Streamlining data collection monitoring activities through the use of control charts: an application to the EU-SILC survey
Outline

- Aim of the work
- Typical control chart
- EUSILC survey
- Defining indicators: some issues
- Procedure to monitor the interviewers’ work
  - indicators
  - charts
- Procedure to monitor the coding of the occupation
  - indicators
  - charts
- Conclusions
Aim of the work

Paradata
Response data

Data collection process of the Italian EUSILC survey

Performance indicators
Control charts

System to monitor fieldwork activities
A typical control chart

Strength: managers are guided in making targeted interventions, without spending time exploring false alarms.
EUSILC survey

- EUSILC (European Union Statistics on Income and Living conditions) is the **EU reference source for comparative statistics on income distribution and social exclusion**

- The survey is carried in Italy by Istat

- **Reference population**: all households (HHs) and their members residing in Italy

- **Sample design**: two-stage scheme (municipalities-households)

- The **annual sample** comprises:
  - HHs of WAVE 1 from population registers of the municipalities selected at the 1\textsuperscript{st} stage
  - HHs of WAVE 2, 3, or 4 interviewed for the 2\textsuperscript{nd}, 3\textsuperscript{rd}, or 4\textsuperscript{th} time

- Data are collected **by an external company** on behalf of Istat
EUSILC survey – Data collection

Technique

- **CAPI**: HHs with **no telephone number**
- **CATI**: HHs whose members are **all foreign citizens**
- remaining HHs

HHs from WAVE 1 are asked for their telephone number

For WAVE 1: longer duration of **interviews**

- **N. OF INTERVIEWERS**
  - CAPI: 34
  - CATI: 158

- **N. OF INTERVIEWS (%)**
  - WAVE 1: CAPI (21), CATI (14)
  - WAVE 2, 3, 4: CAPI (45), CATI (45)
Defining indicators: some issues

Constraints dictated by available data and interview protocol

Effects of undesired source of variability in the process data

Refusals
Breakoffs
Time spent in coding

Data collection mode
Wave
HH size

Information that cannot be used to define indicators

Stratification of indicators and charts
The monitoring strategy

Indicators, based on paradata and response data
- completed interviews
- eligible units
- attempts to complete the interview
- duration of interview

Control chart for each indicator

Step 1. screening chart for all the interviewers working in the reference week

Step 2. in-depth charts to monitor each interviewer flagged at Step 1, over the whole fieldwork period

Tabular reports listing:
- only the flagged interviewers
- any other statistics or information useful to decide on the type of intervention
Indicators and charts to monitor the interviewers’ work

- **Response rate**
  n. of completed interviews / n. of eligible HHs
  by interviewer, data collection mode (CAPI; CATI),
  wave (WAVE=1; WAVE>1)
  \( p \)-chart (chart for proportions)

- **Productivity rate**
  n. of completed interviews / n. of contact attempts
  by interviewer, data collection mode (CAPI; CATI)
  \( p \)-chart (chart for proportions)

- **Adjusted duration of interview**
  time spent to complete an interview (adjusted for HH size)
  by interviewer, data collection mode (CAPI; CATI),
  wave (WAVE=1; WAVE>1)
  \( \bar{X} \) and \( s \)-charts (charts for continuous variables)
Screening chart for *Adjusted duration of interview – 5th week* CATI data – WAVE>1

- **4 flagged interviewers**

- Mean referred to the reference week

![Chart showing mean and standard deviation for interview duration across interviewers, with indicators for flagged interviewers.](chart.png)
In-depth charts for *Adjusted duration of interview* – 5th week
Interviewers 3002 and 3042 – CATI data – WAVE>1

**INTERVIEWER 3002**

- **DATE**
  - 18NOV, 21NOV, 24NOV, 27NOV, 30NOV, 03DEC, 06DEC, 09DEC, 12DEC, 15DEC, 18DEC, 21DEC

- **long interviews**
  - Mean referred to the interviewer

- **DATE**
  - 18NOV, 21NOV, 24NOV, 27NOV, 30NOV, 03DEC, 06DEC, 09DEC, 12DEC, 15DEC, 18DEC, 21DEC

**INTERVIEWER 3042**

- **DATE**
  - 18NOV, 21NOV, 24NOV, 27NOV, 30NOV, 03DEC, 06DEC, 09DEC, 12DEC, 15DEC, 18DEC, 21DEC

- **very short interviews**
  - Mean referred to the interviewer

**Interventions**: de-briefing with interviewer 3002 and strict monitoring of interviewer 3042
Aim: detect any potential improper interviewing behavior that might lead to assign a **wrong code to the occupation** declared by the respondent

**Monitoring the coding of occupation**

- **OPEN-END QUESTION on occupation**
  - if coding takes place **during the interview**
  - if coding takes place **once the interview is over**

- **ONLINE SYSTEM**
  - looking for a match of the description within a database

- **NOTES FIELD**
Indicators and charts to monitor the coding of occupation

- Proportion of classification codes to assess possible concentrations of codes by interviewer, data collection mode (CAPI; CATI), occupational group (1; 2; 3; 4; 5; 6; 7; 8; 9)

  analysis-of-means chart (for proportions)

- Proportion of confirmed codes to assess the tendency to confirm previously assigned codes (only for WAVE>1) by interviewer, data collection mode (CAPI; CATI)

  analysis-of-means chart (for proportions)
Chart of the *Proportion of codes* in the 6\textsuperscript{th} occupational group

CAPI data

6 flagged interviewers

INTERVIEWER ID
Chart of the *Proportion of confirmed codes*

**CAPI data**

not enough commitment in the coding task?

By integrating the results from the other charts, it is possible to understand whether the out-of-control event is linked to a fraudulent attitude.

**above-average efforts**

- lack of understanding of how the coding should be performed
- difficulties in using the assisted coding software
Next steps and lessons learned

The monitoring system is going to be used for the 2021 edition of EUSILC.

Close cooperation among methodologists, fieldwork supervisors, and survey managers to understand how to improve the system.

Similar approach applied to other surveys that need to monitor the interviewers’ work.
thanks

Manuela MURGIA, Alessandra NUCCITELLI, Francesca ROSSETTI

murgia@istat.it, nuccitel@istat, frrosset@istat.it