



**BANCA D'ITALIA**  
EUROSISTEMA

***Survey of Industrial  
and Service Firms***

**BIRD**  
**Bank of Italy Remote  
access to micro Data**

**Sample design  
and methods used**  
Version 3.2



November 2014

## Contents

1.	Cross-references .....	3
2.	The yearly survey on industry and services.....	3
2.1	General aspects .....	3
2.2	The sample design .....	3
2.3	Sample and reference population .....	4
2.4	The observation phase and the questionnaire.....	5
2.5	Imputation of missing data.....	6
2.6	Weighting coefficients.....	6
2.7	The survey results published in the Supplements to the Economic Bulletin ...	7
3.	The Business Outlook Survey of Industrial and Service Firms.....	7

## Sample design and methods used<sup>1</sup>

### 1. Cross-references

This document contains a general description of the methods used in the Bank of Italy's Survey of Industrial and Service Firms. For more information on the design of the sample and the methods used each year, see the **Methodological Notes** section of each Sample Surveys edition of the [Supplements](#) to the Statistical Bulletin devoted to the survey and available on the Bank's website.

### 2. The yearly survey on industry and services

#### 2.1 General aspects

The Survey of Industrial and Service Firms continuously collects data on: firms' individual details, employment, investment (both actual and planned), turnover, capacity utilization, debt and trade receivables. The survey also collects information on other matters, which vary from year to year and are of particular interest for economic research (e.g. corporate strategies and governance, firm size, physical, human and organizational capital, and electric power).

Prior to 1998 the survey only covered firms in the manufacturing sector with 50 or more employees. In 1999 the field of observation was extended to include all manufacturing firms by enlarging the sample with the addition of firms in the following ATECO categories: subsection DF (Coke manufacturing and refineries), Section C (Mineral extraction) and Section E (Electrical energy, gas and water). In 2001 the survey was extended, with a shorter questionnaire, to firms with 20 to 49 employees. In 2002 a similar survey was introduced covering non-financial private service firms with 20 or more employees, excluding the following from market services: credit institutions, insurance companies, public services and other social and personal services).

#### 2.2 The sample design

The population of the survey is divided into strata and from each one a certain number of firms is extracted on a random basis. These firms make up the sample to be observed (one-stage stratified sample design). The strata are combinations of branch of activity, size class (in terms of number of employees) and the regional location of the firm's head office.

The number of workers for the classification of firms by size was measured at the end of the year until the 2003 survey; subsequently reference has been made to the average number of workers during the year (for further details, see part devoted to the [archives](#)).

---

<sup>1</sup> Prepared by Stefania Coscarella, Leandro D'Aurizio, Caterina Di Benedetto and Raffaele Tartaglia Polcini.



The number of firms to be observed in each stratum is determined not in relation to their number in the population (the proportional-sample method) but so as to obtain reliable estimates of the dynamics of investment, employment and turnover. This required the observance of more larger firms than on the basis of their presence in the population (the optimum-allocation-to-strata method, which makes it possible to minimize the standard error of the sample means by over-sampling the strata with the highest variance). For special research objectives, the over-sampling also concerned firms with their head office in the South of Italy, of which there are therefore more than on the basis of their presence in the Italian population of firms.

### 2.3 Sample and reference population

The sample is a panel: the firms observed in the previous edition of the survey are always contacted if they are still part of the population covered by the survey. Those that are no longer willing to cooperate are replaced by others homogeneous in terms of economic activity and size class. Firms that have participated at least once in the survey are considered to have dropped out of the sample if they no longer belong to the relevant universe as a consequence of their going out of business, liquidation, bankruptcy or merger or because the firm ceases to engage in the economic activities or to be in the size classes covered by the survey. In addition, firms drop out of the sample, sometimes temporarily, that are not willing to participate in the survey.

The number of firms in the sample by sector and size can be found in the document containing the [description](#) of the archives. Table 2 shows the number of firms in the different populations.



Table 1

SIZE OF THE POPULATION OF FIRMS, 1984-2011<sup>(1)</sup>

Survey year	Total industry excluding construction		of which: manufacturing		Services		Total industry and services	
	20-49	≥ 50	20-49	≥ 50	20-49	≥ 50	20-49	≥ 50
1984	-	-	-	12,601	-	-	-	-
1985	-	-	-	12,457	-	-	-	-
1986	-	-	-	12,313	-	-	-	-
1987	-	-	-	11,917	-	-	-	-
1988	-	-	-	12,025	-	-	-	-
1989	-	-	-	11,883	-	-	-	-
1990	-	-	-	11,739	-	-	-	-
1991	-	-	-	12,041	-	-	-	-
1992	-	-	-	11,658	-	-	-	-
1993	-	-	-	11,185	-	-	-	-
1994	-	-	-	11,037	-	-	-	-
1995	-	-	-	10,880	-	-	-	-
1996	-	-	-	11,411	-	-	-	-
1997	-	-	-	11,792	-	-	-	-
1998	-	-	-	11,609	-	-	-	-
1999	-	11,712	-	11,502	-	-	-	-
2000	-	12,029	-	11,798	-	-	-	-
2001	27,516	12,629	27,075	12,389	-	-	-	-
2002	26,235	12,095	25,755	11,828	17,509	7,956	43,744	20,051
2003	26,173	12,254	25,713	11,978	18,339	8,338	44,512	20,592
2004	25,685	11,954	25,230	11,677	18,915	8,650	44,600	20,604
2006	24.681	11.858	24.209	11.574	19.925	9.400	44.606	21.258
2007	25.201	11.727	24.711	11.432	21.375	9.776	46.576	21.503
2008	25.189	11.925	24.176	11.168	23.110	10.390	48.299	22.315
2009	23.232	11.427	22.243	10.652	22.351	10.263	45.583	21.690
2010	22.142	10.915	21.147	10.119	22.312	10.137	44.454	22.052
2011 <sup>(2)</sup>	21.809	10.912	20.848	10.119	22.361	10.341	44.170	21.253

(1) Source: Istat (2009). The numbers shown refer only to the population of firms covered by the survey. - (2) Most recent year for the data concerning the population of firms (Istat, Statistical Archive of Active Firms, 2011).

## 2.4 The field and the questionnaire

The data are collected annually between January and April with reference to the previous year (the reference year) in the course of interviews conducted by the Bank of Italy's branches.

Separate questionnaires are prepared for industry excluding construction and for services, for firms with 20 to 49 workers and for firms with 50 and over. The general purpose of this distinction is to avoid having to ask questions that entail a

greater respondent burden for small firms and to allow space, if necessary, to focus on separate topics for industry and services and for large and small firms.

The questionnaires are composed of a fixed part and a variable part. The fixed part contains general information on the firm and its structure, as well as annual data on investment, employment, turnover, result for the year, capacity utilization and debt. The variable part covers different themes each year, which are the subject of detailed cyclical or structural analysis. The questionnaires for service firms are partly different from those used for industrial firms.

The Bank's economists define the content and structure of the questionnaires shortly before they are administered. New questions are tested by the Bank's branches by means of pilot questionnaires designed to assess whether they are easy to understand and whether the information is effectively available from the firms.

In the case of employment, investment and turnover, information is requested for three periods: the year just ended (preliminary results), the previous year (final results) and the following year (forecasts).

Starting from the 2010 survey, data are loaded through a web application or an interactive pdf form compiled directly by the firm.

## 2.5 Imputation of missing data

The percentage of imputed data is usually small. A higher rate of non-response tends to occur with questions involving forecasts, particularly of investment.

Each of the imputed variables is associated with a flag variable that has an empty value if the corresponding value was actually provided by the firm and 1 if it was imputed. The list of imputed variables is available in the document containing the [description](#) of the archives.

The methods used for the imputation of missing data are based on cell averages or regression estimators and are intended to improve the aggregate estimates more than estimating the response the individual firms would have given. Using the flags, it is always possible to identify the imputed values and exclude or replace them with the results obtained with a different model.

## 2.6 Weighting coefficients

Each firm in the sample is assigned a weighting coefficient (**PESO**), which indicates the number of firms in the population that the firm in question represents.<sup>2</sup> Some respondent firms are assigned unit weight. This occurs for self-representative firms when the observation is census-based within the stratum. In a small number of cases a unit weight is also assigned to some outlier firms, which do not consider themselves representative of the other firms in the stratum with regard to the performance of some fundamental variables.

In theory the sum of the weights of the sample design should be equal to the number of firms of the population in the stratum. It was nonetheless decided not to impose this requirement to the complete combinations of sector, economic activity and geographical area, both because some combinations may be without sample units and because such a weighting system could lead to excessive variability of the

---

<sup>2</sup> The archives also contain another variable, PESORISC, containing PESO, multiplied by a factor so that the sum of all the PESORISCs corresponds to the numerosness of the sample. This variable is likely to be useful when using programs that consider the sum of the weights in order to calculate the degrees of freedom.



estimates for some domains of study. Then some stratification cells were identified, made up of class size and sector of economic activity combinations and some post-stratification cells, represented by the geographical area in which the firm is located. In the two sets of cells, independently, the sum of the weights of each cell returns the numerosness of the population of interest it contains. Reference should be made to the [description](#) of the archives for a detailed definition of these variables.

At each survey the weights are recalculated on the basis of the distribution of the population of interest at the latest date available.

The weighting coefficients can be found in the archive (the **PESO** and **PESORISC** variables). To avoid distortions in the estimates, it is recommended that they be used in every calculation. **Since 2007 PESOADD is also available: it reproduces the population's distribution for employees. These weights are particularly suitable for categorical variables, as they take into account the different sizes of the firms.**

## 2.7 The survey results published in the Supplements to the Economic Bulletin

The main results of the survey are reported in the Supplements to the Economic Bulletin series – Sample Surveys: Survey of Industrial and Service Firms; in the Bank of Italy's Annual Report and its Appendix. All the documents can be consulted on the Bank's website.

The differences that may sometimes be found between the published results and those recalculated on the basis of the micro data are due to a variety of factors, such as:

1. the database used for calculations includes any data revised after they were originally published;
2. at the time of publication the weighting coefficients are calculated on the basis of population data that are not updated, since they are not available for the most recent years. When the population data become available, the weights are recalculated;
3. the correction of individual data and changes in the weighting coefficients may give rise to revisions in the calculation of imputed values as well, since they depend on all or some of the other values;
4. since variables expressed at constant prices refer to the most recent year, they are subject to change each year (even though the changes are often very small).

## 3. The Business Outlook Survey of Industrial and Service Firms

The target population of the outlook survey, carried out between September and October has evolved in the same way as the yearly survey. The survey design, weighting scheme and sample units are the same. The two samples may differ either because of the panel attrition or because some units leave the reference population.

The survey asks qualitative data on the firm-level main variables of interest. A pilot study tests the new questions.

The main results of the survey are reported in the Supplements to the Economic Bulletin series - Sample Surveys: Business Outlook Survey of Industrial and Service Firms and in the Bank of Italy's Economic Bulletin. All the documents can be consulted on the Bank's website.

