

Centralizing data collection implementation: the Istat experience

Giampaola Bellini, Novella Cecconi, Loredana De Gaetano, Francesca Monetti, Pasquale Papa and Rita Ranaldi (Istat, Italy)

bellini@istat.it; ncecconi@istat.it; degaetan@istat.it; fmonetti@istat.it; papa@istat.it; ranaldi@istat.it

Abstract and Paper

Italian National Statistical Institute (Istat) traditionally organized the statistical processes according to a classical ‘stovepipe’ model that involved independent, non-integrated setup including all the necessary skills: statisticians, information technology experts, thematic experts, methodologists.

This organizational layout hampered standardization, leading to proliferation and duplication of procedures, methodologies and technologies, producing: i) diseconomies of scale which entail redundancies and inefficiencies, ii) reduction of the potential professional development of the human resources as limited to specific skills and solutions.

During 2016 Istat started a corporate restructuring process that interested all the statistical production structures and that led to a completely renewed organizational setup.

The new model restricts the role of production structures only to the thematic aspects, while the “cross” expertise are all assigned to specialized sectors. The most important innovation concerns the creation of the new Central Directorate for Data Collection that is characterised by a very high degree of specialization of activities and Human Resources.

Regarding the activities run by the new Division for field implementation of data collection from direct surveys, included in the mentioned Directorate, the paper will present a set of significant representative experiences of standardization and harmonization activities that were introduced in different data collection field implementation domains in order to improve the efficiency of data collection processes. In particular, the paper deals with the following representative experiences:

- a) Process innovations introduced and main results achieved in direct data collection from structural and short-term business surveys

Process innovations introduced in structural and short-term business surveys are mainly based, on one hand, on the implementation of infrastructural solutions, as the single access point to data acquisition systems - the “Business statistical portal” - and the centralized Contact center for inbound and outbound services; both solutions are managed through a detailed calendar of each data collection activity. On the other hand, process innovations are based on the standardization and generalization of

each phase of the data collection process and on the specialization of personnel devoted to specific transversal activities. After two years, main results were the significant increase in medium response rates and the reduction of the data collection periods. A detailed description of the innovations and standardization adopted and of the consequent results will be provided.

b) The implementation of a territorial model to improve the efficiency of data collection

A specific project assigned to a selected number of ISTAT territorial offices the new role of running data collection activities on the territory. According to this innovative model a specific territorial office acquires the role to coordinate the running of data collection activity of all the other territorial offices.

The redesign of the data collection for maritime transport survey entrusted a coordination role to the Territorial office Campania and Basilicata. It determined the revision of the flow of interventions on the following issues: a) updating the list of respondents; b) support respondents; c) monitoring data collection on the territory, d) checks of coverage and quality on information collected.

The case study concerning the involvement of the Campania and Basilicata Territorial Office in the DC implementation introduces an innovative way of managing data collection, that is characterized by a distributed throughout the territory approach, opposed to the standard model which provides for a centralized approach. Significant increases in the quality of survey lists and response rates were observed as a consequence of the innovations introduced.

c) The definition of the final disposition codes in the household surveys in the context of the new Integrated Survey Management System

The focus will be on the conceptual structure designed in order to 1) reduce redundancies 2) respect the peculiarities without loss of information 3) harmonize the final disposition codes of the survey units at a higher level of synthesis, in the context of the new Integrated Survey Management System.

d) Personalized reminders: the experience in the Italian survey on the vocational integration of research doctors

The focus will be on the use of personalized feedback to non-respondents according to the different phase in which the survey unit was (who had never made any action, who started the compilation of the questionnaire without completing it, according to the section where he interrupted, etc.) in order to improve the response rate of the survey.

Keywords: data collection centralization, response rates, process efficiency, standardization.

UNECE Workshop on Statistical data Collection 'Resourceful Data Acquisition'

10-12 October 2018, Geneva, Switzerland

Centralizing data collection implementation: the Istat experience¹

Giampaola Bellini - Istat (bellini@istat.it)

Novella Cecconi - Istat (ncecconi@istat.it)

Loredana De Gaetano - Istat (degaetan@istat.it)

Francesca Monetti - Istat (fmonetti@istat.it)

Pasquale Papa - Istat (papa@istat.it)

Rita Ranaldi – Istat (ranaldi@istat.it)

1. Introduction

During 2016 the Italian National Statistical Institute (Istat) launched a wide modernization programme whose main objective was to enrich the supply and the quality of the information produced, improving the effectiveness and efficiency of the statistical processes. Istat designed and implemented a new organizational set-up that was characterized by the centralisation of all the support services, that were clearly separated from statistical production.

The introduction of a structure characterized by production sectors on the one hand and service-providing sectors on the other, pushed towards the "transversalization" and specialization of many services that were managed in a specialized manner. The result obtained consisted of a strong standardization and harmonization of all these services and in particular of data collection.

The introduction of a specialist data collection, led to the review of the organizational structure of data collection processes and the redesign of many of the management procedures adopted.

Before the above mentioned reorganization, the statistical processes were organized according to the classical 'stovepipe' model, that involved independent, non-integrated, statistical processes including all the necessary skills: statisticians, information technology experts, thematic experts, methodologists. This choice, although it was characterized by a high probability of achieving the pre-established objectives, in terms of compliance with the Regulations and compliance with the national dissemination schedules, implies a very low overall efficiency level, due to overlapping redundancies and lack of integration between processes.

The main trends behind the modernization Program are the decreasing number of human resources assigned to the National Statistics Institute, the greater degree of training and specialization of available human resources, the development of communication and information technologies, the computerization of the main survey units where data are

¹ Contributors: G. Bellini paragraphs 2.1.2, 2.1.3; N. Cecconi paragraphs 4.1, 4.3; L. De Gaetano paragraphs 3.1, 3.2, 3.3, F. Monetti paragraphs 2.1.1, 2.2; P. Papa paragraphs 1,6; R. Ranaldi paragraphs 4.2, 4.4, 5.

collected (companies, institutions, individuals and families), the need for greater consistency between the statistical indicators produced, in particular at the level of National Accounting indicators.

The new model restricts the role of production structures only to the thematic aspects, while the "cross" expertise are all assigned to specialized sectors. The most important innovation concerns the creation of the new Central Directorate for Data Collection that is characterised by a very high degree of specialization of activities and Human Resources.

In fact it included four Divisions specialised in the following areas: 1) Division for design of data collection tools; 2) Division for data collection organization; 3) Division for implementation of data collection from direct surveys; 4) Division for integration of administrative sources and registers.

Focusing on the "Implementation of Data collection from direct surveys" Division, the goal of specialization was pursued concentrating a series of activities typical of survey's implementation in a single Division. A further internal subdivision concerned the type of responding units involved (businesses, families and individuals, farms, public and private institutions, others). The integration of data collection implementation processes, previously run independently, promoted standardization, with a view to optimizing and increasing efficiency. As a consequence, several process innovations were implemented.

In field of implementation of data collection it is possible to summarize the following areas of harmonization:

- a) Acquisition of data on the field by conducting direct surveys;
- b) Monitoring Data Collection activities;
- c) Archiving of data collected in a special raw data Repository;
- d) Management of non-thematic relationships with respondents (inbound and outbound contact centers, standard responses on non-thematic aspects, recurring thematic faqs);
- e) Relations with intermediate Data Collection bodies;
- f) Designing generalized survey management systems for data collection;
- g) Collaboration with Territorial Offices and National Statistical System.

The new Division aims to carry out these activities in a centralized and integrated manner, overcoming the previous organizational approach that envisaged single data collection processes, substantially independent in terms of procedures and resources allocated. The new centralized management model for data collection introduced at ISTAT following modernization pushes the possibilities of standardizing processes to the maximum, placing the single limitation on compliance with the sectoral specificities of the processes linked to the types of units involved (companies, families, individuals, institutions, farms) and the collection technique used (e.g. use of a territorial network, intermediate bodies, etc.). The first positive results, obtained following the introduction of centralized data collection, originate mainly in the standardization of data collection processes carried out in the various research contexts.

An example of a reference for process standardization activities is the design and implementation of the Statistical Business Portal (paragraph 2.1.1), which has made various solutions oriented to the integrated management of data collection processes.

In relation to the more traditional communication flows originating from the statistical fulfilments to which the company periodically performs, the Portal offered new services with regard to the possibility, by companies, to communicate any changes in personal data and status of activity, to delegate filling in the questionnaires to personnel / offices inside the company or external consultants, to access the scheduled survey calendar and the state of statistical compliance updated in real time.

Regarding the activities run by the new Division for field implementation of data collection from direct surveys, the paper will present a set of significant representative experiences of standardization and harmonization activities that were introduced in different data collection field implementation domains in order to improve the efficiency of data collection processes. In particular, the paper deals with the following representative experiences:

a) Process innovations introduced and main results achieved in direct data collection from structural and short-term business surveys

Process innovations introduced in structural and short-term business surveys are mainly based, on one hand, on the implementation of infrastructural solutions, as the single access point to data acquisition systems - the "Business statistical portal" - and the centralized Contact center for inbound and outbound services; both solutions are managed through a detailed calendar of each data collection activity. On the other hand, process innovations are based on the standardization and generalization of each phase of the data collection process and on the specialization of personnel devoted to specific transversal activities. After two years, main results were the significant increase in medium response rates and the reduction of the data collection periods. A detailed description of the innovations and standardization adopted and of the consequent results will be provided.

b) The implementation of a territorial model to improve the efficiency of data collection

A specific project assigned to a selected number of ISTAT territorial offices the new role of running data collection activities on the territory. According to this innovative model a specific territorial office acquires the role to coordinate the running of data collection activity of all the other territorial offices. The redesign of the data collection for Maritime transport survey entrusted a coordination role to the Territorial office "Campania and Basilicata". It determined the revision of the flow of interventions on the following issues: a) updating the list of respondents; b) support respondents; c) monitoring data collection on the territory, d) checks of coverage and quality on information collected. The case study concerning the involvement of the "Campania and Basilicata" Territorial Office in the DC implementation introduces an innovative way of managing data collection, that is characterized by a distributed throughout the territory approach, opposed to the standard model which provides for a centralized approach. Significant increases in the quality of survey lists and response rates were observed as a consequence of the innovations introduced.

c) The definition of the final disposition codes in the household surveys in the context of the new Integrated Survey Management System (SGI)

The focus will be on the conceptual structure designed in order to 1) reduce redundancies 2) respect the peculiarities without loss of information 3) harmonize the final disposition codes of the survey units at a higher level of synthesis, in the context of the new Integrated Survey Management System.

d) Personalized reminders: the experience in the Italian survey on the vocational integration of research doctors

The focus will be on the use of personalized feedback to non-respondents according to the different phase in which the survey unit was (who had never made any action, who started the compilation of the questionnaire without completing it, according to the section where he interrupted, etc.) in order to improve the response rate of the survey.

2. Process innovations introduced and main results achieved in direct data collection from structural and short-term business surveys

Process innovations introduced in structural and short-term business surveys are mainly based, on one hand, on the implementation of infrastructural solutions, as the single access point to data acquisition systems - the "Business statistical portal" - and the centralized Contact center for inbound and outbound services; both solutions are managed through a detailed calendar of each data collection activity. On the other hand, process innovations are based on the standardization and generalization of each phase of the data collection process and on the specialization of personnel devoted to specific transversal activities. After two years, main results were the significant increase in medium response rates and the reduction of the data collection periods. A detailed description of the innovations and standardization adopted and of the consequent results will be provided.

2.1 Innovative tools and services supporting DC activities

2.1.1 . The business statistical Portal

The implementation of the Business statistical Portal in the context of the economic surveys involved a new approach in their management that turned from "survey-centered" to "enterprise-centered". While in the past in the management of data collection the prevailing approach placed the needs of statistical surveys in the foreground with the introduction of the Portal the needs of the companies involved in the investigations pass in the foreground. Additionally the Portal introduced for the first time an integrated management of some phases of the DC process.

Main objectives of the Portal can be shown as follows a) Streamline the operations required by respondents to fulfill their response obligations, with an overall reduction of the burden b) Increase both ordinary and extraordinary (e.g. news) communications on the survey events and activities c) standardize and harmonize data collection in order to increase efficiency at the system level.

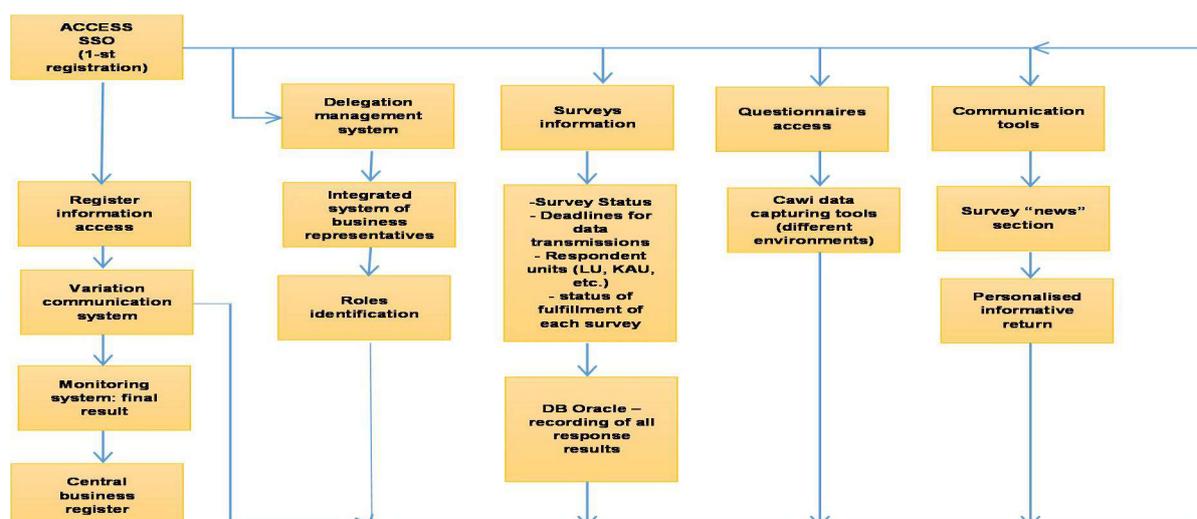
Referring to functionalities available, the Portal is a tool for optimizing the bi-directional communication between Istat (Italian National Statistical Institute) on one side and the companies involved in the statistical surveys on the other side. The section of the Portal dedicated to this function is named Front-Office.

The main features of the section concern the following aspects:

- *Single sign-on and single point of access*, as to say the possibility for the company to access all the questionnaires to be completed and to all the functions of the Portal with unique credentials.
- *Integrated Register changes management*, changes in status and structure communicated by companies are immediately validated by the Register managers and shared with all the structures that carry out the surveys and those that produce samples and survey lists. The Portal also contains a monitoring system for variations processing status and the assignment of final outcomes.
- *The delegation system*, that offers the possibility to manage in a flexible way the assignment of the questionnaires to the appropriate contacts within each company, to involve several offices in completing the same questionnaire, to delegate outside the compilation (e.g accountant), to centralize the requirements at a single office or at a holding company of a group of companies. The contacts can take different types of users profile (administrator, delegate, referent), characterized by different privileges of access to the questionnaires and functionalities.
- *News management*, a specific section dedicated to news regarding the conduction of the surveys (start, closure, reminders, extraordinary events, extensions, technical problems).
- *The state of obligations*, complete and updated framework of all the statistical activities required to each company and of the state of fulfillment of each survey, presentation of additional information of interest to the user (survey status, date of closure of the RD, presence of responding units other than enterprise, etc.)
- *The personalized statistical information return*, a specific section dedicated to the return of personalized statistical information to respondents (main economic indicators and foreign trade statistics). The statistics of foreign trade are used to carry out a “benchmarking” of companies in the markets in which they operate, relating them to the main competitors.

The following block chart points out the main functions and the information flows concerning the front side of the Portal. In the chart some user navigation paths are also shown.

Figure 1. Main functionalities of the Business Statistical Portal



A second section of the Portal (back-office) includes a set of functions to support the management of the survey, with particular regard to help-desk activities, survey administration, monitoring of the data collection process, management of outcomes, reports of variation in register data, data validation, monitoring of linked users, management of the DB of internal and external contacts.

2.1.2. Centralised inbound and outbound Contact center services

The new organization of Division for data collection implementation from direct surveys also implies more specialization of managing the contacts with respondents. In particular, the outsourcing of the activity is entrusted to a specialized company in Contact Center (CC) services. The aim is pursuing progressive centralization of the support and assistance services addressed to the units involved in the surveys (inbound) and of telephone alert and reminders addressed to non-respondent units (outbound). The unique and coordinated management of the service guarantees strong standardization not only within each specific thematic sector but also among sectors, due to the increased transfer of the best practices from one sector to the other.

Settling the service, main effort was devoted to define:

- the information flow among the different actors involved (enterprises, CC operators, DCI personnel);
- the functionality of the sharing tools for the management of the inbound service, the trouble ticketing system, with the implementation of a specific software called “*shared agenda*”, tool used to assign incoming tickets to the appropriate thematic or non-thematic expert;
- the default contact procedure for the outbound service;
- the characteristics of the materials (mainly thematic FAQ) to be produced by thematic sector;
- the FAQ to be provided to the CC operators aimed at ensuring the uniformity of the unit treatment by using a set of harmonised answers in both services;
- the content of the training to be run each time a survey has to be launched;
- the strategies for calendars implementation of the activities to be run.

Main objective of both services is to facilitate the enterprises approach to the questionnaire, reducing the respondent burden, and to remind statistical duties in order to maintain or increase the overall response rate.

In more details, the inbound service provide assistance and support to responding units in the access and navigation of the Business statistical Portal, as well as on the general rules that define the statistical activity and expose the legal obligations for respondents. Finally, it provides answers to the most recurring questions about major instances concerning the survey's content. The assistance is guaranteed by synchronous (free number) and asynchronous channels (dedicated email address). For requests that are not solvable by FAQ, the CC use the "**shared agenda**", a tool that, as already mentioned, presents features useful for managing and sharing the received instances.

The outbound service is realized contacting by telephone the referents stored in the Business statistical Portal and indicated by the responsible of the production survey unit. This service also provides assistance on access to data capturing systems. In the case of the business structural surveys, the contact is carried out in a fixed period of time before the closing of the survey and it is limited only to the most relevant non-respondent units, while for the short-term surveys it is carried out few days after the punctual deadline of the monthly/quarterly Data Collection and during the 'useful' period. The use of a customized contact procedure that is adapted to the specificities of the survey guarantees the uniformity of treatment of the units contacted.

From January 2016 to June 2018, the inbound assistance developed about 250.000 total service requests, of which 90 percent were solved in first and second-level assistance, i.e. for easy resolution requests, and the remaining for assistance on more complex requests, the third level, that involves the intervention by ISTAT personnel. Between the two possible contact channels, the telephone is by far the one preferred by the users, representing about 75 percent of the activated contacts, the remaining quota being made by email.

For the outbound service, from July 2017 till May 2018, 43,580 contacts that led to the recovery of the questionnaires were carried out, 59 percent of them was used for structural surveys (on enterprises or institutions) recall and the remaining for short-term surveys. On average the retrieval of respondents was around 31 percent, being lower for surveys on small enterprises (e.g. SME-SBS) and higher for the ones on larger companies (e.g. SBS).

2.1.3. Innovations introduced in the field of DC implementation

Data collection centralisation required a complete revision of procedures adopted in order to standardise and generalise all the activities run. In the following the main innovative approaches are listed:

- *Harmonized survey lists management.* The preparation of the survey lists was standardized and generalized, by means of a new procedure involving two successive steps of treatment: i) verification of the eligibility of the units included in the survey samples, in order to define the correct and updated lists of units to involve in the survey. These units receive the informative letter, signed by Istat's President, that communicates the start of the survey. Eligibility is assessed taking into account possible recent business transformation events, insolvency proceedings, registrar

modifications and economic activity variations; ii) normalization of the mailing list, verifying for each unit the completeness of register information useful for the correct delivery and integrating possible gaps.

- *Standardization of the contact modalities.* The following standards were adopted: i) single centralized access point both for the data capturing systems (by means of Business Statistical Portal) and for the incoming contacts (free of charge inbound Contact Center) by telephone or by email; ii) system of harmonized standard answers to be used in order to provide efficient and timely resolutions to requests coming from units on non-thematic and recurring thematic issues. The requests are addressed to centralized inbound Contact Center service or directly to Istat DC offices.
- *Strict scheduling for formal and informal communications.* The data collection implementation requires the definition of a strict timetable for the management of the formal and informal communications addressed to the units involved in the surveys. The following timetable has been adopted, following different approaches for structural and short-term surveys.

Table 1. Timetable of formal and informal communications adopted by structural and short-term surveys

TYPE OF SURVEY	DC START		FIELD DATA COLLECTION							
			Survey reminders pre-deadline				Deadline for data submission	Survey reminders post-deadline		
STRUCTURAL	Sending informative letter	Sending ordinary email to survey contact persons	First reminder by certified email (halfway survey period)	Second reminder by certified and ordinary email (around a month before deadline)	Extra reminder by certified and/or ordinary email (for surveys with low <i>rr</i> or short data collection period)	Telephone recall (about twenty-one to seven days before deadline)	Informative letter deadline (data capturing system closure)	-	-	-
SHORT-TERM			Reminder by ordinary email (about two days before informative letter deadline)					Date of informative letter deadline	Reminder by certified email (about two to ten days after deadline)	Telephone recall (about five to ten days after deadline)

The massive submissions are carried out through a specific software application (named Archiflow) that allows the creation and sending by certified email, scheduling the starting time, of dynamic text messages; massive ordinary email dynamic text messages are managed by means of a proprietary Web application named MMM (Mail Massive Manager).

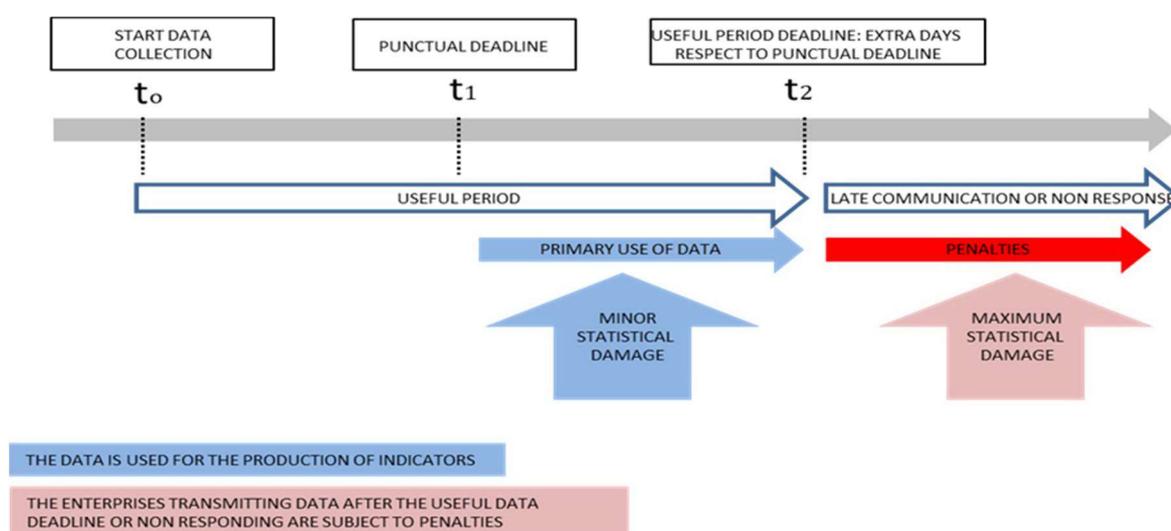
- *Procedures and tools for monitoring the data collection process.* Automatic and generalized procedures were implemented in order to monitor the entire data collection process. The aim is implementing timely corrective actions to control non-respondents, such as extra reminders in addition to those already scheduled (by ordinary or certified email, or phone). These procedures allow cyclical (comparing to the previous period of the same year) and structural comparisons (same period of the previous year) on the basis of specific indicators (e.g. response rates). On this regard, particular relevance assume the management tools provided by the back office of the Business Statistical

Portal that allows a detailed analysis of the response rates for territorial level, economic activity and for specific employees classes.

- *Data capturing and security systems.* Data capturing takes place in a safe mode through generalized Web systems which allow the storage of raw data in a separate and centralized logical environment that allows monitoring of all deliveries to different recipients.
- *Harmonised penalties management procedure.* The integrated approach to the CDC management allowed the generalization of the procedures used for the generation of the lists of the units subject to penalties. The lists are produced at the end of the DC period, after appropriate check of the most recent register information.

In particular, in the context of the short-term surveys an important innovation was introduced, aimed at redefining the procedure for the identification of units subject to penalties. The new procedure was implemented with the aim to produce timely and quality statistical information while trying to minimise the statistical damage charged on Istat; the statistical damage has been assessed on the basis of the response behavior by the units involved in the surveys in relation to the phases and timings, as reported in the following Figure 2.

Figure 2. Penalties management criteria in short-term surveys



2.2. The response rates of economic surveys pre and post the start of the modernization program

During the last two years Istat experienced a clearly increasing trend in response rates (rr) both in structural and short-term economic surveys. The increase of the rr was normally associated to a significant reduction of the data collection period. Particularly for main structural² economic surveys, the results show that generally speaking the rr increased of

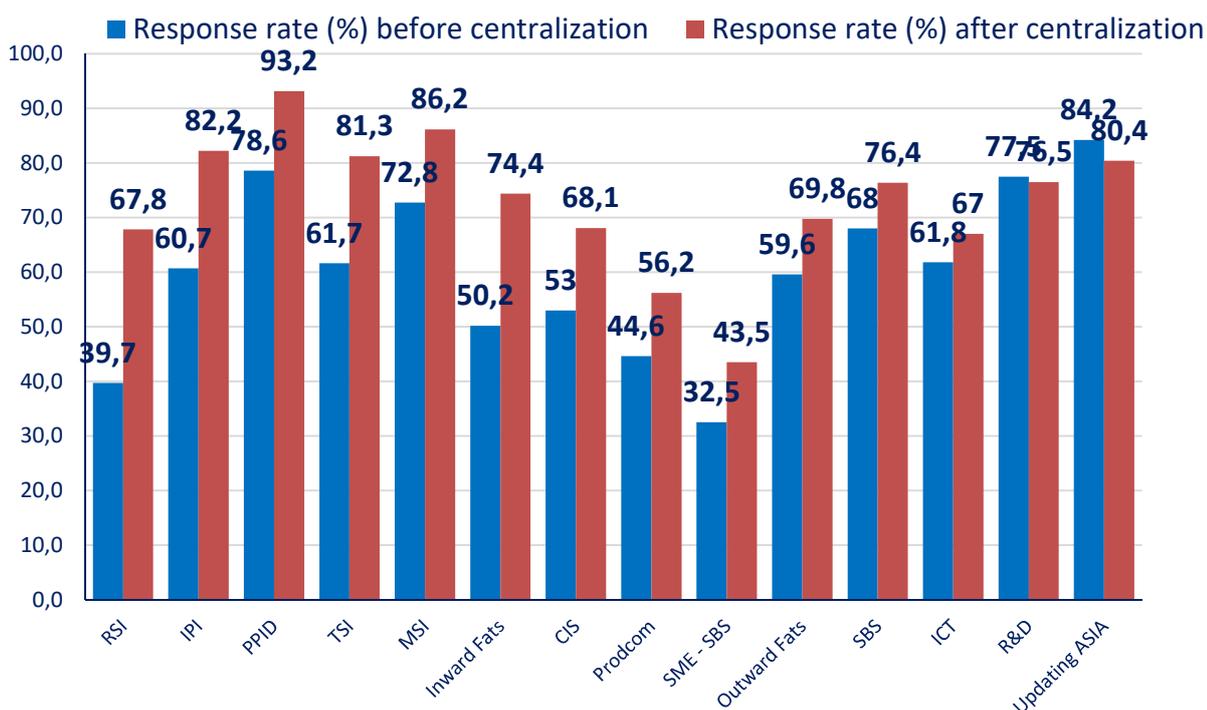
² Community innovation survey (CIS), Statistics by product (Prodcom), Small and medium enterprise survey -SME (including professional and artistic activities) (SBS), Survey on information and communication technology in enterprises (ICT), Survey on enterprise accounting system (SBS), Survey on Research and Development in enterprises (R&D), Statistical Business Register (SBR) update and local units (LUs) Survey, Survey on the activities of foreign controlled enterprises resident in Italy (Inward Fats), Survey on abroad foreign affiliates activities controlled by Italy (Outward Fats).

about 11 percentage points (pp), whereas for short-term³ surveys of about 20 pp. The comparisons refer to surveys on enterprises carried out before and after the centralized Data Collection (CDC) management was implemented in Istat. For structural surveys, the comparison was carried out between last concluded survey (2017-2018) and the last run before CDC introduction (2015-2016); for the short-term ones, the comparison was carried out considering 1-st semester 2016 and 1-st semester 2018.

The following Figure 3 shows that the average response rates (*rr*) for single surveys carried out before and after CDC implementation.

Compared with the first semester 2016, the *rr* shows a positive average variation of 19,4, particularly relevant are the increases of 28,2 percentage points (pp) registered for the Monthly survey on retail sales (RSI) and of 21,5 for the Monthly survey on industrial production (IPI). The Monthly survey on producer prices for industrial products sold on the domestic market (PPID) and the Quarterly survey on turnover in the services (TSI) also show significant increases of 14.5 pp and 19.6 pp respectively.

Figure 3 - Structural and short-term surveys: average response rates before and after CDC



Source: Elaboration on data extracted from Business Statistical Portal

³ Monthly survey on retail sales (RSI), Monthly survey on industrial production (IPI), Monthly survey on producer prices for industrial products sold on the domestic market (PPID), Quarterly survey on turnover in the services (TSI), Monthly survey on sales and orders (MSI)

2 The implementation of a territorial model to improve the efficiency of data collection

3.1. Introduction

In the new organizational framework Istat started a new project that assigned to territorial offices⁴ a central role, in carrying out several cross-cutting data collection activities on the territory. In particular, Istat territorial offices were entrusted with several tasks typical of centralized Data Collection offices: checking and updating the lists of companies involved in the surveys, carrying out and monitoring data collection on the territory, providing support and assistance to users. The new model assigns a specific role to one leader territorial office that carries out a coordination of the data collection activities for all the offices in the territory. The same lead office also maintains relations with the structure dedicated to managing the implementation of data collection at central level (office conducting data collection from direct surveys- RDC). The selection of the lead office is based on the experience accumulated in the specific subject area investigated. So, for the first time in Italy the activity of the Territorial offices in the field of data collection was not limited to the territory under its jurisdiction but is extended nationwide. During the start-up phase, the new management project for data collection on the territory is limited to a small number of activities: survey on Maritime transport, survey of Road accidents, Demographic statistics. At the conclusion of this opening phase, it is planned to extend it to other thematic sectors such as structural business statistics, tourism, environmental statistics. In order to better describe the new role of Territorial Offices (UUTT) during the data collection phase it will be presented in the next paragraph the case study of the Territorial Office for Campania and Basilicata (RMH) that, starting from the survey edition 2018, co-ordinates the activities of data collection on Maritime Transport statistics, in collaboration with the Istat Section for the implementation of data collection from direct surveys (RDC) in the Central Directorate for data collection. During the start-up phase of the project, all the data collection activities taken over by the central structure were transferred to the territorial office through videoconferences and face-to-face meetings held at the headquarters in Rome. In particular, the transfer concerned the use of the tools used for data capturing (Tramar), for monitoring the DC, for the management of reminders, for assistance and support to respondents. Then they were entrusted with the main activities typical of the centralized data collection such as the task of updating the lists of respondents, carrying out and monitoring data collection on the territory, provide technical support for the correct filling out of the survey questionnaires, ensuring the consistency and quality of the data acquired. In fact the Territorial offices also carry out an activity of first level checking the information reported in the ISTAT TRAMAR models, compiled by the Maritime Agents, Forwarders, Recipients.

⁴ As the Central Directorate for data collection also the Istat's Territorial offices (UUTT) were located in the new Department for data collection and development of methods and technologies for the production and dissemination of statistical information, so establishing the foundations for closer cooperation. The new organizational structure provides the following 8 offices located on Italian national territory: - Piemonte, Valle D'Aosta and Liguria RMB;- Lombardia RMC; - Veneto and Friuli Venezia Giulia RMD; - Emilia Romagna, Toscana and Umbria RME; - Marche, Abruzzo and Puglia RMF;- Lazio, Molise and Calabria RMG; - Campania and Basilicata RMH; - Sicilia and Sardegna RMI. Each territorial office may include one or more regional offices.

Since for the purposes of the survey it is foreseen a periodic sending of reminders for the transmission of monthly data to respondents, they will be contacted by relevant Istat Territorial Offices in order to collect updated contact information (i.e, phone and/or email addresses), useful to reach people authorized to carry out the operations of declaration of boarding and landing of goods and passengers.

3.2. A case study of process innovation: a coordination role to the Territorial office Campania and Basilicata for maritime transport survey

The reorganization of the data collection for Maritime transport survey, that entrusted a coordination role to the Territorial office Campania and Basilicata, has determined the redesign of the flow of interventions to support respondents and of the checks of coverage and quality on information collected. Firstly, the redesign required the revision of methods for monitoring the data collection trend by redefining the reports deducible from the dedicated web application (named Tramarint), that is both an application for monitoring the survey trend, and a collaboration tool between Territorial Offices and Maritime Agencies. It offers facilities to manage the exchanges of information between ISTAT and the Maritime Agencies, in particular for the requests of password for accreditation to the data capturing system, for the modifications of the delegation powers, for restoring questionnaires to be reinserted or modified. This last functionality allows, by operating on Tramarint back-office system, to support respondents for changes and restoration of previously entered questionnaires and which must be made accessible again to the Agency's completion. Tramarint is currently accessible only by the Territorial Office, to which it returns the views on the summaries of the inserted questionnaires and a summary of the compilation of the relevant questionnaire sections in relation to the specific characteristics of the ships surveyed, the goods moved, the passengers transported.

The Tramarint monitoring system, when its management was taken over by the Campania and Basilicata Territorial office, presented only partial reports. It was therefore necessary to expand the contents of the reports and designing and implementing a new layout in order to reconstruct the entire survey year for every single movement of ships arriving or departing. With the new layout, the information refers to the whole registry of the marine agency and ship's marine data and in particular the ship code IMO (International Maritime Organization), a unique code assigned to the vessel when the keel was laid by the IHS Fairplay⁵, ex-Lloyd's Register - Fairplay.

The monitoring activity is carried out on the basis of two archives: the first is TRAMAR, the ISTAT site which collects the flow of the data of the survey with ISTAT ownership, according to Regulation (EU) No. 1090/2010 prescriptions, the second consists of the ADES (Arrivals Departures Enhanced Statistics) archive, which is fed by the flow deriving from the compulsory administrative records for each arrival and departure of ships from/to any Italian port by the Shipper. The Recipient is the legal figure assuming all the responsibilities related

⁵ Persons performing ship recommendation activities, assisting the master in respect of local authorities or third parties, receiving or delivering goods, embarking and disembarking passengers, acquiring freight, concluding transport contracts for goods and passengers with release of the relative documents, as well as any other activity for the protection of the interests assigned to him.

to the various aspects of port security and tax declarations on goods and passengers transported, with an additional requirement for cruise ships to communicate the list of passengers to the Ministry of the Interior for national security purposes. The Statistical Archive ADES is transmitted to Istat by the Statistical Office of the Ministry itself that acquires the elementary data (single trip of each ship identified by a code that is called `visit_id`) in possession of the Port Authorities and available for large and medium-sized Italian ports through the PMIS Portal.

All the complexity associated with the identification of ships and agencies to be subject to reminder or quality control mainly concerns the enucleation of the following cases: alleged duplication, compilation of only one of the two types of travel (arrival only or departure only), excess of declarations of empty vessels in cabotage movements for which it is presumed that the ship is unlikely to arrive or depart effectively empty. It follows the need for a continuous exchange between the National Coordination Office and all the local offices, concerning the issue of the various territorial completeness checks of ships operating between the same ports of origin and destination. In such situations, very frequently it is necessary to determine both the correspondence of the number of journeys and the correspondence in the definition of the type of ships and, as previously mentioned, the truthful correspondence of the indication of "empty ship". For this reason, in order to support information exchanges the Campania and Basilicata Territorial Office deemed it essential to activate a specific collaboration web area using the opportunity offered by the Istat Intranet. In addition to the need for frequent exchanges between the various territorial offices of the files monitored and the results, it allows to discuss any changes found by Istat territorial referents before requesting the Maritime Agency to formally communicate the changes. Only after the opportune exchanges of information between the territorial offices and the Istat lead office of Campania and Basilicata the formal communication channel is activated. The collaboration area is also used to share on the territory the survey administrative documentation among all the actors involved in all the phases of the survey, notably between ISTAT and the maritime agencies (Example of this documentation are the informative letter to start the survey, the monthly alerts and the quarterly and annual reminders).

3.3. Preliminary results

The case study of the Campania and Basilicata Territorial office also demonstrated the active role of the UUTT in the DC activities as a national reference, in fact the first results point out the capacity of a Territorial office to assume the role of leadership in the management of the DC activities, of coordinating the activities of other offices / territorial entities as well as the role of reference to the central structures that deal with the management of the collection. A first phase involves initial critical issues related to the consistent use of resources of the Central Data Collection Directorate, due to the rigidity of the existing management systems; the difficulty of extending the harmonization and rationalization objectives typical of data collection on the territory and the tendency to "mix" thematic and non-thematic topics. In the experience of maritime transport survey, the Campania and Basilicata office also demonstrated the ability to manage links with any intermediate bodies involved (other than ISTAT). The first results obtained showed a strong

rationalization of the collection process with deep revisions of the modalities of relations with the agencies involved in the survey, in the roles of the external and internal actors involved and in the checks on the data made during the DC phase. The case study considered, moreover, represents a prototype model that can be of reference for the application in other research contexts and it may also be extended to other statistical surveys in the future.

3 The definition of the final disposition codes in the household surveys in the context of the new integrated management system in Istat⁶

4.1 The system of survey outcomes underlying the workflow of the new integrated management system for the surveys (SGI)

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 design of an 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 implementation of the activities of implementation, monitoring and control of the survey.

In fact, the workflows that SGI regulates 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 unit 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

⁶ The authors, N. Cecconi and R. Ranaldi, thank M. Morricone for the valuable contribution provided in the research.

have a specific disposition code (outcome of the attempt, sometimes called "temporary disposition code")⁷; 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, ...).

4.2 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, about 180.

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

⁷ Fieldwork activities include not only the activity carried out by the interviewers in contacting the survey units to obtain the interview, but also the inbound and outbound activities carried out by the contact center operators aimed at providing assistance and clarifications to the respondents, and sometimes even to make reminder actions against non-responding units; in particular, this activity is of fundamental utility in the assignment of the outcomes in the CAWI surveys, which do not provide for the presence of an interviewer.

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).

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 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).

The following is a proposal for standardization of synthetic outcomes.

Table 2 – Description of the synthetic final result

Code	Synthetic final disposition code	Description
1	Complete questionnaire	It includes: -completed questionnaire -not completed questionnaire but valid (with sufficient information)
2	Refusal	It includes the refusal both at introduction / before interview and during the filling in the questionnaire causing a break-off
3	Unit unable to participate	It includes the unavailability to participate due to: -serious illness -away throughout field period -unit composed by all minors -IT barrier (does not have / use internet) -unit living in another Italian municipality (if it is eligible, but survey methodology does not provide for the tracing rules)
4	Worked-out call-counting meter	It includes situations in which the minimum number of attempts required to obtain an interview was reached. It includes both the situations in which eligibility has been ascertained (for example also units for which the completion of the questionnaire has already begun, but not completed due to excessive unsuccessful appointments), and the situations in which eligibility has not been ascertained.

5	Unreachable unit	<p>It includes both</p> <p>Unreachable eligible units, i.e.:</p> <ul style="list-style-type: none"> -the unit lives in the municipality but the precise address is not know -the unit lives in another Italian municipality (but without precise information) <p>and</p> <p>Unreachable units with unknown eligibility, i.e.:</p> <ul style="list-style-type: none"> -the unit is unreachable for unknown name/address/telephone -Unable to enter building/reach housing unit (i.e. natural disasters, atmospheric events)
6	Not eligible unit	<p>It includes not eligible units:</p> <ul style="list-style-type: none"> A. because they no longer exist (eg death) B. for changes in status (eg unit transferred abroad) C. because out of the sample, i.e.: <ul style="list-style-type: none"> --seasonal/vacational/temporary residence, --unit living permanently in an institution /community, --group of single people who do not constitute a household, --shop, enterprise, company, office, -- qualification not achieved in the reference period, --quota filled
7	End of the fieldwork period	<p>Outcome assigned by office and corresponding to units for which contact attempts have begun and it has not been possible to assign a final outcome before the end of the fieldwork period. It includes all those units that at the end of the fieldwork period the status of contacts is "Attempts in progress" and the eligibility status "Eligible" (i.e. units for which the completion of the questionnaire has started, but not completed before the end of the fieldwork) or "uncertain eligibility".</p>
8	Not attempted unit	<p>Outcome assigned by office and corresponding to units assigned but without any attempt of contact, because they are distributed to the interviewers network but never attempted by the interviewers; in particular, this last case includes all those units that at the end of the fieldwork period have the assignment status "assigned to the final interviewer" and status of contacts "Attempts to be started"</p>

9	Unit not worked by an intermediate body of data collection (municipality, private company, etc.)	It includes units assigned but without any attempt of contact, because they are not distributed by the intermediate bodies of data collection (eg due to unavailability of an interviewer); in particular, this last case includes all those units that at the end of the fieldwork period have the assignment status "assigned to intermediate body" and the status of contacts is "Attempts to be started"
10	Not assigned unit	It includes units not assigned by ISTAT to intermediate bodies of data collection eg for natural disasters, errors in allocation procedures, etc.

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 systems⁸ (Information System on the Documentation of surveys) 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.

4.3 Specifications for the integrated management system and work in progress

The activity on the survey outcomes system has so far led to the preparation of technical specifications for the new integrated management system, for the implementation of the survey Diary, currently in use for the population census. Furthermore, a specific document

⁸ SIDI is a structured system of quality support whose aims are: 1) integrated management of metadata and quality indicators, 2) support for quality control for survey managers, 3) support for decisions on quality improvement strategies, 4) quality communication to users

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.

In particular, in order to assign the analytical final disposition codes, during the survey process it is necessary to populate / update: 1. the assignment status of the unit which indicates at which assignment stage each unit of the theoretical sample is located, 2. the outcome of the contacts indicating the stage in which the contacts with the assigned units are located, 3. the state of eligibility of the unit and 4. the status of the questionnaire which indicates the level of completion of the questionnaire. The different combinations of status and temporary outcomes determine the final disposition codes.

The Istat Division for "Data collection implementation from direct surveys" is also called in the short term to contribute to the specifications on the modules 'Summary Reports' and 'Payments', in the definition of the functionalities to provide for ensuring the correct use to the outcomes system.

Meanwhile, the started work has already allowed the first interventions also on the acquisition tools of the survey outcomes themselves. In fact, looking forward to the progressive use of the new management system, gradually for all the surveys, it is now necessary to revise and / or rethink the acquisition tools for the survey outcomes (and the related flow and unload/download rules).

In fact, on this side too, we are in the presence of a highly variegated situation. Depending on the characteristics of the survey (sampling design, survey mode and interviewer network) the degree of standardization and automation adopted in the management of contacts and in the outcomes data entry can be more or less high. In the case, for example, of the Census, the municipal interviewer manually records the contacts outcomes in the survey diary, while in the case of other CAPI surveys outsourced to private companies (with own interviewers and data collection systems not developed in the Institute), the outcomes – temporary and final - are gradually assigned in compliance with precise rules, which derive from the use of different tools (interviewer diary and, above all, contact forms, that are real questionnaires implemented in the system that regulate contact flows and which contain questions aimed at verifying the eligibility of respondents, the compliance with the survey timing and with the replacement criteria of the units).

4.4 Conclusions

The process of comparing and integrating the outcome systems is finalizing. In addition, initial changes to the questionnaire for contact forms are already being introduced, albeit gradually, to make them suitable for the needs of the new general survey outcomes system, in order to standardize the label of the same categories of outcomes, while respecting the specificities of the survey.

4 Personalized reminders: the experience in the Italian survey on the vocational integration of research doctors⁹

5.1 Introduction

In the first half of 2018, Istat carried out the survey on the vocational integration of research doctors with the aim to know the times and the methods of the vocational integration of those who concluded a research doctorate course in the years 2012 and 2014, at a distance of 3 and 5 years from the achievement of the degree.

The survey was conducted using CAWI mode through a web portal that provided to the respondents in addition to the online questionnaire, also informative materials and support functions. The 22,241 research doctors involved in the survey received the invitation to participate and the instructions for filling in the questionnaire through a letter signed by the Istat President sent by mail.

A private company, on behalf of Istat, provided a contact center service by means of a dedicated free number and a dedicated email address: the company provided an inbound service of support to the respondents for requests of information and clarifications about the survey and its aims and for difficulties showed up in the phase of the access to the web portal or during the compilation of the questionnaire; in the same time the private company provided also an outbound service by means of reminders by email and by telephone to the non respondents. With reference to this last aspect, during the whole fieldwork period, the survey design provided for a massive weekly submission of an email remainder and a constant telephone reminder to those who had not yet sent a completed questionnaire.

This is the first edition of the survey conducted after the recent (April 2016) creation of the Istat Central Directorate for Data Collection, and has represented one of the fields for the development of strategies aimed at harmonizing and making more efficient the data collection operations for the surveys conducted by the Institute.

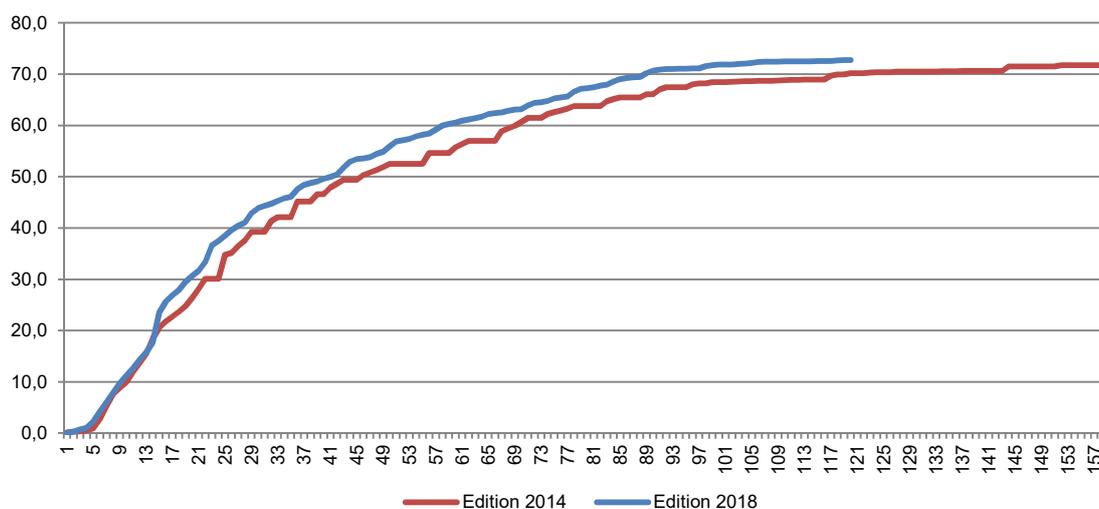
In this context, the activity of the Istat staff is included, which has constantly monitored the field data collection as well as the activity of the contact center; in particular, its role has been central in the drafting of personalized feedbacks to non-respondents and the identification of specific targets.

5.2 Data collection performance

Before dealing with the activity of the contact center reminders, some of the main results achieved are reported below. Field data collection lasted about 3 months and a half, from 15 February to 31 May 2018. At the end of the data collection, after approximately 120 days, 16,185 completed questionnaires out of 22,241, the 72.8%, have been received. The result is certainly positive if compared to the one obtained in the previous edition of the survey in 2014, where the final response rate was 71.8% in 158 days of fieldwork. Figure 4 compares the cumulative daily response rates for the 2014 and 2018 survey editions.

⁹ The author, R. Ranaldi, thanks E. Vergura and E. Mercuri for the valuable contribution provided in the research.

Figure 4 – Cumulative percentage of questionnaires daily transmitted by number of fieldwork day – Editions 2014 and 2018



5.3 The email remainder activity of the contact center

The outbound service provided by the private contact center took place between 15/02/2018 and 15/05/2018 and concerned about 102,300 emails sent to non respondents and 165,000 telephone contacts.

Specifically, email outbound activity has provided for 12 weekly reminders according to a schedule designed by Istat.

The first reminder was sent massively to all the doctors who had not yet sent a completed questionnaire. All the email addresses available in the registry were used, both those provided by the universities and those found by Istat with ad hoc searches. For the doctors who registered but had not completed the questionnaire, the reminder was also sent to the email address entered by the respondent during registration phase.

The text of the emails, different for each of the 12 reminders, in addition to remembering the objectives of the survey, provided again the log-in information and invited those who were in trouble to complete the questionnaire to contact the toll-free number or write an email to the dedicated email address describing the problem encountered.

Figure 5 – Number of questionnaires transmitted by day and date of the email reminders

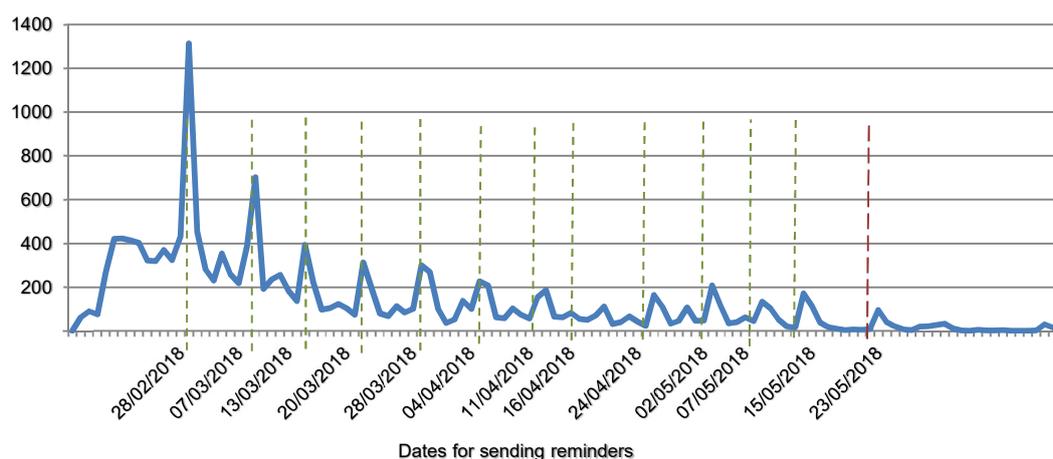


Figure 5 shows the trend of the questionnaires transmission by data collection day and the dates in which the email reminders have been sent. We can see an absolute peak in the number of daily questionnaires sent in correspondence of the first email reminder and in any case relative peaks in the days immediately following the sending of the other email reminders. On the date of the 4th email reminder (20/03/2018), 34 days from the beginning of the fieldwork, 45.8% of the questionnaires were sent. After May 15, the web portal was still open to respondents until the end of the same month to allow the Istat staff to receive the latest completed questionnaires, as well as to carry out some specific action (23/05/2018), for example to doctors who during their contacts had expressed their refusal to participate.

After the first reminder, sent in a massive and undifferentiated way to all the doctors who had not yet sent the questionnaire, much care has been given to the preparation of personalized texts of the subsequent reminders. The objective was indeed to remember the survey participation to the research doctors, but above all to convince him to participate as quickly and easily as possible, providing in the email text all the information necessary to proceed from the moment of receiving the e-mail. For this aim, three main targets have been identified based on the stage in which each doctor was. After receiving the initial invitation letter, the research doctor had to initially register following a certain procedure, then start filling in the questionnaire, and then conclude and transmit the questionnaire. A specific text has been therefore prepared for reminders to doctors who still had to register (*reminder to unregistered doctors*), with the link to the web portal, the log-in information and the procedure for the registration; a second specific text has been prepared for the reminder to the doctors who registered but did not start filling in the questionnaire (*reminders to registered doctors*), with the link to the web portal and the new log-in information (different from the previous ones), as well as the invitation to filling in the questionnaire; finally, a third personalized text has been designed to remind the doctors who had registered and started filling in the questionnaire but without terminating it (*reminders to doctors who started the compilation*), with the link to the web portal and the invitation to the conclusion of the questionnaire.

The reminders have been sent to all the email addresses available in the registry, both those provided by the Universities and those found by Istat with ad hoc searches; for the doctors who instead registered, the reminder has been also sent to the email address entered by the respondent himself at the first access, since in this phase it was mandatory to insert an email address.

Figure 6 – Number of reminders sent to the doctors who had not yet submitted the questionnaire (absolute figures)

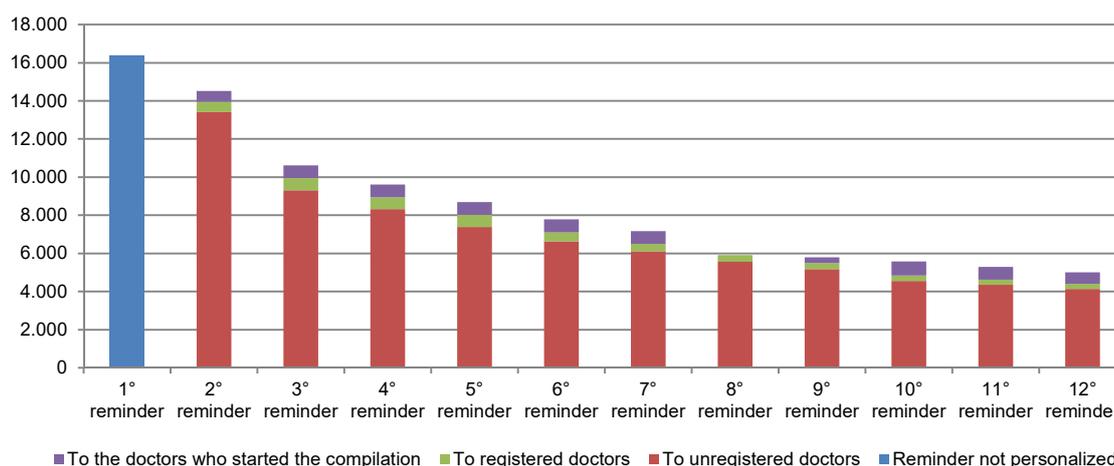


Figure 6 shows the absolute number of reminders sent on 12 occasions depending on the target group, that is the type of customization included in the reminder text. In an initial phase of the fieldwork, the share of personalized reminders sent to those who had not done any action on the web portal was preponderant, and then decreased with the conclusion of the survey.

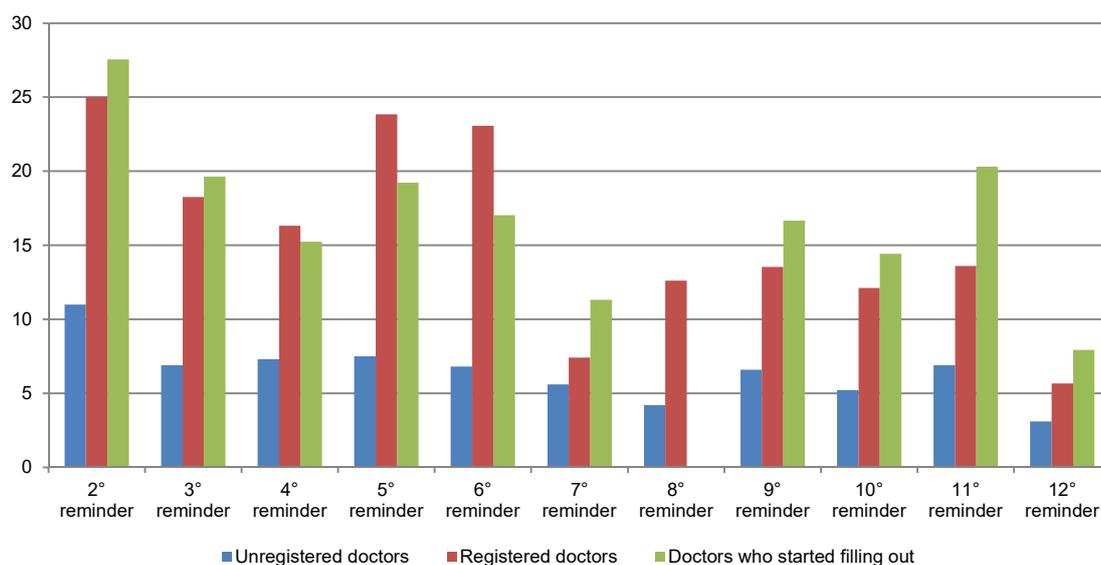
5.4 Reminders effectiveness

In order to estimate the effectiveness of each reminder, the percentage of those who submitted a complete questionnaire before the next reminder has been calculated on the total number of doctors to whom a specific reminder was sent. In this case it can be assumed that the questionnaire submission is the result of the specific reminder.

Figure 7 shows the percentage of questionnaires submitted following the various types of reminder. In particular, it is noted that the percentage of questionnaire submissions is lower by research doctors who showed little interest in survey participation from the beginning, confirming it also after a lot of reminders received. The percentage, with the exception of the second reminder, has been always less than 10%.

Reminders sent to the doctors who had already shown a certain willingness to participate in the survey, having registered the web portal or having started filling in the questionnaire at the time of the reminder, have a greater "effectiveness".

Figure 7 – Percentage of submitted questionnaires by personalized reminder type and reminder occasion (percentages per 100 doctors to whom the reminder was sent)



The quality of the email addresses list used to send reminders also impacts on these results. In fact, while for the former, only the directory provided by the Universities, that is a directory sometimes obsolete also because of the specific nature of the target population, characterized by a strong territorial mobility¹⁰; for the second and third type of reminder, the email address provided by the doctor during registration phase was also used and therefore certainly valid.

The constant monitoring of the survey had revealed a "hard core" represented by an almost constant number (about 600) of questionnaires started but not completed. The reasons that led the doctors not to fill in the questionnaires are mainly due to problems in answering some questions. In this regard, a detailed analysis of the interrupted questionnaires was carried out by Istat staff, which allowed to identify some critical questions. The questionnaire was divided into a brief introductory section and five thematic sections. The analysis has highlighted some critical issues: in section 2 that collects some information on the employment status of doctors, in particular the use of a web navigator for the choice of the code corresponding to the occupation and the answering the question concerning the research activity; in section 4, which collects some information on the territorial mobility before and/or after the doctorate; in section 6, the last one of the questionnaire, in which it is necessary to verify the correctness of the data entered and to proceed to the definitive submission of the questionnaire.

The reminders to the doctors who had expressed their willingness to participate because they had begun to fill in the questionnaire, were further tailored above all to help them

¹⁰ The reference population includes research doctors of any nationality, who concluded a research doctorate course in an Italian university.

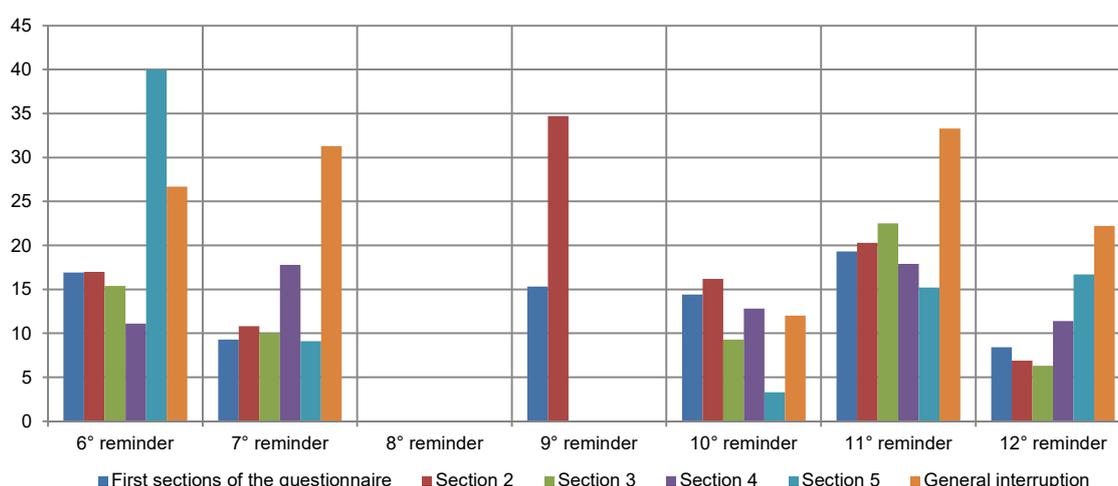
overcome any problems they encountered in the sections where they interrupted the filling in.

In particular:

- for the 6th and 7th reminder a mobile text was added which varied according to the last section saved; the exact point of the questionnaire in which he had interrupted the compilation was then communicated to the doctor and how he had to proceed;
- the 8th reminder was not sent to avoid excessive statistical burden;
- the 9th reminder was sent only to doctors stopped in the first sections and in section 2;
- the 11th and 12th reminders were provided for more detailed texts, mentioning one by one, the blocking questions and how to proceed.

Figure 8 shows the percentage of questionnaires submitted following of these targeted reminders. The 6th and 7th reminders were particularly effective for the doctors who had stopped filling in section 5 and for those who had interrupted in a part of the questionnaire not considered critical (generic interruption); the 9th reminder has unblocked the submission of many doctors still in section 2; particularly effective on all the sections of the questionnaire was the 11th reminder, the first with the most personalized texts and the solutions to be adopted to overcome the critical questions.

Figure 8 – Percentages of submitted questionnaires by personalized reminder type according to the last saved section (percentages per 100 doctors to whom the reminder was sent)



5.5 Conclusions

The adoption of email reminders has been very effective (certainly more timely than a telephone reminder) and even more the adoption of personalized reminders on a weekly basis throughout the fieldwork period.

The solution to use reminders with personalized texts has certainly helped to resolve some critical issues in a targeted manner, obtaining higher response rates, especially over a shorter fieldwork period. The personalized reminder, in addition to being a support to the doctor in the compilation, makes him feel more involved and therefore more inclined to continue.

A design of repeated reminders has on the one hand the advantage of encouraging dormant doctors to fill in the questionnaire, on the other hand the risk of annoying reticent doctors for participating. To obviate this, the text of the email object has also been modified on each sending occasion and this has at least intrigued the respondent to read a text apparently different from those previously received.

In a future perspective in the conduction of similar surveys, we intend to set up experimental designs useful for testing the effectiveness of one text rather than another one on the same targets.

5 General conclusions and future developments

The introduction of the new organizational model launched by Istat in 2016, which provided a specialized approach to the management of cross-cutting services and the creation of a new Department exclusively dedicated to the Data Collection has produced important results in terms of increasing response rates significant reductions in the data collection periods, product and process innovations. The results are independent of the platforms used for web data capturing and are extended to all types of surveys. Among the factors that most explain these increases has to be considered the standardization of data collection processes that led to significant increases in efficiency.

Focusing on the activities concerning the "Implementation of Data collection from direct surveys" the main solutions concern *innovative tools and services* supporting DC activities and process innovation and optimization that involved significant gains in terms of process efficiency

The experiences reported in this document represent examples of solutions oriented to the standardization and harmonization of data collection management processes in very diversified contexts:

- The development of centralized systems for the acquisition and return of data to the units involved in the surveys and the design of a single inbound and outbound contact center are examples of reference made in the context of economic surveys;
- The case study of the introduction of an innovative model for the management of data collection at the territorial level is another example that can be generalized to many other contexts;
- The standardization of the synthetic disposition codes for the classification of contacts and status, in the context of the design of a single management system for the implementation of field data collection, is another example of reference. The categories of outcomes vary in number and meaning and depend on the technical, methodological and organizational characteristics 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.

- The solution to use reminders with personalized texts according to the problems the users encountered in filling in the different sections of the questionnaire, has certainly helped to resolve some critical issues in a targeted manner, obtaining higher response rates, especially over a shorter fieldwork period. This result paves the way for further applications in other contexts and in particular in data collection for surveys managed with CAWI Technique

The examples shown represent only a minimal selection of solutions deriving from a model that envisages centralized data collection, but it is possible to produce many others in the most diversified contexts. The common denominator of all these experiences is due to the results in terms of process efficiency that the management of statistical reporting processes in an integrated and standardized way can produce.

These efficiency gains free up resources to be used in process and product innovation activities, in the quality of the outputs and to respond to new needs for statistical information expressed by users. Furthermore, the results achieved in terms of increasing response rates can trigger a reduction in the statistical burden on respondents.

Even in the presence of the above mentioned undoubted results, the new organization of the processes has also shown some critical issues that can be resolved in the medium term: i) resistance to change and increase in the conflict between transversal and production structures, mainly deriving from the "subtraction" of some activities that were traditionally managed within the production processes; ii) strong fragmentation of DC processes; iii) permanence of overlaps and doubts about "who does what" in the transversal structures and in particular Data collection. The main challenges for the future concern the methods and the solutions to be adopted to consolidate the transition process towards the new model: i) development of integrated and generalized platforms for data capturing from units belonging to different sectors; ii) design and implementation of a unique generalized system of integrated management of surveys; iii) greater integration between inbound and outbound contact center services; iv) development of acquisition Portals to increase the efficiency of data collection processes from survey units belonging to different sectors; v) identify solutions to be applied at an organizational level in order to reduce the processes fragmentation, while respecting the principle of specialization and standardization of the activities involved.

7. References

Bellini G., Monetti F., Papa P. (2018), The impact of a centralized data collection approach on response rates of economic surveys and data quality: the Istat experience. Q2018, European conference on quality in official statistics, Kraków

De Gaetano L., Digrandi A., Papa P. (2018), A territorial model for data collection implementation. Q2018, European conference on quality in official statistics, Kraków

Fazio N.R, Murgia M., Nunnari A. (2013), The business statistical portal: a new way of organizing and managing data collection processes for business surveys in Istat, Unece - Conference of european statisticians.

Istat (2017), Mapping delle attività della DCRD nell'ambito dello schema concettuale di riferimento internazionale GSBPM. Delibera D16 49 DIRM2017.

Istat (2016), Istat's modernisation programme,
https://www.istat.it/it/files//2011/04/IstatsModernisationProgramme_EN.pdf

Lise Rivais, Marc St-Denis, Susan Lensen (2013), Centralising data collection at Statistics Canada. Seminar on Statistical data collection. Unece - Conference of european statisticians.

Saraiva dos Santos P. and Moreira A. (2013). Creating a data collection department: statistics portugal's experience. Seminar on Statistical data collection. Unece - Conference of european statisticians.

Signore M. (2017) GSBPM and other international standards MedStat training on GSBPM, Istat - Rome (2017-07). Available at <https://statswiki.unece.org/display/GSBPM/GSBPM+Training+Materials>

Signore M. (2017) GSBPM how to use and implement MedStat training on GSBPM, Istat - Rome (2017-07). Available at <https://statswiki.unece.org/display/GSBPM/GSBPM+Training+Materials>

Marske R. And Stempowski D. M. (2009), Company-Centric Communication Approaches for Business Survey Response Management, Statistics Canada Symposium 2008: Data Collection: Challenges, Achievements and New Directions. Available at <https://www150.statcan.gc.ca/n1/en/pub/11-522-x/2008000/article/10983-eng.pdf?st=0-lwAKvz>

Unece Statistics wiki. Generic Statistical Business Process Model – GSBPM. <http://www1.unece.org/stat/platform/display/metis/The+Generic+Statistical+Business+Process+Model>