

*UNECE Workshop on Statistical data Collection 'Resourceful Data Acquisition'*

*10-12 October 2018, Geneva, Switzerland*

# **Centralizing data collection implementation: the Istat experience**

Giampaola Bellini - Istat ([bellini@istat.it](mailto:bellini@istat.it))

Novella Cecconi - Istat ([ncecconi@istat.it](mailto:ncecconi@istat.it))

Loredana De Gaetano - Istat ([degaetan@istat.it](mailto:degaetan@istat.it))

Francesca Monetti - Istat ([fmonetti@istat.it](mailto:fmonetti@istat.it))

Pasquale Papa - Istat ([papa@istat.it](mailto:papa@istat.it)) (Speaker)

Rita Ranaldi – Istat ([ranaldi@istat.it](mailto:ranaldi@istat.it))

# The Istat Modernization Program: main characteristics

During 2016 the Italian National Statistical Institute (Istat) launched a wide Modernization Program

The program designed and implemented a new organizational set-up characterized by the centralization of all the support services, clearly separated from statistical production

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 "transversalization" of many services pushed towards specialization, standardization and harmonization of all these services and in particular of data collection

# The Istat Modernization Program: main characteristics

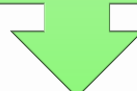
The introduction of a specialistic data collection, led to review of the organizational structure of data collection processes and to redesign of many of the management procedures



Before reorganization, statistical processes were organized according to the classical 'stovepipe' model, that involved independent, non-integrated, processes including all the necessary skills



The old approach, effective in terms of achieving the objectives set, involved low overall efficiency level, due to overlapping redundancies and lack of integration among processes



Among the main Program there is also the valorization of administrative sources for statistical purposes and the construction of an integrated system of registers

# Creation of the new Central Directorate fo Data Collection

Creation of the new Central Directorate for Data Collection: high degree of specialization of activities and Human Resources It included four Divisions specialised in the following areas

**Division for  
design of data  
collection tools**

**Division for data  
collection  
organization**

**Division for  
implementation  
of data collection  
from direct  
surveys**

**Division for  
integration of  
administrative  
sources and  
registers**

# Implementation of Data Collection from direct surveys

Focusing on the "Implementation of Data collection from direct surveys", the goal of specialization was pursued concentrating activities typical of survey's implementation in a single Division

- Internal organization was carried out according to the type of responding units involved (businesses, families and individuals, farms, public and private institutions, others)

# Representative experiences of process innovation introduced in different data collection domains

It is possible to identify several examples. Representative ones in different sectors were selected

*1. Process innovations introduced and main results achieved in business surveys*

*2. The Implementation of a territorial model to improve the efficiency of data collection*

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

*4. Personalized reminders: the experience in the Italian survey on the vocational integration of research doctors*

# 1. Process innovations introduced and main results achieved in business surveys

Process innovations introduced in business surveys are mainly based on:

1. The design and implementation of innovative **infrastructural solutions**
2. **Standardization and generalization** of each phase of the DC process

The Business Statistical Portal (BSP)

Main objectives of the Portal are:

- Streamline the operations required by respondents to fulfill their response obligations, with an overall reduction of the burden
- Increase both ordinary and extraordinary (e.g. news) communications on the survey events and activities
- Standardize and harmonize data collection in order to increase efficiency at the system level

# Innovative infrastructural solutions

BSP: Front office section



Optimizing the bi-directional communication between Istat and the companies involved in the statistical surveys.

The main features of the section concern the following aspects:

Single sign-on and single point of access	Integrated Register changes management	The delegation system	News management	The state of obligations	The personalized statistical information return
---	--	-----------------------	-----------------	--------------------------	---



# Innovative infrastructural solutions

## BSP: Back-office section

- Includes a set of functions to support the management of the survey:
  - Help-desk activities
  - Survey administration
  - Monitoring of the data collection process
  - Management of outcomes
  - Reports of variation in register data
  - Monitoring of linked users
  - Management of DB of internal /external contacts

# Innovative infrastructural solutions

## Centralised inbound and outbound Contact center services

- Progressive centralization of:
  - support and assistance services addressed to the units involved in the surveys (inbound)
  - telephone alert and reminders addressed to non-respondent units (outbound).

The coordinated management of the  
service guarantees strong standardization

# Centralised inbound and outbound services

## Inbound service

The inbound service provides 1-st level assistance and support on the following areas:

- access and navigation of the Business statistical Portal
- general rules that define the statistical activity and the legal obligations
- 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).

### Tools:

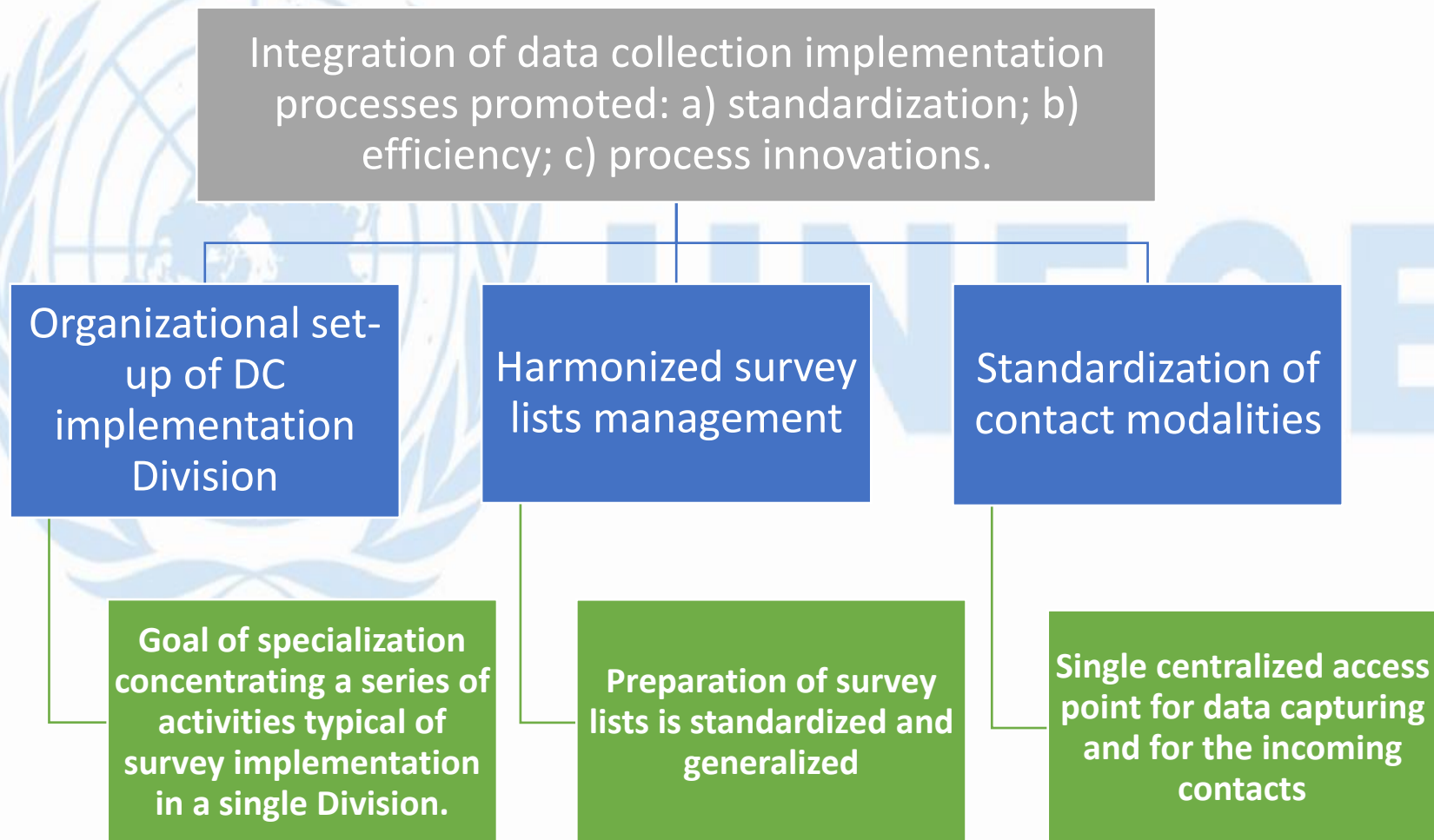
- - Set of “Standard answers” and “FAQ”
- - “**shared agenda**”, for managing and sharing the received instances with Istat’s experts (thematic, register, technical, legal experts)

# Centralised inbound and outbound services

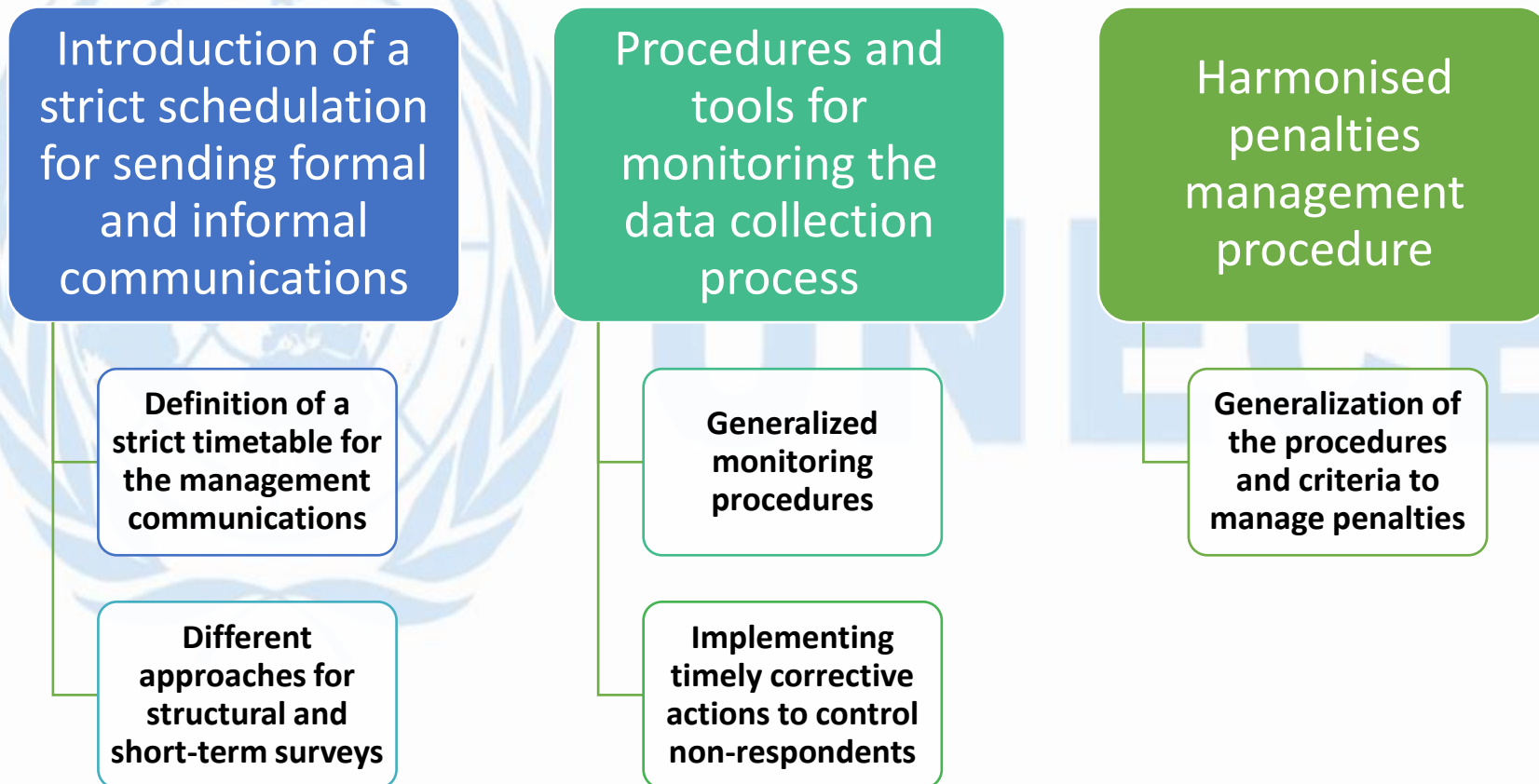
## The outbound service

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

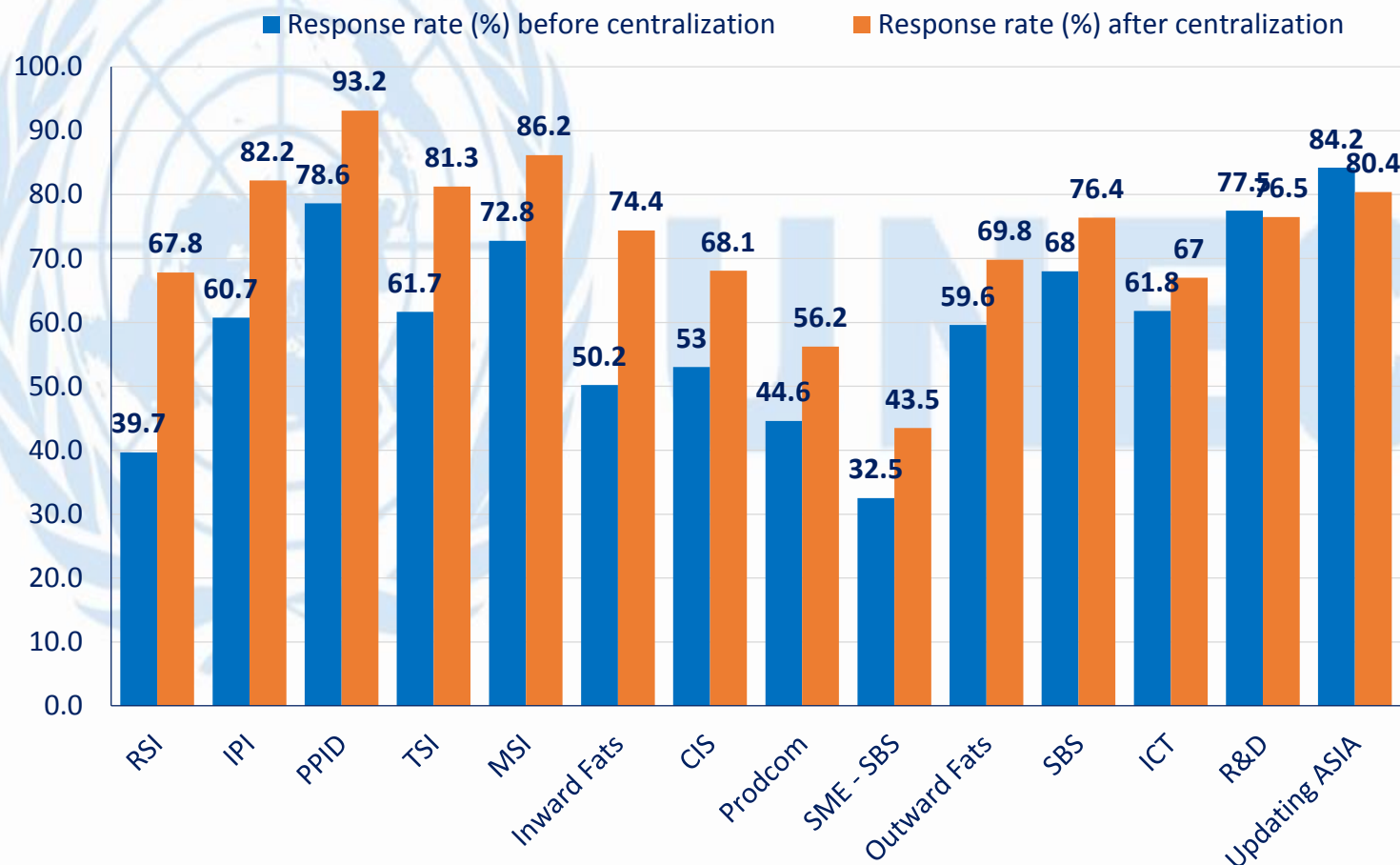
# Standardization and generalization of each phase of the DC process



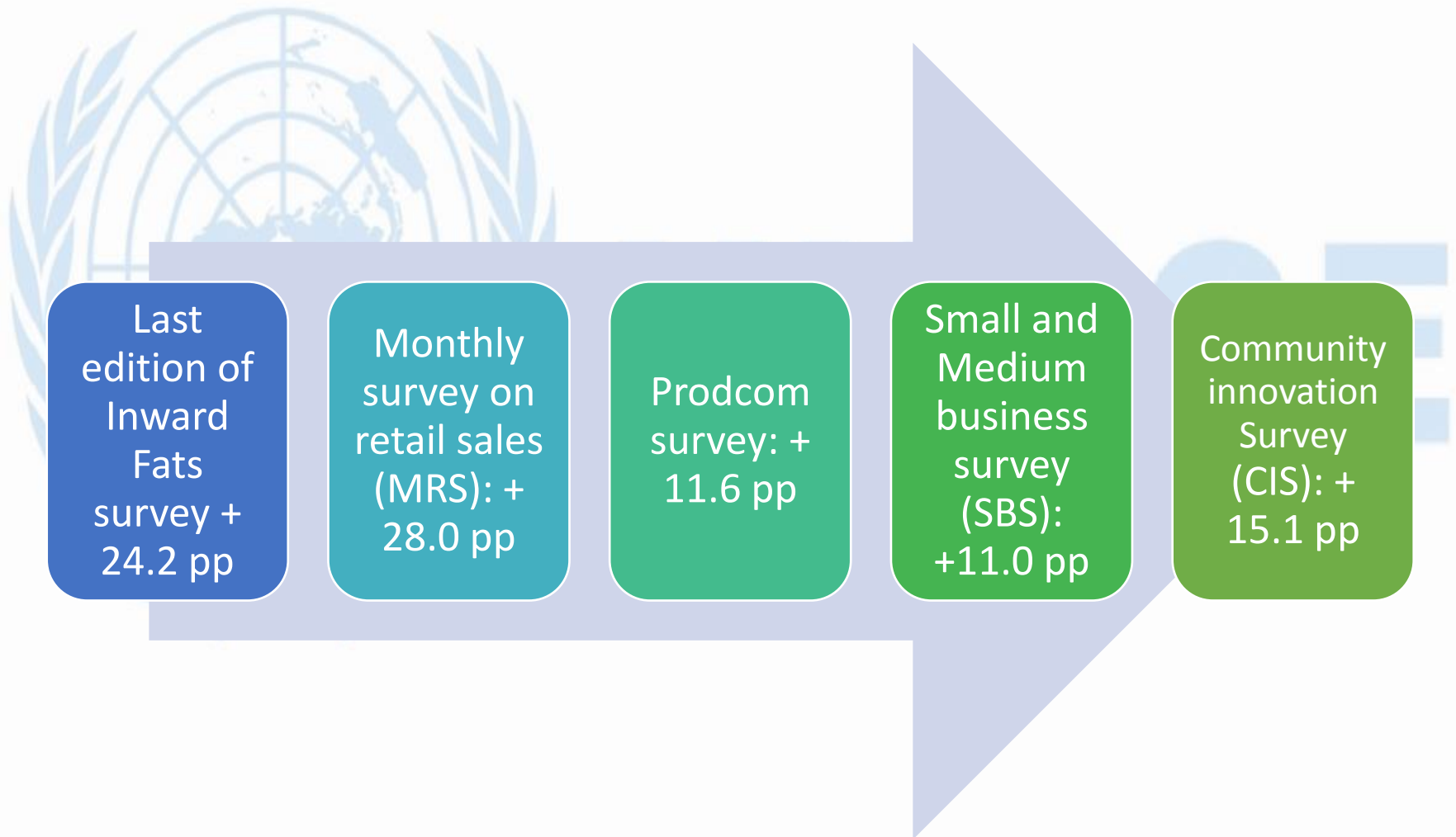
# Standardization and generalization of each phase of the DC process (2)



# Structural and short-term business surveys: average response rates before and after CDC

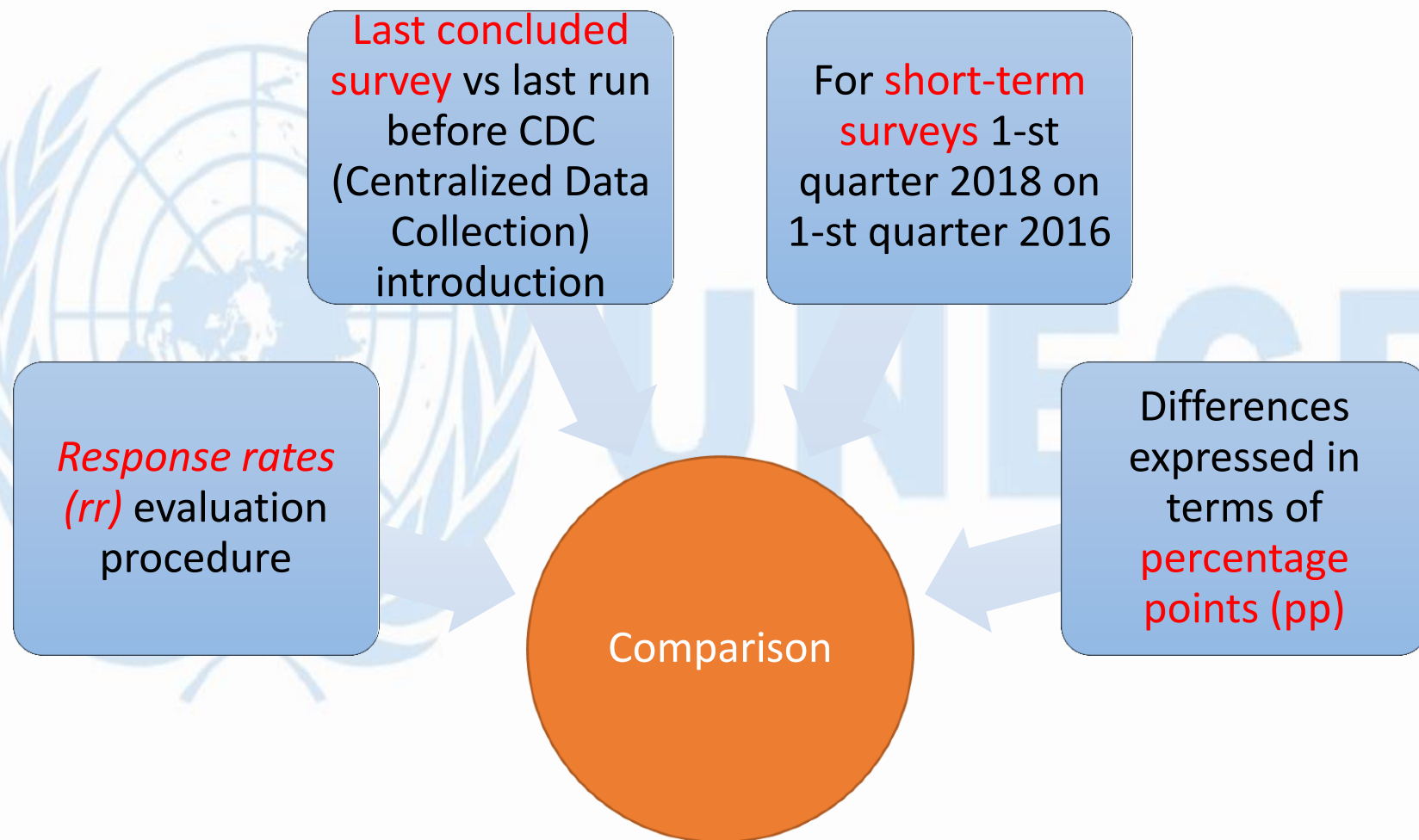


# Main results: examples of significant increases in response rates of business structural surveys





# Comparison method



# Effects on data collection periods

Average reduction  
structural surveys

• 37.2 solar days (*d*)

Small and medium  
enterprise accounting  
system survey - SBS  
Regulation)

• - 118 d

Community innovation  
survey (CIS)

• -92 d

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

In the new organizational framework an innovative project assigned to territorial offices the role to carry out several cross-cutting data collection activities on the territory

Main characteristics: of the project:

- Specific role to one **leader office** that carries out a coordination of the data collection activities for all the offices in the territory
- Lead office maintains relations with the **central structure** dedicated to managing the implementation of data collection
- 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 implementation of a territorial model to improve the efficiency of data collection

The selection of the lead office is based on the experience accumulated in the specific subject area investigated.

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

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

# A case study of process innovation: the coordination role to the Territorial office “Campania and Basilicata” (RMH) for maritime transport survey

Starting from the survey edition 2018, co-ordinates the activities of data collection on Maritime Transport statistics


## Start-up phase of the project:

- The data collection activities taken over by the central structure were transferred to the territorial office through videoconferences and face-to-face meetings
- The transfer concerned the use of the tools employed for data capturing (Tramar), for monitoring the DC, for the management of reminders, for assistance and support to respondents.

The survey monitoring system (named Tramarint) required 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)

# Preliminary results



First results confirmed the capacity of a Territorial office to assume the role of leadership in the management of the DC activities

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 case study considered, moreover, represents a prototype model that can be of reference for the application in other research contexts

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

The new Central Directorate on Data Collection promoted the design of an **integrated survey management system** (named SGI – Sistema Gestione Indagini)

The new system is used for the first time for the 2018 Permanent Population and Housing Census and will be afterwards extended to other surveys

The design of the new system enhanced the necessity 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 the definition of the **status and of the final disposition codes for all the survey units**

# The workflow of the new integrated management system for the surveys (SGI)

The **workflows** that SGI regulates are determined by

- the system of the survey **disposition codes** (temporary and final)
- 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

The outcomes system permeates the whole functioning of the integrated management system, 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



# Steps for producing final disposition codes

It is possible to identify the following steps

- Uploading of theoretical sample into the integrated management system
- Assignment of the units so that the fieldwork can begin

All the units assigned must have an outcome, which derives from the fieldwork. This result is achieved with a series of contact attempts that involves a specific disposition code (**temporary disposition code**)

The combination of the attempts outcomes, according to specific rules, produces a **final disposition code** for the unit

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

# Conceptual structure of outcomes and status

Each survey 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**

**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**

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**

# Harmonized synthetic survey outcomes

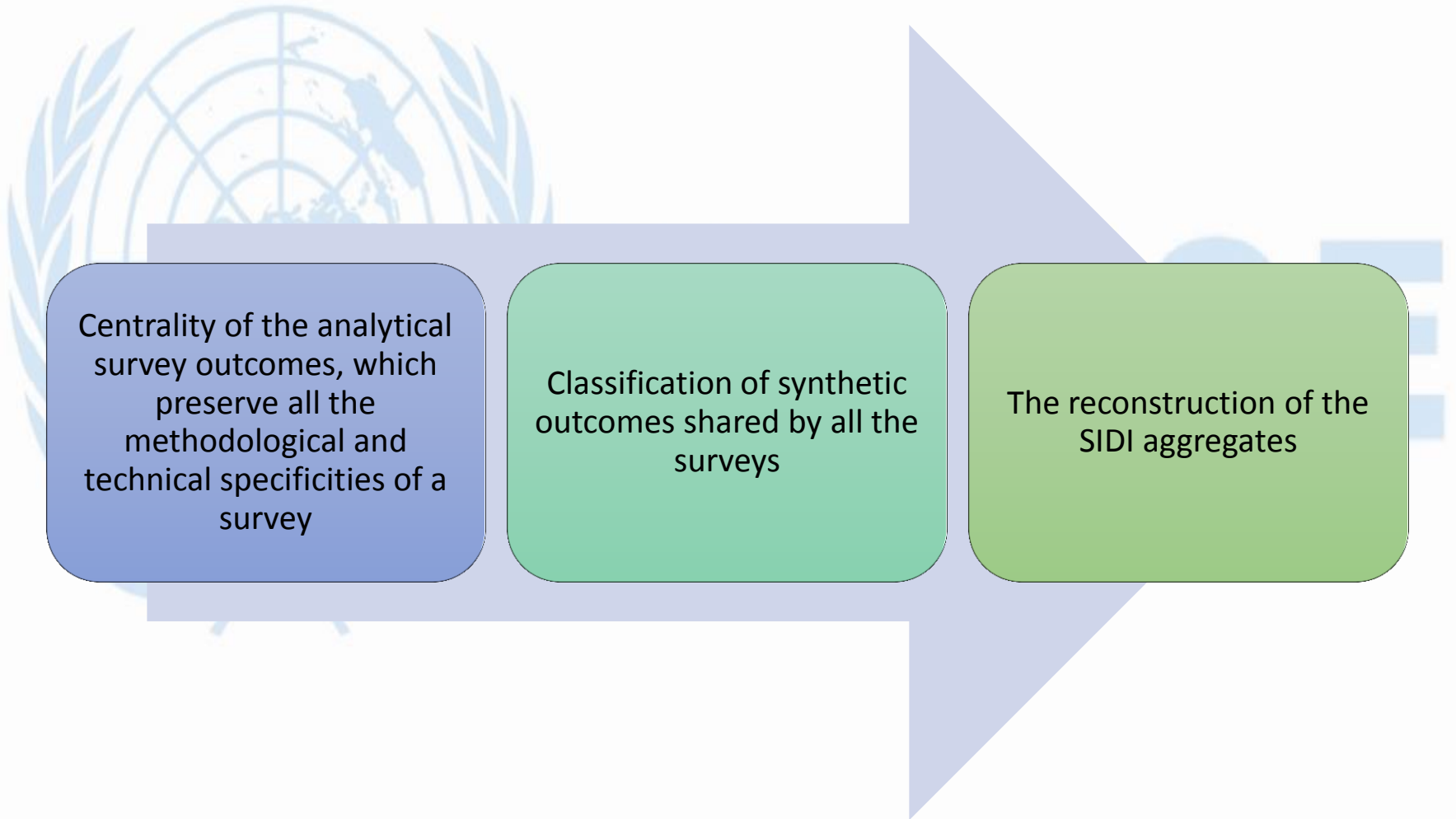
The need for the "operational" harmonization at least at the macro level is strong and useful for calculating **summary indicators of the fieldwork** and also for making a possible **comparison among different surveys**

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

Review of the existing **transcoding tables** between the analytical final disposition codes of each survey and the aggregates of the SIDI system (Information System on the Documentation of surveys) was carried out

**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

# Necessity to compound the needs of the survey management and to the generalization and standardization of the outcome categories



Centrality of the analytical survey outcomes, which preserve all the methodological and technical specificities of a survey

Classification of synthetic outcomes shared by all the surveys

The reconstruction of the SIDI aggregates

# Specifications for the integrated management system

In order to assign the analytical final disposition codes, during the survey process it is necessary to populate / update:

- the assignment status of the unit which indicates at which assignment stage each unit of the theoretical sample is located
- the outcome of the contacts indicating the stage in which the contacts with the assigned units are located
- the state of eligibility of the unit
- 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

# Acquisition tools of the survey outcomes

Standardization and automation adopted in the management of contacts and in the outcomes data entry depends on the characteristics of the survey (sampling design, survey mode and interviewer network)

The acquisition can be carried out:

- manually in the in the survey diary (e.g. by municipal interviewer)
- in compliance with precise rules, which derive from the use of different tools (interviewer diary, contact forms)

Contact forms are real questionnaires implemented in the system that:

- regulate contact flows
- contain questions aimed at verifying the eligibility of respondents, the compliance with the survey timing and with the replacement criteria of the units

## 4. Personalized reminders: the experience in the Italian survey on the vocational integration of research doctors

In the first half of 2018, Istat carried out the survey on the vocational integration of research doctors

The aim was to know the times and the characteristics of the vocational integration of doctors at a distance of 3 and 5 years from the achievement of the degree.

The survey was conducted using CAWI mode The 22,241 research doctors involved in the survey

A private company, on behalf of Istat, provided a contact center service inbound and outbound

During the fieldwork period, the survey design provided for a massive weekly submission of an email reminder and a constant telephone reminder

# Email reminders management

12 email reminders were planned, the text of the emails, different for each of them:

- It remembered the **objectives** of the survey
- provided again the **login** information
- offered **assistance** by toll-free number or by dedicated email address

First reminder was sent in a massive and undifferentiated way to all the doctors who had not yet sent the questionnaire



## Email reminders management (2)

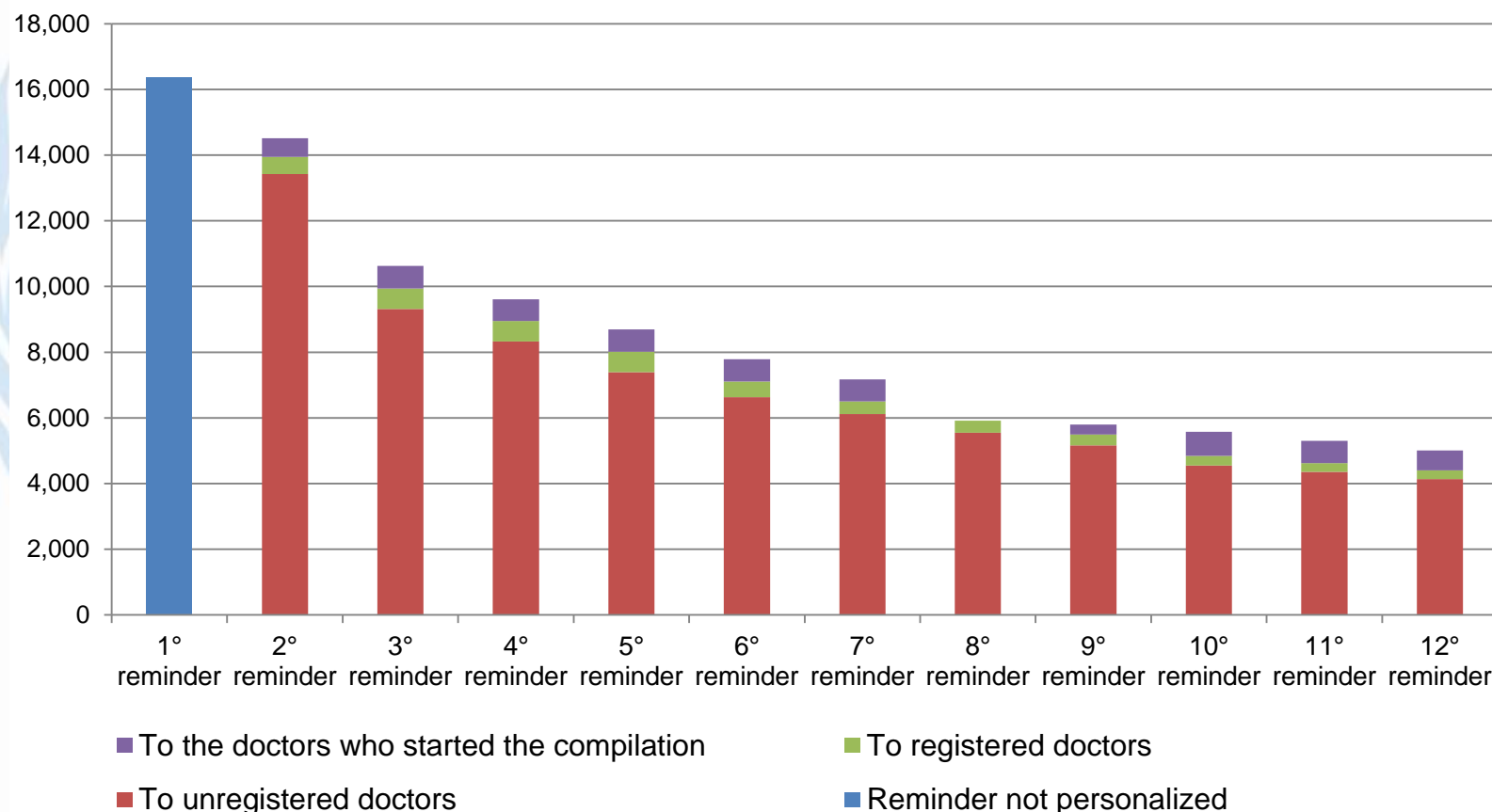
For the subsequent reminders, three main targets have been identified based on the stage in which each doctor was:

- **Login stage** (after receiving the initial invitation letter, the research doctor had to initially register following a certain procedure)
- **filling in** the questionnaire
- **concluding** and **transmitting** the questionnaire

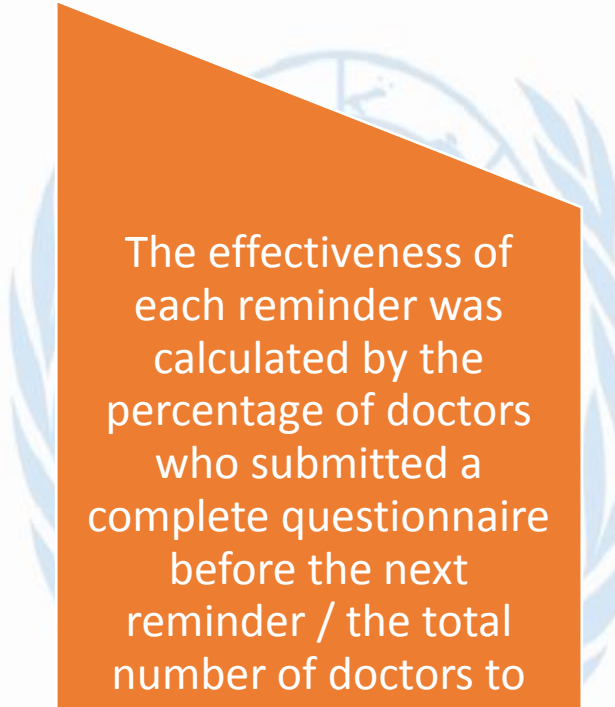
Each stage involved specific texts:

- *reminder to unregistered doctors*, with the link to the web portal, the login information and the procedure for the registration
- *reminders to registered doctors*, with the link to the web portal and the new login information (different from the previous ones), as well as the invitation to filling in the questionnaire
- *reminders to doctors who started the compilation*, with the link to the web portal and the invitation to the conclusion of the questionnaire

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



# Reminders effectiveness



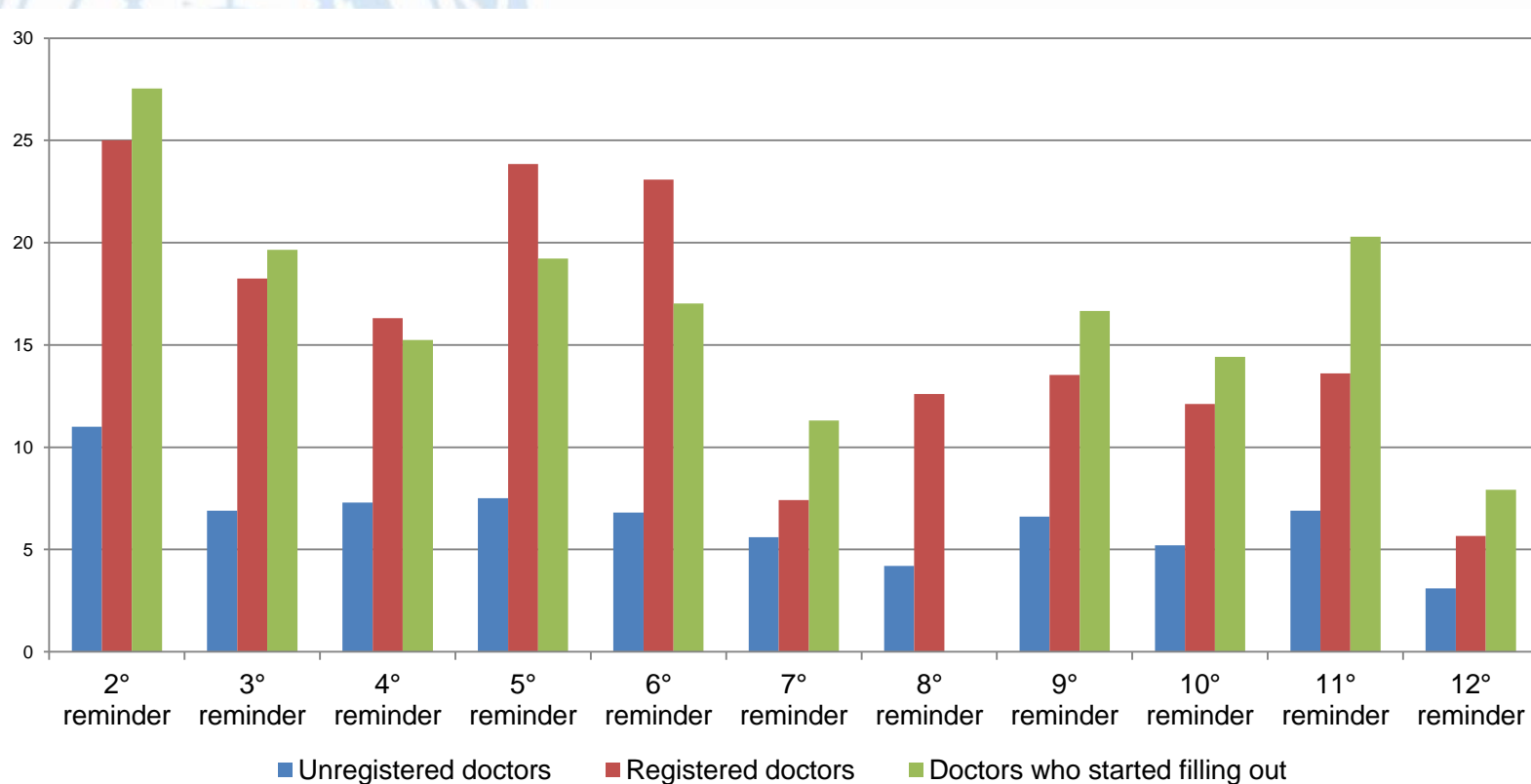
The effectiveness of each reminder was calculated by the percentage of doctors who submitted a complete questionnaire before the next reminder / the total number of doctors to whom a specific reminder was sent

It is noted that the percentage of questionnaire submissions is lower by research doctors who showed little interest in survey participation

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

# Reminders effectiveness

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



# Analysis of interrupted filling in

The constant monitoring of the survey had revealed a 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 **specific** questions

A detailed analysis of the interrupted questionnaires was carried out by Istat staff, which allowed to identify some **critical questions**:

- 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 were **further tailored to help them overcome any problems** encountered in the specific sections

# Personalised reminders

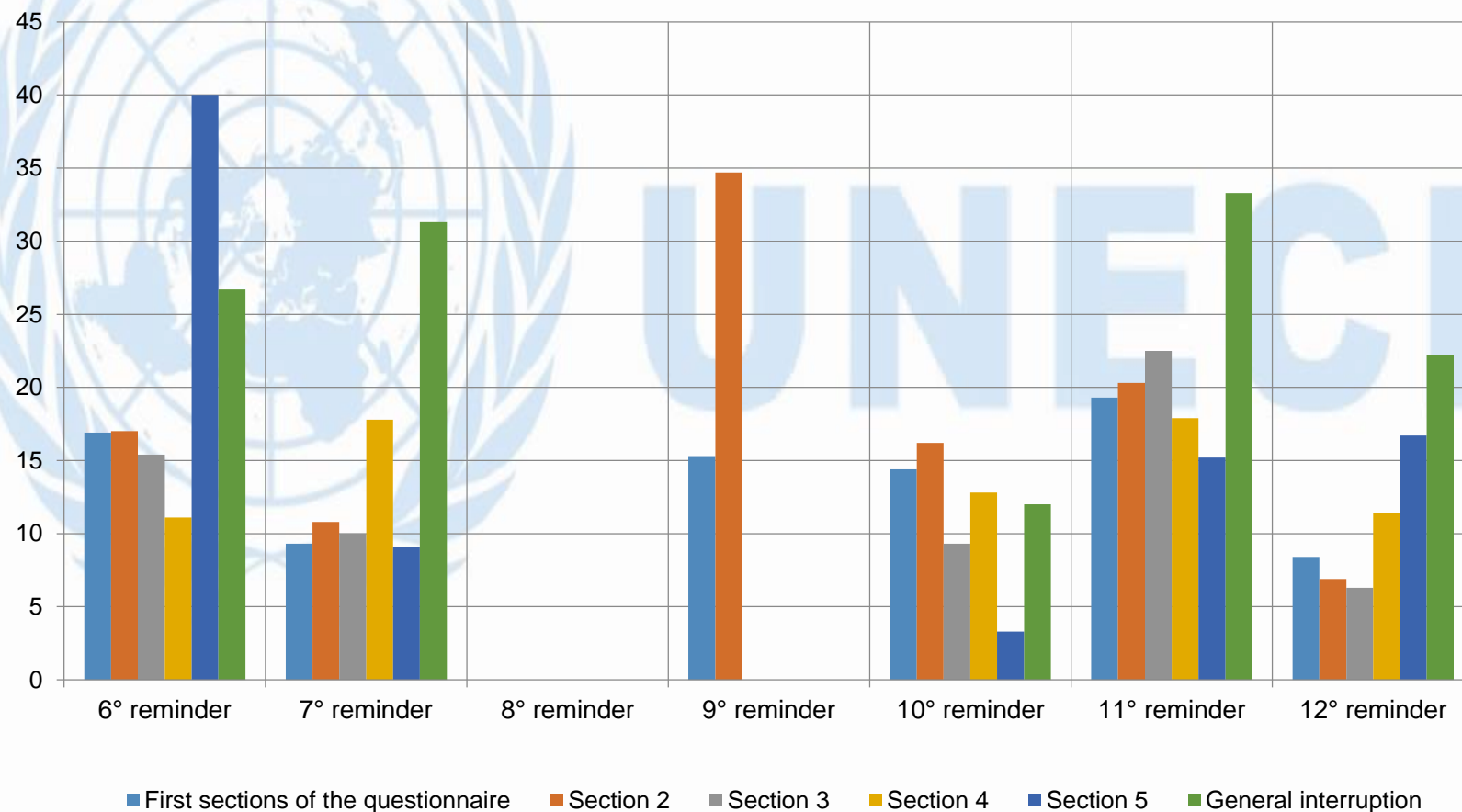
For the 6th and 7th reminder a mobile text was added which varied according to the **last section saved**; the exact point of the questionnaire in which he had interrupted the compilation was then communicated to the doctor and how he had to proceed

The 8th reminder was not sent to avoid excessive statistical burden;

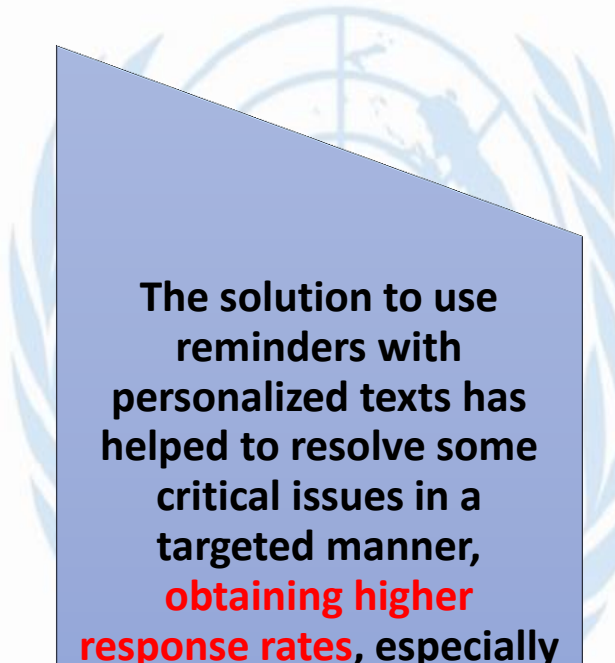
The 9th reminder was sent only to doctors stopped in **the first sections and in section 2**

The 11th and 12th reminders were provided for more detailed texts, **mentioning, the blocking questions** and how to proceed

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



# Conclusions



The solution to use reminders with personalized texts has helped to resolve some critical issues in a targeted manner, **obtaining higher response rates**, especially over a shorter fieldwork period

The text of the email object has 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 on the same targets



# General conclusions

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 produced **significant results in terms of increasing response rates reductions in the data collection periods, product and process innovations**

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

**Efficiency gains** can be re-used in further process and product innovation activities, in the **quality** of the outputs and to respond to new **needs** of statistical information. They can also represent the base for statistical **burden** reduction

The experiences reported in the presentation are all examples of innovative tools and services or of process innovation, oriented to process efficiency and quality of the outputs

# General conclusions

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) fragmentation of DC processes
- iii) permanence of overlaps and doubts about "who does what" in the transversal structures

# The main challenges for the future

- 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) Development of a **centralized contact center service** and greater integration between inbound and outbound contact center services
- iv) Development of a **single centralized acquisition Portal** to increase the efficiency of data collection processes from survey units belonging to different sectors
- v) Identify organizational solutions to be applied in order to **reduce the processes fragmentation**, while respecting the principle of specialization and standardization of the activities involved

# The impact of a centralized data collection approach on response rates of economic surveys and data quality: the Istat experience



# UNECE

Giampaola Bellini - Istat ([bellini@istat.it](mailto:bellini@istat.it))

Novella Cecconi - Istat ([ncecconi@istat.it](mailto:ncecconi@istat.it))

Loredana De Gaetano - Istat ([degaetan@istat.it](mailto:degaetan@istat.it))

Francesca Monetti - Istat ([fmonetti@istat.it](mailto:fmonetti@istat.it))

Pasquale Papa - Istat ([papa@istat.it](mailto:papa@istat.it)) (Speaker)

Rita Ranaldi – Istat ([ranaldi@istat.it](mailto:ranaldi@istat.it))

***UNECE Workshop on Statistical data Collection 'Resourceful Data Acquisition'***

***10-12 October 2018, Geneva, Switzerland***