THE ITALIAN CONTINUOUS CENSUSES: MANAGING THE COMPLEXITY AMONG REGISTERS, STATISTICS AND TERRITORIES[1]

Antonella Bernardini, Daniela Bonardo, Alessandra Dentini, Maura Giacummo, Matteo Mazziotta, Alessandra Preti

1. Introduction

The recent process of modernization of the Italian Institute of Statistics (Istat) is based on the increasing use of databases from administrative sources in order to support and integrate sample surveys; this reduces the statistical burden for respondents (citizens and economic units) and the costs borne by the community, guaranteeing the production of statistical information on an annual, biennial or three-year basis, depending on the census considered, and no longer as ten years as in the traditional census model.

The transition from the ten-year census to the continuous one represents an element of great innovation, with a huge reduction in the costs of the operation and an increase in the frequency of production and dissemination of data. The transition to a new census model is made possible by the acquisition, processing and use for statistical purposes of administrative sources that generate, through data validation processes, statistical registers updated with high time frequency.

The challenge is to maintain a high quality of the data by increasing the frequency of publications. This objective can be reached by breaking the classical scheme based on the exhaustive survey, the counting of population units, the estimation of the non-sampling error.

The use of administrative sources and sample surveys to support the correction of the sources themselves represents a real revolution in official statistics. The actors on stage increase and the functions differ, creating a new work paradigm that needs a different organizational structure than in the past. The simultaneous conduction of some censuses needs to share competences, roles and functions in order to create economies of scale that limit the waste of time and resources.

[1] Acknowledgments: The paper is the result of the common work of the authors. In particular, Antonella Bernardini has written sections 2.1 and 4.3, Daniela Bonardo has written sections 1 and 4.2, Alessandra Dentini has written sections 3 and 4.1, Maura Giacummo has written sections 2 and 4.4, Matteo Mazziotta has written section 2.3 and 4.5 and Alessandra Preti has written sections 2.2 and 4.

The paper aims to present some characteristics of the organizational management of the five Italian censuses (Population, Business, Agriculture, Public Institutions and Non-profit Institutions) in which a new organizational structure composed by the authors of this paper and called "Contact Group" aims to standardize common processes of the censuses allowing the connection between different functions and a large internal diffusion of strategic information.

2. New Strategies and Old Issues

The old censuses were characterized by an organizational scheme based on many actors that interacted with Istat: local authorities, trade associations, interviewers, universities, etc.

The survey covered all the units of the reference population, occurring at the same time and every 10 years. The statistical errors of a census were calculated at the end of the survey and they are called non-sampling errors, since they do not depend on the partial selection of units from the population (Istat, 2013; Istat, 2016).

The new censuses are very different: they do not necessarily cover all the units at the same time, they are also based on sample surveys and statistical registers and a frequency of results publication every year (or two or three, depending on the census) is provided. In fact, with the new continuous Censuses information on the main socio-economic characteristics of the country will always be more detailed and continuously updated. Thanks to the integration of administrative data with those provided by the surveys, costs will be significantly reduced compared to the past and more useful information will be obtained to plan more efficient interventions and services.

2.1. Households

In October 2018 the continuous Census of Population and Housing started. For the first time the surveys are annual and "on sample": they involve 1 million 400 thousand families and about 2,800 municipalities. By 2021, all the municipalities of Italy will have participated at least once in the surveys. The areal sample survey is useful to enumerate the individuals who really insist on the territory and, from methodological point of view, it represents the real value to compare with the individuals of the population register. In the past, in many developed countries,

areal surveys have measured the quality of the population census and this type of survey is called Post Enumeration Survey (PES). In Italy there is an excellent tradition of conducting these investigations: the experience gained can help to design and carry out new areal surveys to constantly improve the quality of the population register.

2.2. Companies

Starting from 2019 also the Business Census becomes continuous: every three years, to increase the quality of the information offer, the information of more than 250 thousand productive units will be collected. In this case, the reference register has a multi-year history; in fact, the statistical register of active companies (ASIA) has been perfected over time and today it represents an essential reality for all surveys on Italian companies. From this point of view, the "road" of the continuous census is simpler than the population one because the detection units are simpler to capture and to collect through administrative sources.

The first continuous Census of Non-Profit Institutions was already implemented in 2016, with the survey of about 43 thousand units. The next edition is scheduled for 2019 and every three years (2022, 2025 and so on) with a sample size of 70 thousands units. There are positive prospects for administrative data in support of the construction of sectorial registers (third sector reform toward a single register). The data collection technique is based on mixed mode in which the analytical information is structured by stable parts and variable modules.

The continuous Census of Public Institutions has a biennial frequency and is the only one, among the new censuses projected by Istat, not to be a sample; in fact it involves all the institutions present in the country (about 13 thousands institutions and related local units). The second survey of the new course ended at the end of 2018 and the next will be in 2020. The collection of analytical information on structure, operation, organizational, innovation, human capital, relations with the market is based on stable parts and variable modules. The integrated publication of the results with information from basic and extended registers will be progressively implemented. In fact, the extended register of Public Institutions (Frame PI) is in progress. The data collection technique is CAWI and then the cost is reduced.

The Agricultural Census will change starting in 2021: the last general census of agriculture with ten-year frequency, in which all Italian farms will participate, is in fact planned for 2020. Subsequently, the farm register will be corrected with sample surveys every three years (2023, 2026 and so on). The Integrated publication of the results with information from basic and extended registers will be progressively implemented on an annual basis. The size of 2020 survey is about

1.5 millions of farms. In the future, the burden and the cost reduction on respondents and the integrated use of multiple collection techniques will be the main characteristics of the new world of agricultural surveys.

2.3. Errors

In this context, the statistical error will be composed of a sample and a non-sample part and that the sources of the error are generated, in addition to the actors involved, also by the use of new administrative data. For this reason, it seems essential to focus attention on quality checks, some of which are included in the surveys themselves. Such a great revolution must be supported by a new management of skills and roles within the complex process of producing the final data, providing active management structures to solve problems and to facilitate dialogue between the actors involved.

3. Management of "Functional Processes"

The Contact Group, established in the spring of 2017, is made up of seven Istat researchers and technologists who present different skills in order to cover different scientific and organizational functions. In fact, skills range from management to information technology, from statistics to accounting; furthermore, the specificities of each member can adapt to the specificities of such diverse censuses. The main function is to provide a management approach to the most complex phases by finding a sustainable solution that can be used for all censuses.

Although the specificities of each census are very different, the goal of the Contact Group is to find common points in order to converge the pillars defining the general strategy. In particular, some urgent recommendations are proposed:

- outlining in a single Gantt the planning processes of the various censuses;
- developing shared guidelines for the General Census Plan (PGC) that take into account the continuity of the new censuses;
- synchronizing some tender's procedures, in particular those concerning IT equipment or in any case provide for the optimization of the invested resources;
- encouraging the circulation of methodological notes and information materials on the processes that generate statistical registers;
- reasoning on the transversal and on the organizational optimization of the Advisory Committees;

- evaluating the resources to be committed on the various processes, in particular for the study of procedures for exploiting Big Data and data for the geolocation of the units;
- accurately defining the development of communication plans;
- deepening and coordinating the methods of data publication and more generally planning the IT management of the censuses.

For almost two years, the Contact Group has assisted the Istat structures to carry out their tasks, has organized technical meetings and organized seminars on frontier issues: the group has dealt with the general census budget, produced technical reports for specific commissions and provides methodological support in order to calculate the statistical error of the censuses. Today it has become a reference point for the exchange of relevant information towards a new era for official statistics.

4. Transversal Functions

We are aware of the different states of progress about the design of the five censuses, as well as of the different conditions in which we can realize them (specific sources, alliances and norms, experiences and internal resources etc.).

With different speed we must however establish a point in time in which to converge on the foundations that define the strategy of continuous censuses. The year on which to focus and succeed in aligning the founding elements of the strategy is 2021.

As mentioned in the previous sections, the decisive role of this great revolution in official statistics is the construction of statistical registers and the relations among themselves, samples and census surveys. The basic register is the result of a vast whole, to be continuously fed even with new acquisitions of administrative sources, together with periodic coverage surveys in order to evaluate the quality (no-sampling error). Since the target is to define the uniqueness of the size of the populations of interest in order to support the statistical Institute's production and the analysis of the processes. In fact, the identification of the fundamental administrative source (s), whose availability must be stable, foreseen by law, and whose correction on the basis of the basic Register may be of institutional interest (for example, municipal register, Chambers of Commerce, Land Register and so on).

In this context, some transversal functions are common to the five censuses and the work of the Contact Group is useful for sharing the different experiences in order to optimize phases and processes of the survey. The main transversal functions are listed below.

4.1. Administrative aspects

The Contact Group assisted the Istat organizational structures in order to achieve the planning objectives of some important administrative aspects concerning the censuses.

The first fundamental task was a survey of the status of administrative procedures such as the budget, the tenders for the awarding of services, the General Census Plans (PGC's) and other legal aspects concerning the obligation to reply to the questionnaire and the penalties for non-respondents. In the month of May 2017 the Contact Group has begun the recognition of the financial needs for the realization of the permanent censuses connecting with the different interested structures. The needs highlighted by the production structures and the transversal structures, also taking into account the scheduling of the tenders, were structured by macro item of expenditure and by single census. The expenditure forecasts have been analyzed for each year, starting from 2018, and reports have been prepared highlighting the expenses articulated for two years, distinguishing the period 2018/2019 and 2020/2021 from the following period (2022 - 2031). In order to allow the aggregation of the items and the comparison between the censuses, the subdivision of the expenses by macro-code has been prepared homogeneously for all the censuses. The expenses were divided into current expenses and capital expenditures, and further divided into macro-items of expenditure. The financial tables produced were the basis for writing the census law which was subsequently approved by the Italian Parliament.

Another task, developed in accordance between Contact Group and Istat thematic sectors, is the writing of the General Census Plans. The PGC is a planning act of a general nature that dictates the guidelines regarding the planning, organization and execution of operations carried out at each census. The General Census Plan, drawn up in accordance with the European regulation scheme, describes the legal foundations of census, the organization of the survey network and the tasks of the census bodies, the main technical and methodological aspects of the survey and the calendar of the operations. It represents a point of reference both for Istat and for all the organizations and actors involved in census operations. Moreover, in view of the permanent censuses, the PGCs assume even more

importance because they regulate many operations that are repeated over time with an annual frequency (or biennial or three-year) and not a decade.

4.2. Training

The training of the interviewers and of the various actors on the field is one of the main aspects for the success of such complex surveys as the censuses. The general principle underlying the design of a training process is to allow the detection networks to acquire the skills and abilities necessary to detect high quality data. In fact, training is one of the main strategies for preventing non-sampling errors. In detail, it is important to:

- provide all the tools necessary for the activity to be carried out;
- provide well-calibrated tools for recipients;
- ensure uniform training in the areas;
- provide adequate tools for continuous learning also through the support of the Moodle platform;
- maintain contact with the network for sharing critical issues and solutions;
- provide self-assessment tools and learning assessment;
- provide tools for process monitoring;
- provide tools for the evaluation of the training project.

4.3. Data Collection

The data collection phase is certainly among the most delicate because a bad management of the process can cause errors that cannot be corrected. In this regard, briefly, the main aspects to consider are the following:

- take into consideration the quality of the survey lists;
- adopt operational solutions for the implementation of a possible mix-mode;
- define the integrated tools for monitoring data collection;
- guarantee assistance to respondents and to the survey network.

In this context, it seems necessary to: anticipate any critical issues related to the conduct of data collection in the field, mainly related to any mix-mode; establish virtuous interrelationships between Istat actors and the survey network;

design of integrated and efficient management tools; give greater impetus to the use of administrative sources in a continuous census.

4.4. Survey Management System

The function of data collection is centralized in a single direction, giving the possibility to rationalize and integrate the data acquisition and monitoring phases of the operations. From a technological point of view (IT mainly) this has led to a standardization of the processing phases allowing the possibility to design generalized platforms consistent with well-defined processes. In particular, for the data collection and management phase, two web applications have been developed called the Survey Management System (SGI) and PANDA, as acquisition system, which allowed the first continuous census of population and housing to be carried out and they will also be reused for the census of agriculture and census of non-profit. The acquisition system allows the management of the multi-technique, it guarantees the management of the guided flows of the questions, it manages the consistency and the formal correctness rules. The Survey Management System follows all the collection phases in the field: establishment of the census offices, creation of its own survey network, assignment of work to surveyors, management of the field survey and monitoring of the system. From a technical point of view the new platforms are:

- designed according to the user centricity approach in order to facilitate their usability towards users;
- homogeneous from the point of view of the software infrastructure: this choice helps the IT management to manage any corrections and evolutionary maintenance in a more rational and immediate way;
- developed in compliance with the architectural principles of a service approach;
- integrated with the Institute's metadata systems to allow centralized management of all available information.

4.5. Data Quality

All register-based censuses cannot disregard the quality measurement and the correction of the starting list. The phases in which the statistical error can occur are many and it is necessary to apply different statistical methodologies in order to keep both the sampling part and the non-sampling part under control (Bernardini

and Mazziotta, 2013). The incidence of non-sampling errors, particularly in complex investigations such as census in which a considerable effort of memory and knowledge of the phenomenon are required, can seriously affect the reliability of final results. The quality takes on the meaning of precision that is expressed as an inverse function of the error statistics. The aim of Istat is to provide accurate estimates of the main non-sampling errors, particularly in complex investigations such as the Census (Grossi and Mazziotta, 2016). The researchers must take into consideration many phases, such as the construction of the census list (eligibility of the units, reduction of over/undercover errors and measurement of the error on the microdata), provisional data release, check and data correction and so on. Specifically, the measurement of coverage of the statistical register is a fundamental phase of the whole process of modernization of the official statistics in Italy. In order to evaluate the census coverage, an areal sample survey is carried out. In fact, the aim is to estimate the real number of units really existing on the Italian territory and the coverage rate is the ratio between the number of units pointed out in the register and the number of units really existing on the Italian territory at the same time of register construction. Particular estimation models allow for estimating the number of units for each reference territory of the census with very short time intervals (for example, every year for the population census. For more details, see section 2). In this phase, it is possible to correct the reference register and always have an updated situation of the statistical units present in the territory (individuals, farms, companies). Moreover, the register of places will allow to accurately localize the units providing an indispensable tool to the scientific community and to the public decision maker.

5. Conclusions

The era of big data and the use of administrative sources for statistical purposes is probably the biggest revolution in official statistics. Istat has initiated a process that through the construction of statistical registers completely changes the philosophy of data collection. The Contact Group represents a point of union of skills and professionalism available to the different sectors of the institute in order to standardize common processes towards a quality statistical product.

The Contact Group is a consistent choice with the modernization program, which is based on a common commitment to all the censuses in terms of the fundamentals of the production and output process.

In fact, the modernization is a great opportunity for:

- implementing the registers in an organic way with the different institutional processes;
- harmonizing the survey processes and implement unified solutions for problems common to various contexts;
- developing skills, advanced services, integrated products;
- strengthening the relevance of census products through a structured dialogue with stakeholders and other institutions (Ministries, Municipalities and Regions, Privacy Authority).

Moreover, some strengths are: usefulness of the operation; benefits in terms of cognitive content, relevance of the information produced; increase in timeliness, punctuality and accessibility of data; maintenance of data accuracy and comparability standards; high rate of innovation; reuse of the statistical information and consequent reduction of the burden; great cost reduction compared to traditional censuses.

New census translates into new products: supply of new and diversified annual products in new ways. In order to achieve this objective, the Contact Group has made available different skills in order to be able to cover numerous transversal functions in order to be able to co-deploy the Institute's resources in connections with internal and external structures. This task made it possible to standardize many operational functions by facilitating the exchange of information among the many actors involved in census operations. This is just the beginning of a long process of change of official statistics and censuses towards the joint use of direct surveys and statistical registers. This process also needs an adaptation of the organizational structure in order to optimize workflows and allow the different phases to become an efficient mechanism.

References

BERNARDINI A., MAZZIOTTA M. 2013. Atti del 6° Censimento Generale dell'Agricoltura – La valutazione della qualità, Vol. 6, Roma: Istat. https://www.istat.it/it/files/2014/02/5.La-valutazione-della-qualit%C3%A0.Atti-6%C2%B0-Cens-agr..pdf

GROSSI P., MAZZIOTTA M. 2016. Atti del 15° Censimento Generale della Popolazione e delle Abitazioni - La valutazione della qualità, Vol. 6., Roma: Istat. https://www.istat.it/it/files//2016/08/AttiPOP_Fascicolo-6-web.pdf

SUMMARY

The Italian Continuous Censuses: Managing the Complexity among Registers, Statistics and Territories

The tradition of censuses in Italy is very strong and, over the years, the experience has allowed us to arrive at exhaustive measurements that guarantee a high quality of the data. The challenge of the coming years will be to maintain the same level of quality for the same territorial domains by increasing the frequency of publications. This will be possible thanks to the combination of sample surveys and statistical registers; the use of administrative data, a new chain of information construction and a renewed organizational management will be the challenge that official statistics will have to win.

Antonella BERNARDINI, Istat, anbernar@istat.it
Daniela BONARDO, Istat, bonardo@istat.it
Alessandra DENTINI, Istat, aldentin@istat.it
Maura GIACUMMO, Istat, magiacum@istat.it
Matteo MAZZIOTTA, Istat, mazziott@istat.it
Alessandra PRETI, Istat, preti@istat.it