The definition of the final disposition codes in the household surveys in the context of the new integrated management system in Istat

Keywords: Final disposition codes, Household surveys, Integrated management systems for the surveys.

1. INTRODUCTION

The Istat Directorate for Data Collection, from its recent constitution (April 2016) to now, is focusing on some strategic projects, aimed at harmonizing and making more efficient the data collection operations for all the surveys conducted by the Institute. Among these projects, there is the integrated management system for the surveys (SGI), already being implemented for the 2018 Permanent Population and Housing Census. The IT development of the system requires a reflection, in particular from a theoretical point of view, in order to identify and implement all the functionalities necessary for the management of each survey, within a single conceptual framework. A fundamental aspect of the design of this integrated management system is certainly the definition of the status and of the final disposition codes for the survey units. Accompanied by an appropriate set of rules of assignment, they are the basis of some important functions connected to the execution of the activities of implementation, monitoring and control of the survey. To overcome the difficulties related to the peculiarities of the single surveys, characterized by different sampling designs and survey modes, we proceeded to make an in-depth comparison between the systems of the outcomes currently used, which led to: reduce redundancies; respect the peculiarities without loss of information; harmonize the final disposition codes of the survey units at a higher level of synthesis. The status and the outcomes of the survey units thus allow to monitor the phases of their initial assignment, of implementation of the field work and of final validation. Furthermore, the final outcomes (after validation) allow the calculation of indicators on survey quality (mainly coverage and non-response rates), in compliance with national (SIDI - Information System for Survey Documentation [1]) and international (AAPOR - American Association for Public Opinion Research [2]) standards. The paper will describe the conceptual structure on the basis of the states and the final disposition codes of the survey units, in the context of some functionalities of the integrated management system (interviewer diary, monitoring).

2. THE NEW INTEGRATED MANAGEMENT SYSTEM FOR THE SURVEYS AND THE SYSTEM OF SURVEY OUTCOMES

The workflows, regulated by SGI, are determined by the system of the survey disposition codes (temporary and final) and of the different status (status of assignment of units, outcome of contacts and of contact attempts on the unit, in progress status of the questionnaire): the different combinations of states and outcomes determine the visibility of the functions in each phase of the fieldwork process and for each user profile. To understand how the outcomes system permeates the whole functioning of the integrated management system, it is important to keep in mind that the outcomes are populated and perfected during different phases of the process, from the phase of unit assignment, to the fieldwork phase, up to the validation phase. The logical sequence is as follows. Once the theoretical sample has been uploaded into the integrated management system, the units must be assigned so that the fieldwork can begin. All the units of the theoretical sample

that have been assigned to a final operator (directly from ISTAT or intermediate body) or to a "fictitious" operator in the case of CAWI surveys, for example, must have an outcome, which comes from the fieldwork. In particular, this activity is achieved with a series of contact attempts from or to the unit, each of which will have a specific disposition code (outcome of the attempt, sometimes called "temporary disposition code"); the combination of the attempts outcomes, then, according to specific rules, can give rise to a final disposition code for the unit. The final disposition codes coming from the fieldwork, afterwards, must pass a validation phase, during which it is plausible that a response (complete interview) turn into non-response due to the fact that, for example, subsequent checks have resulted in fake interviews, or paper questionnaires received did not meet the acceptance standards. At the same time, non-responses could turn into responses, as, for example, incomplete questionnaires have a level of compilation deemed acceptable to be considered valid, or a complete paper questionnaire, that has not been registered as such during the fieldwork for reasons of malfunction of some components of the system or for forgetfulness, has been received. The outcomes, therefore, within the integrated management system, are implicitly and explicitly called into question in various modules and functions. For example, they are among the variables that are displayed in the survey Diary and which guide the actions to be taken towards units; they are the basis for the calculation of the main aggregates and indicators that appear on the summary reports for monitoring; moreover, they represent the calculation basis for payments to be provided to municipalities and private companies entrusted with carrying out the interviews. Therefore, the outcomes are among the fundamental defining elements that are acquired by the system from the first moment of the configuration of a survey, together with all the other elements that constitute the initial information set on the survey (survey mode, list of theoretical units, ...).

3. CONCEPTUAL STRUCTURE OF OUTCOMES AND STATUS: THE CENTRALITY OF ANALYTICAL OUTCOMES, BETWEEN SYNTHETIC OUTCOMES AND SIDI AGGREGATES

In the design of the new integrated management system, with the ambition of generalization of the management functions, the complexity to be faced was immediately evident, with a very high number of surveys to be managed. Each survey, in fact, has its own system of outcomes and related tools suitable for tracking them. The categories of outcomes vary in number and meaning and depend on the technical, methodological and organizational characteristics specific for each survey, such as the survey mode, the survey design and the sampling unit, the eligibility criterion, etc. For example, some types of outcomes, such as non-responses due to worked-out call-counting meter, have existed since computer assisted techniques were introduced, in which the data collection systems have become able to 'count' the attempts made on the survey unit; other types of survey outcomes exist only for some surveys, such as those outcomes that identify the non-responses that occur in particular phases of the fieldwork (for example, in the Household Budget Survey, non-responses due to break-off may occur both during the initial and final interviews and between the two phases of the survey, so in the classification of outcomes there is not a single code for break-off, but as many codes as there are all these cases); finally, some outcomes identify the non respondent units due to the non-possession of a specific eligibility requirement, such as age (for example, in the Adult Education survey, a specific survey outcome identifies the sample units, who can not participate as older than the established age threshold). The task of systematization of the outcomes has required (and is requiring) the sharing of languages and methodologies. The obligatory step was to operate the transcoding between the final disposition codes, starting from those used for surveys on households and individuals and, by analogy of survey units, for population census. The choice to start from these types of surveys is

mainly due to the consideration that it was necessary to start from the most complex surveys, so as to be able to later transfer most of the deductions even on surveys with no interviewers network and more streamlined questionnaires. Moreover, the imminent start of the Census has also dictated the priorities of work on this side. The challenge was to find agreement on the most appropriate conceptual categories to describe all the specific outcome systems. Examining, therefore, the analytical survey outcomes and looking for the correspondences between the labels-codes of these, the impossibility to eliminate the specificities present in the classifications adopted so far has clearly emerged, if not to the detriment of the ability to intercept all the critical issues that may occur during data collection. Furthermore, the survey outcomes of the units in the fieldwork phase must be "telling" for the interviewers, therefore these can only be the analytical outcomes specific for each survey. Due to their high specificity, a complete standardization of the analytical final disposition codes can not be reached, but their harmonization is feasible only at the aggregate level. At the same time, since the need for the "operational" harmonization at least at the macro level is strong, it has been possible to arrive at harmonized synthetic survey outcomes, useful for calculating summary indicators of the fieldwork and also for making a possible comparison between different surveys. It is good to underline that the frontier to be faced is represented by the search for solutions suitable to the management of data collection in an increasing transversal perspective, jointly monitoring the quality of surveys carried out with the same technique or in the same time period or through the same interviewer network and so on. In light of the foregoing, within the integrated management system, in the configuration of each survey the analytical final disposition codes must be uploaded with the respective linking table, which connects each analytical outcome to a synthetic outcome (the latter standardized among all the surveys).

In addition to the aforementioned operation of systematization of the analytical survey outcomes into synthetic outcomes categories, at the same time a review of the existing transcoding tables between the analytical final disposition codes of each survey and the aggregates of the SIDI system (Information System for Survey Documentation) was carried out. The conceptual framework of SIDI, and the relative hierarchical classification (tree) of the aggregates, represented a fundamental point of reference. However, SIDI indicators are not suitable for monitoring the fieldwork phase, while they are suitable very well for being calculated at the end of the validation process, as they are mainly aimed at giving an indication of the final quality of the survey and at guaranteeing international comparisons. Having paid attention in parallel to the needs of the survey management and to the generalization and standardization of the outcome categories, has therefore led to an outcome architecture based on the centrality of the analytical survey outcomes, which preserve all the methodological and technical specificities of a survey from a harmonized point of view and that, on the one hand, are synthesized in a classification of synthetic outcomes shared by all the surveys and, on the other, allows the reconstruction of the SIDI aggregates. Survey status and survey outcomes thus allow to monitor the phases of initial assignment of the unit, of fieldwork and of final validation. Furthermore, the final disposition codes (after validation) allow the calculation of indicators on the survey quality (mainly coverage and non-response indicators), in compliance with national (SIDI) and international (AAPOR and Eurostat) standards.

4. CONCLUSIONS

The deep considerations on the survey outcomes system has so far led to the preparation of technical specifications for the new integrated management system, in particular for the implementation of the module on survey Diary, currently in use for the Permanent Population and Housing Census. Furthermore, a specific document on the outcome

system has already been formalized, which illustrates the meaning of the specific and synthetic outcome categories and which also describes the conceptual categories of survey status and their relations with the outcomes. Other technical specifications will be drafted with reference to other modules of the system that interact in some way with the outcomes system described above.

REFERENCES

[1] Brancato, G., Pellegrini, C., Signore, M., G. Simeoni (2004), "Standardising, Evaluating and Documenting Quality: the Implementation of Istat Information System for Survey Documentation – SIDI", Proceedings of the European Conference on Quality and Methodology in Official Statistics, Mainz, Germany.

[2] AAPOR, Standard Definitions. Final Dispositions of Case Codes and Outcome Rates for Surveys, (2016), https://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf