

METHODOLOGICAL NOTES

Target Population

All the enterprises with 10 or more employees operative in the Eu core sectors (NACE Rev. 2 sections B, C, D, E, H, K and in divisions 46, 58, 61, 62, 63 and 71), plus some additional activities chosen in order to ensure data comparability over time (NACE Rev. 2 sections F and I, and divisions 45, 47, 59, 68, 72, 77).

Statistical Units

The enterprise, as defined in Council Regulation No 696/1993 on statistical units and defined in the national statistical business register.

Reference Period

This round of survey covers the period 2006-2008 (with reference year: 2008).

Type of Survey

A combination of sample and census survey. The census refers to the enterprises of with 250 + employees. For the rest of population, a stratified random sample has been built. The strata are defined by size (in terms of number of employees), by sector (NACE 2-digits codes) and by region (at NUTS2 level).

Data Collection

Data collection of the Italian CIS 2008 has been carried out by means of a web survey and a questionnaire self-compiled by enterprises and completed online. Some reminders were sent to non-respondent enterprises to minimise the unit non response.

Business Register

Observed enterprises are extracted from the most up-to-date edition of Istat's Business Register, called Asia (Archivio Statistico delle Imprese Attive).

Survey questionnaire

The Italian Innovation Survey uses the European harmonised survey questionnaire (CIS2008) and is the same for all NACE sectors.

Data Processing

An updating of a SAS software appositely designed by Eurostat for Cis has been used for data processing. The micro and the aggregated data were checked and corrected for inconsistencies. In case of missing values, before carrying out automatic imputation, a large use of administrative, historical (CIS 2006 survey) and R&D surveys has been done. To correct for item non-response, different techniques were used for metric variables and ordinal-nominal variables. For metric variables, a weighted mean of each metric variable, by NACE and size class, is calculated and applied as a ratio to the enterprises with the missing values, within the stratum concerned. Imputation of ordinal and nominal variables, done after the metric estimation, made use of nearest-neighbour hot decking technique.

Estimation Process

This activity referred to the overall respondents sample, obtained by merging the respondents samples come from the three different statistical sources. The survey results are weighted in order to adjust for the sampling design and for unit non-response to produce valid results for the target population. Calibration estimators methodology (Deville and Särndal, 1992), was used for the estimation process. The software used is GENESEES, a generalised software implemented in SAS language by ISTAT researchers.

Final Respondent Sample and Response Rate

Final respondent sample is composed of 19,904 units, which are representative of 208,638 enterprises of the CIS2006 target population. The response rate is 52.1 percent.