

# **Practical Strategies for Minimizing Sampling and Non-sampling Errors in Telephone Surveys: a Case Study using the “Sample Survey on Births”**

**C. Castagnaro, A. Guarneri, S. Prati and F. Rinesi**

## **Introduction and aim of the work**

It is well-known that Italy is one of the Countries with the lowest fertility level in Europe. Understanding the causes of low fertility becomes more and more of a concern, therefore surveys devoted to this topic are particularly relevant.

Starting from October 2011, a new edition of the Istat “Sample Survey on births” has been going on. It is a computer assisted telephone interviewing (C.A.T.I.) survey, where 15.000 mothers are interviewed 16-30 months after delivery. Individual questionnaire includes information on partnership status, reproductive career, delivery, fertility intentions, both partners’ working conditions, socio-economic status and housing condition.

The population of interest is all live-births registered in the Italian Population Register. A record linkage between this list and the listing of domestic telephone numbers (ex-directory numbers included) has been performed. The matching keys considered are name and surname of the newborn parent’s and of the head of the household, and household address at delivery time. The sample is drawn by the linkage output.

The advantages of C.A.T.I. survey versus different data collection techniques are well known, as well as their limits. In this paper we focus on specific strategies that can be adopted in order to minimize both the sampling and the non-sampling errors when telephone survey method is adopted.

## **Strategies adopted for minimizing the sampling error**

The continually increasing percentage of households with no domestic telephone number raises concern among researchers. The crux of the matter is that the main socio-demographic characteristics of households with no number can vary significantly from the one of households with telephone number. Generally speaking, deprived persons and households, as well as those with high spatial mobility, are less likely to have fixed phone and therefore they may be under-represented in telephone surveys. It follows that disregarding this aspect can lead to inaccurate results.

Our first analyses aim to assess how much this issue can affect the accuracy of the estimates derived from the ongoing “Sample Survey on Births”. To hit this goal we exploited the latest edition of “Aspetti della Vita Quotidiana” (AVQ), an Istat yearly face-to-face survey, that measures, among other aspects, the main socio-demographic information of households and its

members as well as the presence of fixed phone (target variable). This provides a unique opportunity to compare the characteristics of household with/without phone. As a matter of fact, by using Decision Trees procedures it was possible to establish which variables are significantly associated with the presence of phone. We consider a set of 19 independent variables and the analysis was restricted to households with at least one child under the age of 4. Interestingly, only 5 variables (higher educational level between partners, number of household components, home-ownership, geographical area of residence and geographical distance from other relatives) are significantly associated with the presence of fixed phone. Those results deserve high consideration in sample weights' calculation to produce not-biased estimates across all groups, under-represented populations included.

### **Strategies adopted for minimizing the non-sampling error**

An other crucial aspect that must be considered when dealing with not self-assisted questionnaire is the “interviewer effect” that impacts both the response rate and the data quality. Several tools are used to monitor and correct the interviewers' behaviour. Particularly, the continuous monitoring process consists of two lines of action. By analysing the reports on the daily interviewers' performances, it is possible both to compute aggregate standardized indicators and to conduct in-depth analysis on interviewers' behaviour at individual level.

Concerning the first line of action, a set of indicators (Response rates, Cooperation rates, Refusal rates and Contact rates) are daily computed and updated. Those rates, calculated using the standardized definitions developed by the American Association for Public Opinion Research (AAPOR), make it possible to monitor constantly the outcome rates and to perform corrective actions (if needed). The second line of action taken to minimize the interviewer effect consists in analyzing the interviewers' behaviour at individual level. Several aspects are considered: firstly, we regularly compute the indicators developed by AAPOR at micro level; secondly, we daily supervise the interviewers' performance over a set of questions that regards difficult aspects to address and over filter questions that direct interviewers to skip big set of questions that do not apply to the specific interviewee. The interviewers' behaviour over specific questions is assessed by multilevel regression analysis. If we recognize that interviewee are nested within interviewers, an intercept-only model can be estimated: within this context the intra-class correlation coefficient can be interpreted as the interviewer effect. Interviewers that show a random effect that significantly differs from zero are rapidly detected and further trained in order to raise data quality.

### **Conclusions**

C.A.T.I. surveys represent a cost and time saving way to collect socio-demographic data. At the same it is becoming increasingly difficult to ignore

the fact that less and less households have a fixed phone number and this can lead to an under-representation of selected population subgroups. The study of the main characteristics of households with no fixed phone becomes therefore fundamental. At the same time, also the interviewer effect can lead to inaccurate estimates. A continuous monitoring of data collection process and the use of proper methodological tools contribute to data quality enhancement.