

UNECE 2019 Data collection Workshop
Session 5: Modernisation of Data Collection
Geneve, 14-16 October 2019

Effects of Centralised Data Collection approach on the reduction of the Total Survey Error: experiences in data collection implementation field

Giampaola Bellini (bellini@istat.it)
Silvia Binci – Istat (binci@istat.it)
Loredana De Gaetano (degaetan@istat.it)
Francesca Monetti – Istat (fmonetti@istat.it)
Pasquale Papa – Istat (papa@istat.it) (Speaker)

CENTRALISED DC IN ISTAT: MAIN CHARACTERISTICS

During 2016 Istat launched a **Modernization Program**

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

The new model restricts the role of production structures only to **thematic aspects**, while the “cross” expertises are all assigned to specialized sectors

The "transversalization" of many services fostered **specialization of HR and harmonization of procedures** notably in the field of Data Collection (DC)

Valorization of **administrative sources** and the construction of an integrated system of registers

OBJECTIVES OF PRESENTATION



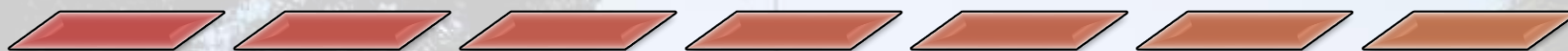
The adoption of a CDC approach involves a **positive effect on the TSE (Total Survey Error, Grooves and Lyberg 2011, Biemer 2010, Lyberg 2019)** for the surveys carried out by the NSI

Other effects concern: control of TSQ (Total Survey Quality), efficiency of DC processes, cost reduction, acceptability by the statistical system and reduction of burden on companies, convergence towards an "advanced" multi-source approach to DC

Empirical evidence confirms the above mentioned effects

TRENDS IN NATIONAL STATISTICAL INSTITUTES

Decreasing number of Human Resources (HR)



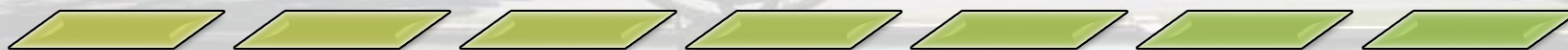
Greater degree of training and specialization of HR



Development of communication and information technologies



Convergence towards more advanced data collection approaches



THE CHALLENGE OF TSE MEASUREMENT

Direct estimates of TSE are difficult to obtain and can be expensive

An operative approach tends to decompose TSE in smaller error component evaluations, in order to facilitate the task

Individuation of specific metrics for process characteristics that are critical to quality

A possibility is applying quality criteria to each error source, and providing ratings by criterion (e.g. according to the scale: poor, fair, good, very good, excellent).

That makes it possible to attribute a specific score to each error source and to obtain a total score

Possibility to reduce TSE for a NSI consists of using specific models and solutions to improve organizations and processes, for clusters or for all surveys

The implementation of CDC can be a solution that goes in this direction

INNOVATIVE PROCEDURES AND TOOLS INTRODUCED IN DATA COLLECTION IMPLEMENTATION: BUSINESS SURVEYS

Harmonized **survey lists** management

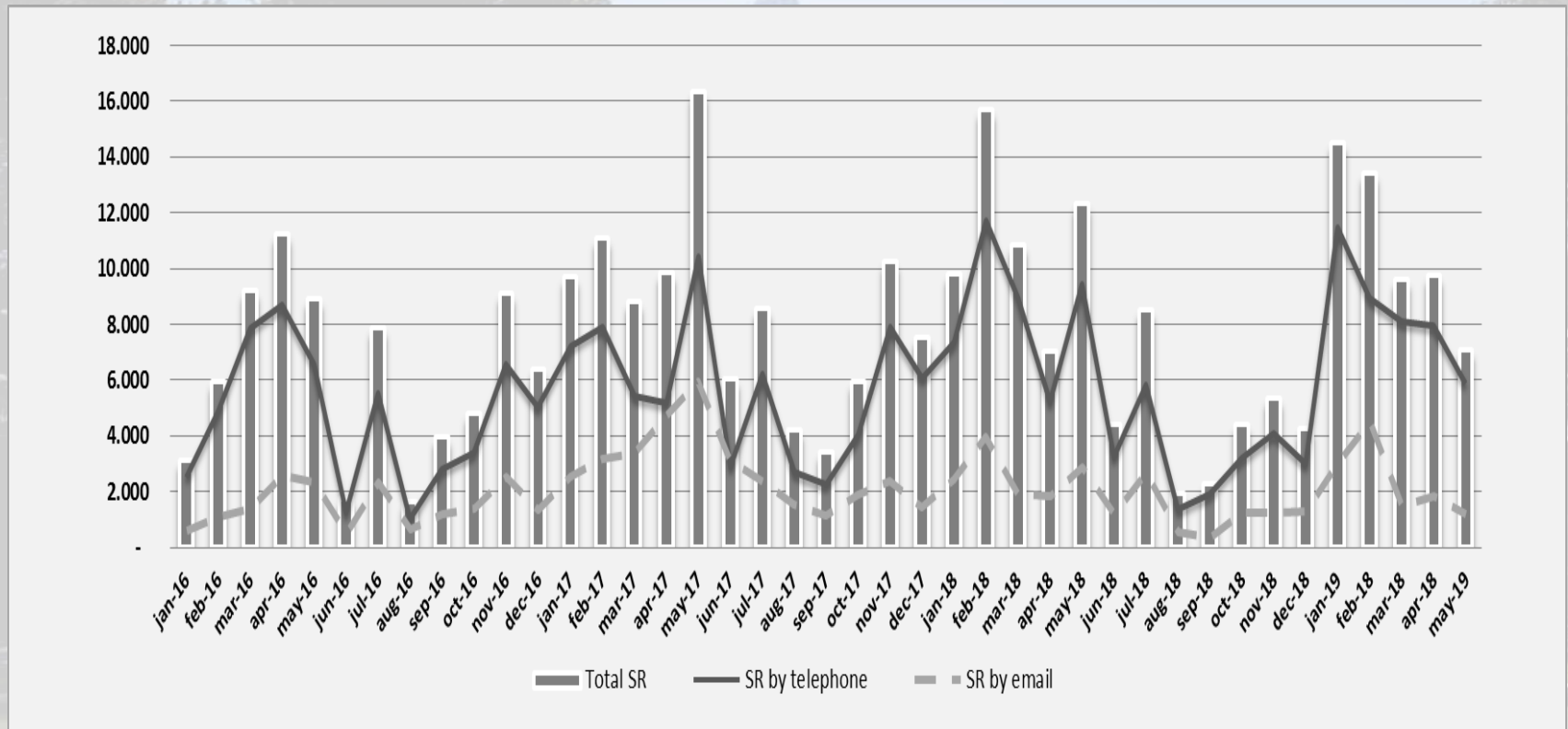
Implementation of **standardized communication and survey management tools and services** (Business Statistical Portal, Centralised Contact Center)

Strict **scheduling** for formal and informal communications (informative letters, reminders,....)

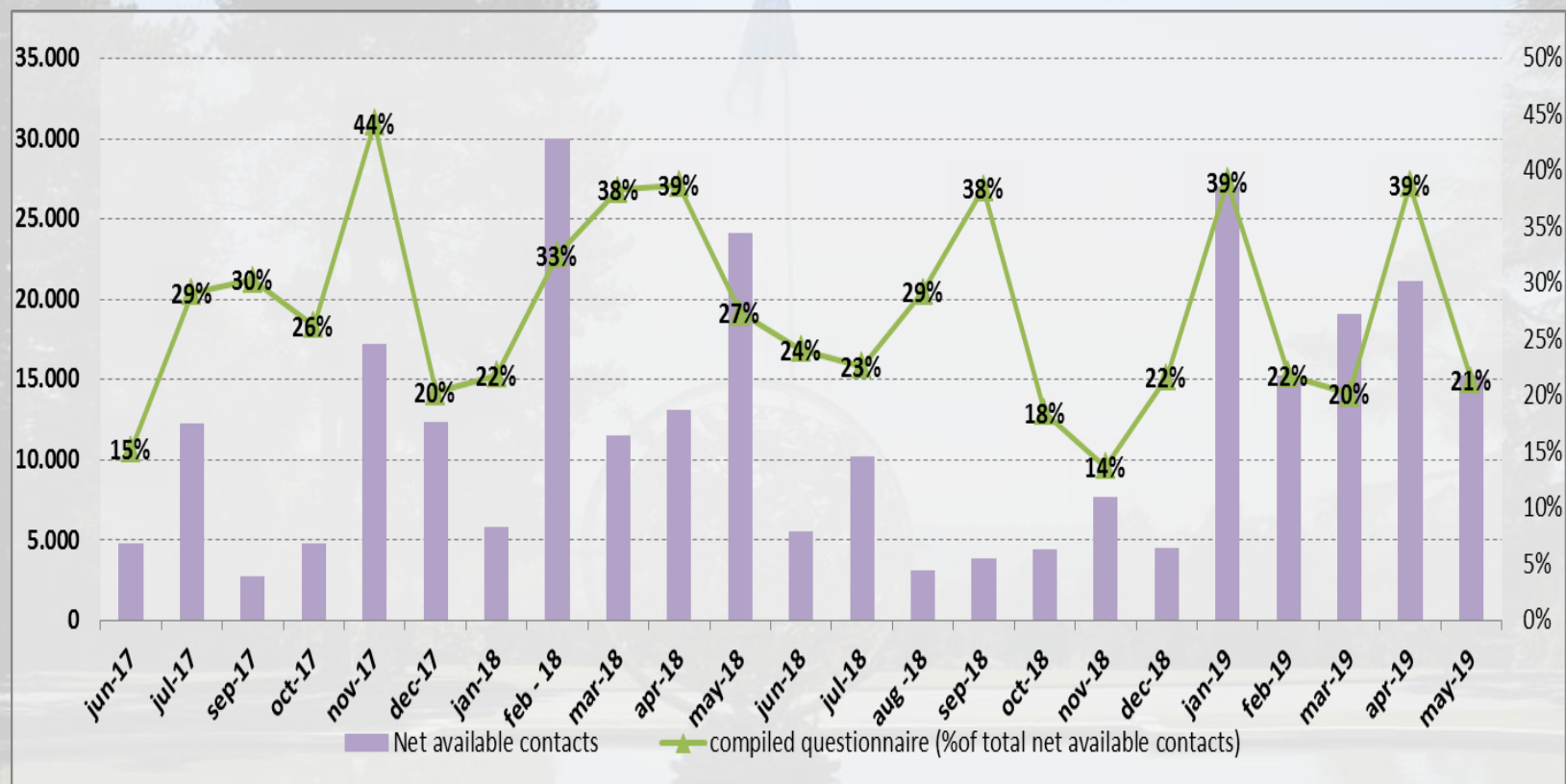
Procedures and tools for **monitoring** the data collection process

Harmonised **penalties** management procedure

INBOUND ASSISTANCE REQUESTS (TOTAL, BY CHANNEL) PER MONTH — PERIOD JAN2016-MAY2019



OUTBOUND SERVICE AND COMPILED QUESTIONNAIRE BY MONTH – PERIOD JUN 2017-MAY 2019

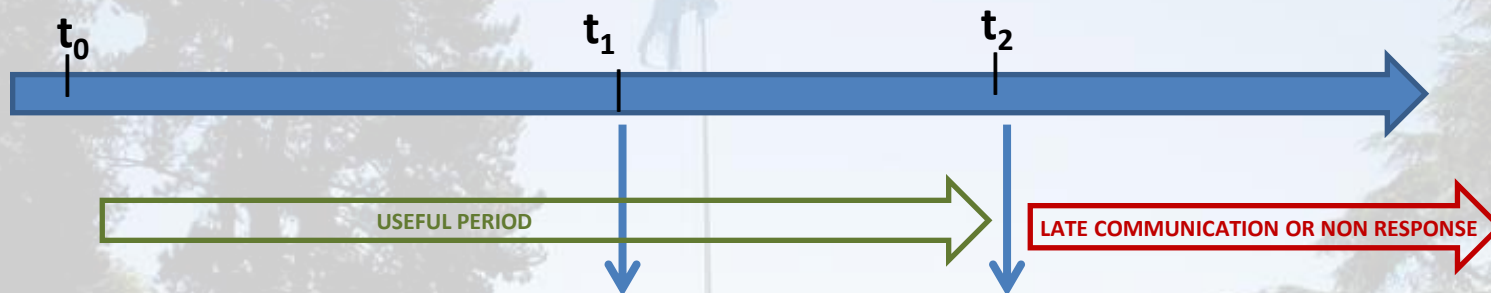


PENALTIES MANAGEMENT CRITERIA IN SHORT-TERM SURVEYS

START OF DATA COLLECTION

PUNCTUAL DUE DATE

USEFUL DATA AS DUE: EXTRA DAYS
RESPECT TO PUNCTUAL DUE DATE

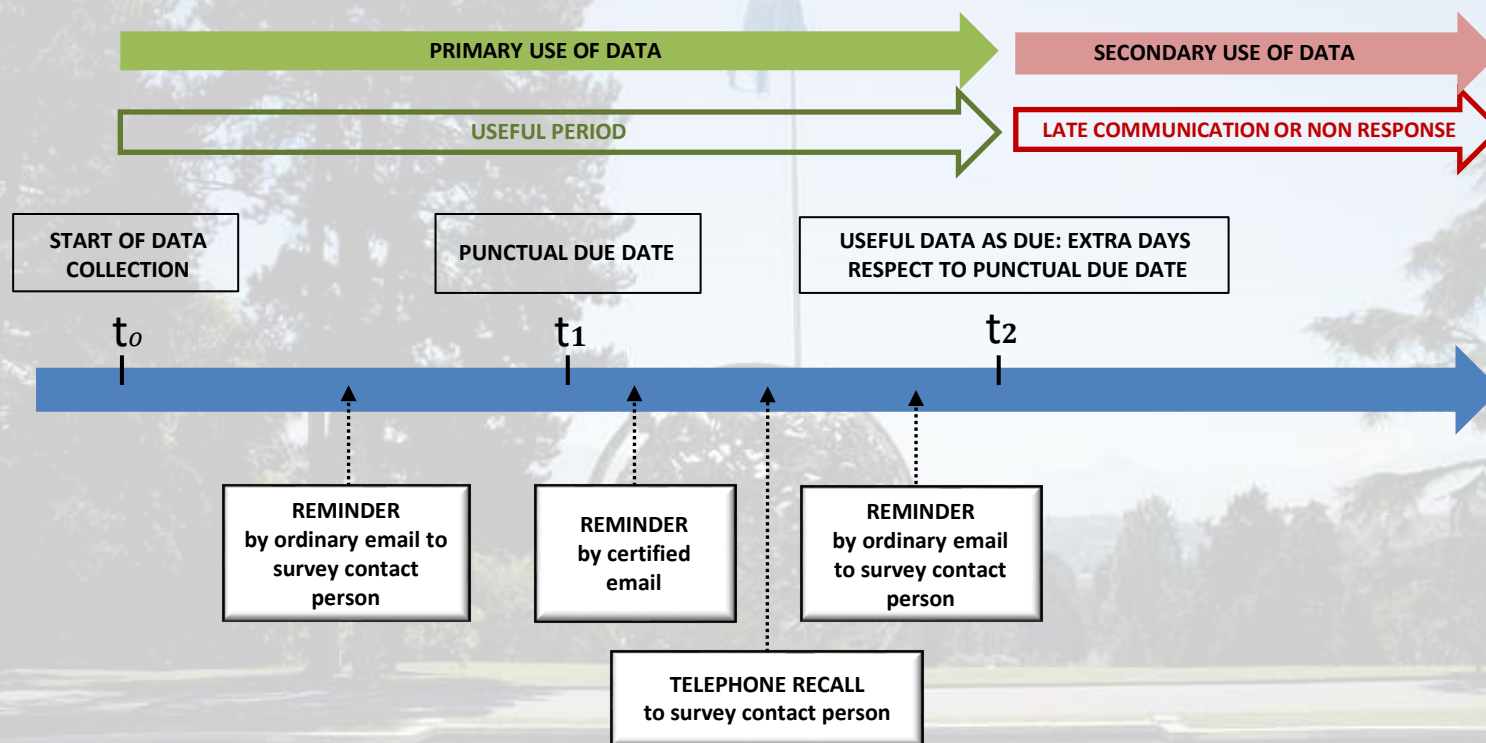


First deadline indicated in the informative letter for each period (month or quarter) for supplying the requested data

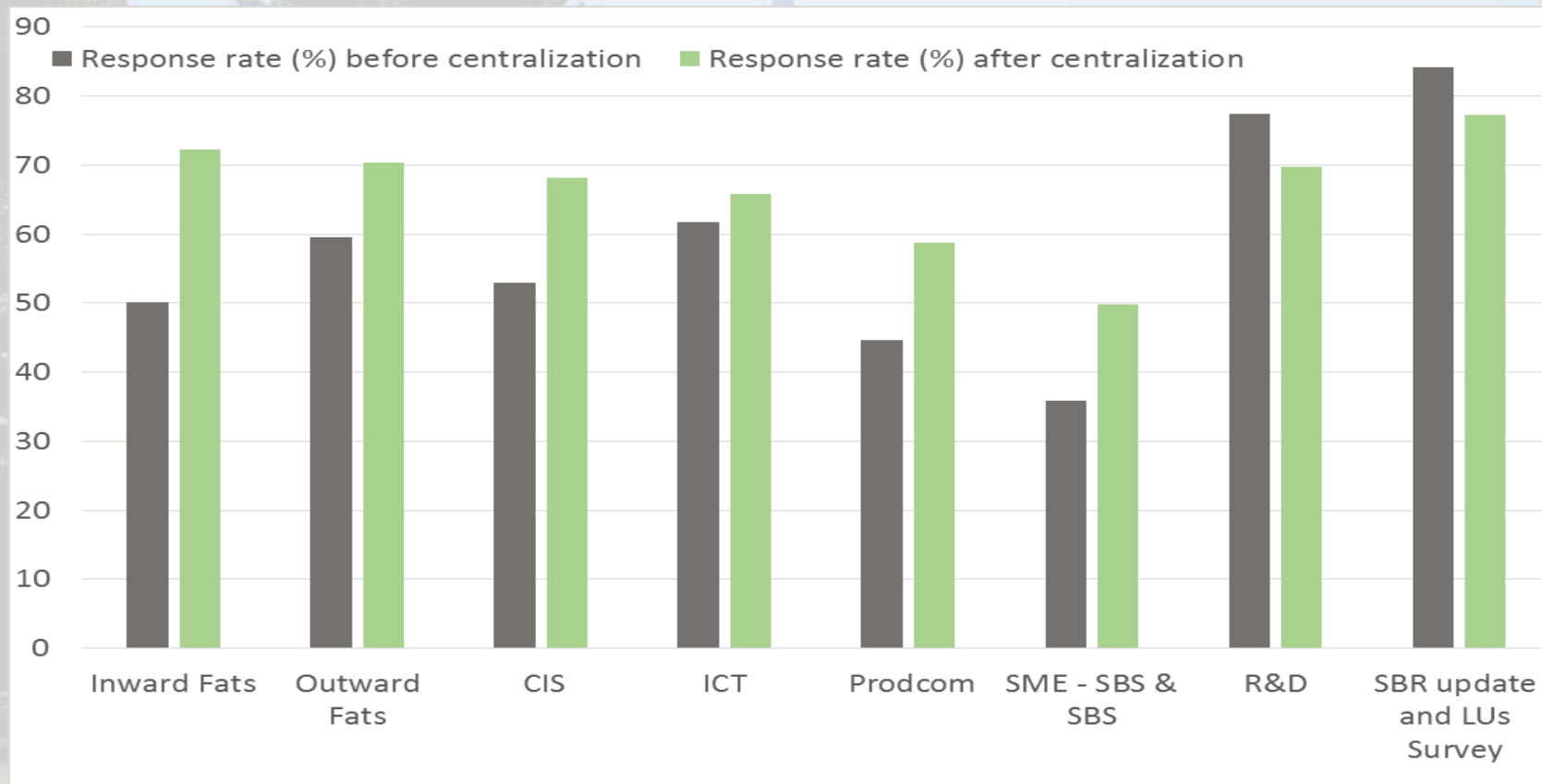
Deadline set after some extra days for sending data, allowed with respect to the specific punctual deadline

Additional criteria: annual cumulative tolerance for supplying data respect to the punctual deadline threshold, lower than the sum of the monthly/quarterly tolerances

PROCESS INNOVATION INTRODUCED: FORMAL AND INFORMAL COMMUNICATION MANAGEMENT CRITERIA IN SHORT-TERM SURVEYS - TIMETABLE



BUSINESS STRUCTURAL SURVEYS, AVERAGE RESPONSE RATES BY SURVEY AND DCI CONDITIONS – YEARS 2014/2015 AND 2018/2019



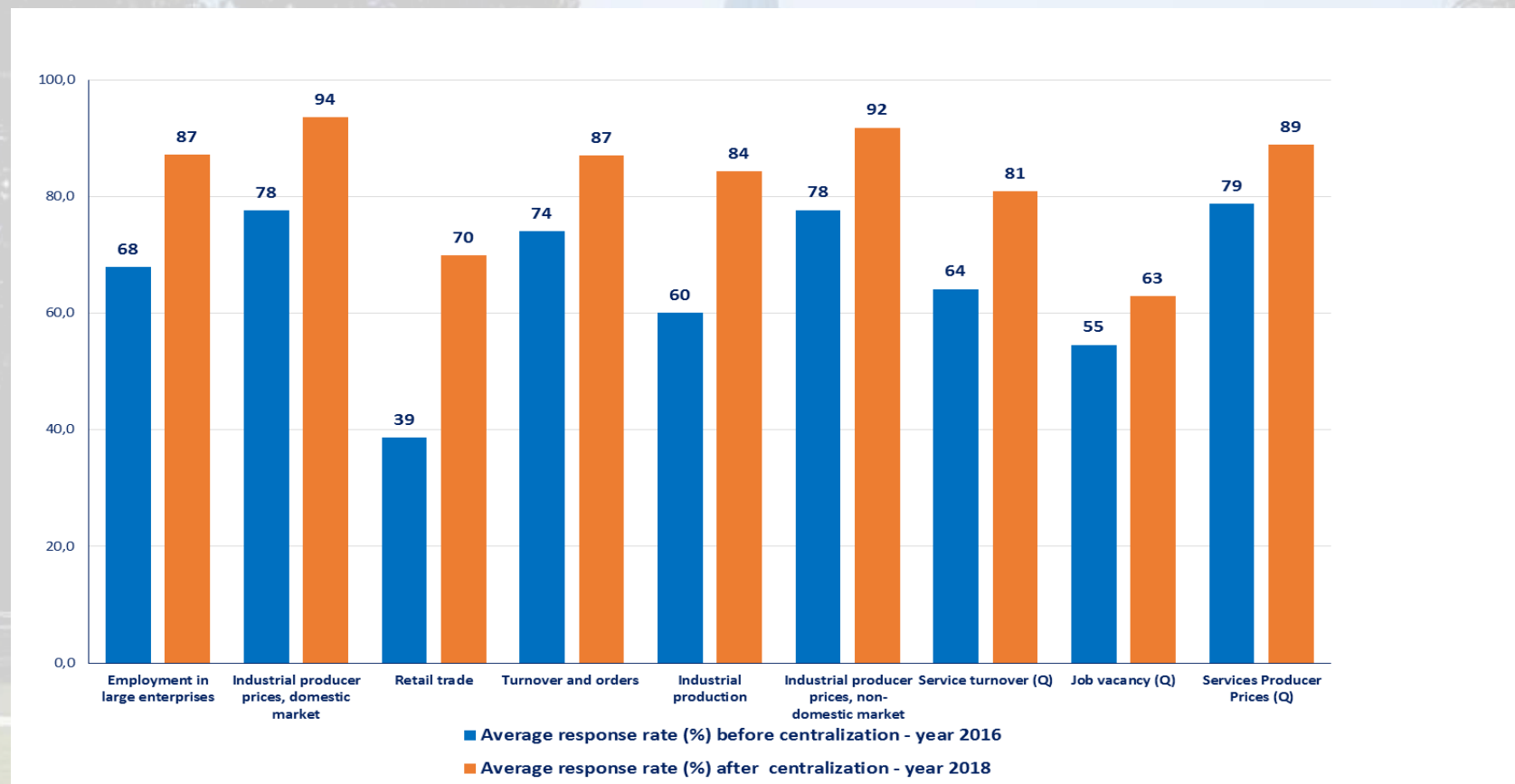
BUSINESS STRUCTURAL SURVEYS, RESPONSE RATES VARIATIONS BY SURVEY AND TYPOLOGY OF ENTERPRISES – YEARS 2014/2015 AND 2018/2019



BUSINESS SHORT-TERM SURVEYS, AVERAGE RESPONSE RATES OF ENTERPRISES VIRTUALLY SUBJECT TO PENALTIES – YEARS 2016 AND 2018

Survey	Year 2016 (%)	Year 2018 (%)	pp variation
Employment in large enterprises	68,7	88,2	19,4
Industrial producer prices, domestic market	74,7	95,8	21,1
Retail trade	62,4	87,5	25,2
Turnover and orders	86,0	94,8	8,8
Industrial production	63,2	91,3	28,1
Industrial producer prices, non-domestic market	75,1	95,4	20,3
Service turnover (Q)	72,8	87,8	15,0
Job vacancy (Q)	77,1	88,6	11,5
Services Producer Prices (Q)	66,7	78,5	11,8

BUSINESS SHORT-TERM SURVEYS, AVERAGE RESPONSE RATES— YEARS 2016 AND 2018



THE EFFECTS OF CDC ON AGRICULTURAL SURVEYS

CDC brought significant **innovations** in data collection processes of agricultural survey processes, all oriented to efficiency and error reduction



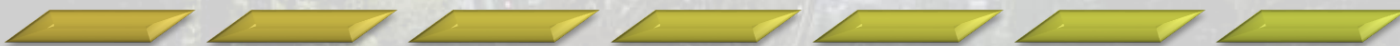
Rationalisation in three main areas:



a) Migration of all the surveys addressed to **industrial companies** in the Business Statistical Portal



b) Application of mixed-mode techniques (CAWI-CATI) to the agricultural surveys **directed at farms**



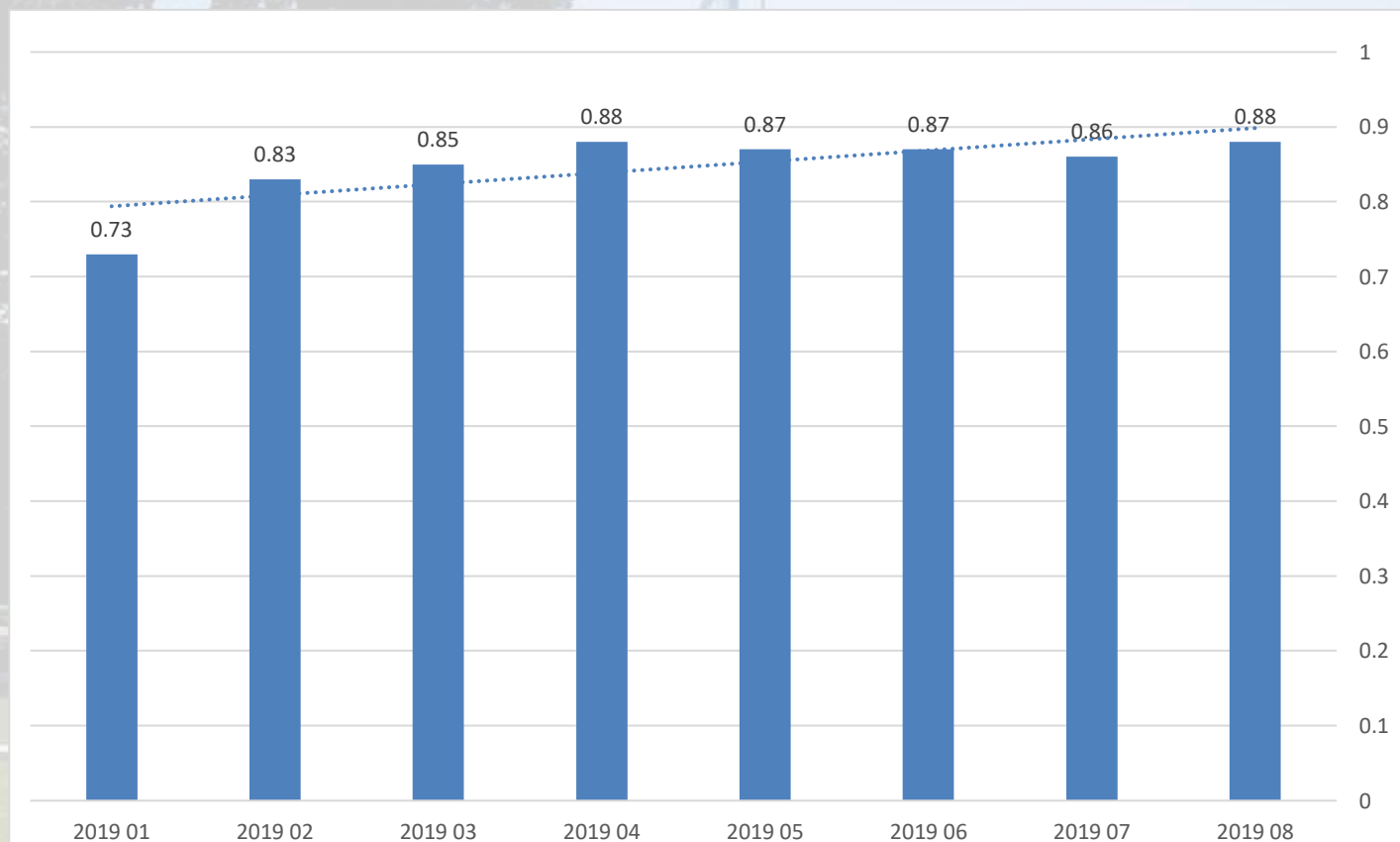
c) Introduction of generalized and integrated criteria for the preparation of **survey lists**, for all surveys



THE EXPECTED EFFECTS

Solution	TSE dimension involved (prevalent)	Quality dimension involved (prevalent)	Statistical system acceptability	Impact on advanced DC
Solution 1. Migration in the BSP of surveys aimed at industrial companies: slaughtering, milk and dairy products	Non-response error	Accuracy, Coherence	Services support to the respondent, costs	Efficiency, Burden
Solution 2. Adoption of mixed-mode techniques for surveys aimed at farms	Non-response error Measurement error	Accuracy timeliness	Possibility for respondents to choose the mode to provide data, costs	User acceptability
Solution 3. Improvements of survey lists quality	Frame error	Accuracy	Efficiency image, costs	Integration with other sources

MIGRATION IN THE BSP OF SURVEYS AIMED AT INDUSTRIAL COMPANIES: MILK AND DAIRY PRODUCTS, 2019 MONTHLY RESPONSE RATES



FIRST CYCLE OF APPLICATION OF MIXED-MODE DATA COLLECTION TECHNIQUE IN AGRICULTURAL SURVEYS – YEAR 2018

SURVEYS	SURVEY TECHNIQUE	NUMBER OF WEEKS OF DATA COLLECTION			CATI-CAWI WEEK BREAKDOWN		DIFF.
		PROVISIONAL	ACTUAL	DIFFERENCE	PROVISIONAL	ACTUAL	
Annual on the intentions of sowing some herbaceous crops	CAWI-CATI	7	9	2	3 CAWI - 4 CATI	2 CAWI - 7 CATI	-1
Annual on the use of plant protection products (Corn)	CAWI-CATI	4	6	2	2 CAWI - 2 CATI	2 CAWI - 4 CATI	0
Annual on the use of plant protection products (Potato)	CAWI-CATI	4	6	2	2 CAWI - 2 CATI	2 CAWI - 4 CATI	0
Annual on milk and dairy products	CAWI-CATI	4	9	5	2 CAWI - 2 CATI	5 CAWI - 4 CATI	3
Livestock consistency on 1-st December	CAWI-CATI	5	7	2	2 CAWI - 3 CATI	2 CAWI - 5 CATI	0
Livestock consistency on 1-st June	CAWI-CATI	5	7	2	2 CAWI - 3 CATI	2 CAWI - 5 CATI	0

CAWI AND CATI RESPONSE RATES FOR THE FIRST EDITION OF CURRENT AGRICULTURAL SURVEYS CARRIED OUT BY A MIXED MODE DATA COLLECTION TECHNIQUE – YEAR 2018

SURVEYS	CAWI Resp. Rates (%)	CATI Resp. Rates (%)
Annual on the intentions of sowing some herbaceous crops	6,2	76,1
Annual on the use of plant protection products (Corn)	8,0	48,4
Annual on the use of plant protection products (Potato)	6,2	76,1
Annual on milk and dairy products	6,1	47,9
Livestock consistency on 1-st December	9,6	61,0
Livestock consistency on 1-st June	8,8	51,8

IMPROVEMENTS OF SURVEY LISTS QUALITY FOR AGRICULTURAL SURVEYS

The process of creating the lists by DC structures included the following steps:

Acquisition of the survey list from the thematic production

Link to the personal data, status of activities, insolvency procedures or corporate events, from “current” registers

Check and revision of territorial references, updated periodically by ISTAT

Formal control, standardization and correction of the certified email addresses, standard emails and telephone numbers, base

integration and updating of the most recent certified email addresses through specific DBs (e.g. Chambers of Commerce)

Integration of the missing ZIP codes for farms without a certified email

CONCLUSIONS

Empirical results confirm positive effects of CDC on TSE in economic and agricultural surveys

Other effects concern: control of TSQ (Total Survey Quality), efficiency, costs, acceptability, burden, convergence towards an "advanced" approach to the DC

The introduction of a CDC setup produced significant results in terms of **increasing response rates** of business surveys

The increasing was in most cases associated to a significant **reduction of data collection periods**.

The basic motivation at the base of these results lies in a set of **product and process innovations**, which were introduced in association with CDC.

The process innovations introduced in the field of **agricultural surveys** represent further examples

The main challenges for the future: to **complete and consolidate CDC** implementation

Promote **Advanced DC convergence**

Implement TSE **evaluation criteria**