can also exist for themselves, for example through series of appointments, through refusal which does not interrupt the process, through interviewer strategy of keeping in touch expecting a later cooperation, or through avoiding unpromising target.