A systematic and standardised burden measurement system for surveys on businesses

Annamaria D'urzo, Angela Golino, Stefania Macchia, Alessandra Nurra, Carmela Pascucci, Mariagrazia Rinaldi, Fabiana Rocci, Carla Schiattone, Giorgia Simeoni [1]

Abstract

Following what emerged from the second round of the peer reviews on the ES Code of Practice carried out in 2015, Istat decided to propose a systematic and standardised burden measurement system, to be implemented for surveys on business through the web-based data collection.
This paper describes the study conducted to identify a representative set of indicators and the methodology to elaborate them. The results of a test run on several businesses surveys are also described and some reflections about possible strategic actions to contain burden are presented.

Keywords: response burden, business surveys.

1 Annamaria D'Urzo (adurzo@istat.it); Angela Golino (golino@istat.it); Stefania Macchia (macchia@istat.it); Alessandra Nurra (nurra@istat.it); Carmela Pascucci (pascucci@istat.it); Mariagrazia Rinaldi (mrmarina@istat.it); Fabiana Rocci (rocci@istat.it); Carla Schiattone (schiatto@istat.it); Giorgia Simeoni (simeoni@istat.it), Italian National Institute of Statistics – Istat.

The views and opinions expressed are those of the authors and do not necessarily reflect the official policy or position of the Italian National Institute of Statistics – Istat.

1. Introduction[2]

This paper describes a study carried out by an internal Working Group[3] of the Italian National Institute of Statistics – Istat aimed at defining and implementing a system of indicators for the measurement of respondent burden for surveys on businesses.

One of the recommendations resulting from the second round of peer reviews on the implementation of the ES Code of Practice[4], underlined that Istat was already collecting information on response burden in some business surveys but not in a systematic and standardised way. Thus an improvement action for the definition and implementation of a systematic and standardised burden measurement system was formulated by Istat. The Working Group, set up at the end of 2015, had the aim of complying with such improvement action.

The purpose of this report is to describe the set of indicators on response burden that Istat has designed for surveys on businesses in a strategic perspective, so as to release them in a systematic and standardised way through a generalised software procedure. The results obtained on a first set of three surveys are presented, and some reflections about the possible strategic actions to contain burden are mentioned.

The following Section 2 specifies some context information related to the modernisation of the production process that Istat has been udergoing since 2016 (Istat, 2016). Section 3 reports some general concepts and definitions on response burden measurements and the general decisions endorsed by the Working Group on what kind of burden will be measured and how. Section 4 presents the set of indicators that the Working Group has defined in order to measure the response burden by survey and Section 5 shows the preliminary results obtained on three surveys, on which the standardised process was tested according to their different features (short term / structural survey, short/long

2 Although the article is the result of a joint work, the single parts are authored as follows: Sections 1, 2, 3, and 4 by Stefania Macchia; Paragraph 5.1 by A. Nurra; Paragraph 5.2 by A. Golino and F. Rocci; Paragraph 5.3 by C. Pascucci and F. Rocci; Section 6 by M. Rinaldi and C. Schiattone; Section 7 by A. D'Urzo; Section 8 by G. Simeoni.

3 Coordinators: S. Macchia, N. R. Fazio. Members: S. Cuomo, A. Golino, R. Mazzucco, M. Murgia, A. Nuccitelli, A. Nurra, A. Nunnari, P. Papa, C. Pascucci, S. Pietropaoli, M. Rinaldi, F. Rocci, C. Schiattone, G. Simeoni.

4 For further information on the peer reviews on the ES Code of Practice please refer to http://ec.europa.eu/eurostat/web/quality/peer-reviews.

questionnaires). Section 6 describes the Automatic Procedure for Burden Indicators by Survey which implements the standardised process, while Section 7 is devoted to the percistency indicators. Finally, some preliminary suggestions to contain burden are identified.

2. The transition to the new model for business surveys

It is worth mentioning that since 2016 Istat, in the context of the newly established Business Architecture Activity Model (Istat, 2016), has been reunifying all the scattered resources and processes pertaining to the data collection phase under a new dedicated Directorate. Thus extending the principles underpinning the reorganisation which the business survey system has undergone in recent years. The main goal of that reorganisation was to abandon the so-called 'stovepipe' model, in which each 'stovepipe' identifies a specific field of statistics and its corresponding production system, to implement a new model in which the production of statistics is no longer expected to operate through independent processes, but rather as a single, consistent and integrated pool of non-redundant building blocks (enterprise-centred model).

Cornerstone of this new system is the Istat Business Statistical Portal (N. Fazio, M. Murgia, A. Nunnari, 2013) an integrated system for the management of data collection processes, which is, at the same time, an attempt to streamline the organisation and production processes of business surveys as a whole.

The Portal acts as a single entry point for web-based data collection from enterprises, according to a 'business-centric' perspective. It provides new integrated functions supporting respondents in several areas: survey unit management and updating, data collection process management, direct and proxy compilation of electronic questionnaires. The environment also provides a software tool to develop/design electronic questionnaires: GX (Generalised Italian (Data) Collection System XML), an in-house product using XML, to represent the main survey's structure, *i.e.* survey metadata, survey variables, questionnaire structure, check plan and skipping rules.

To fully achieve its goals, this new architecture relies heavily on fast-tracking the semantic and syntactic harmonisation of survey questionnaires, *i.e.* both in terms of concepts and design, with the perspective of improving quality and containing the respondent burden.

Up to now, surveys questionnaires were implemented with different, often dedicated/*ad hoc* software systems, so the migration of all of them in the Business Portal data collection system is gradually undergoing.

This emerging context of integration and common IT solutions is encouraging the definition of standards and recommended practices for questionnaire design and for all functions belonging to the data collection phase, including the sets of indicators aimed at monitoring and enhancing quality of the data collection processes.

3. The burden to be measured

The literature on the measurement of statistical burden on businesses is quite wide and a review of it can be found in the Memobust handbook (2014, Eurostat). Building on this, in order to identify indicators to be produced, it was needed to provide some details to better substantiate what was intended to measure. It is possible to narrow down the factors that contribute to create burden to two main classes:

- actual/objective factors, mainly due to time spent to provide responses;
- subjective factors, connected with what is 'perceived' as burden by respondents.

The first kind of factors can be measured in terms of time, but also of costs. For this purpose, a model has been evaluated, the so called SCM - Standard Cost Model (2004, Eurostat). On balance, this was not considered as a viable, cost-effective option, as SCM requires analysts to collect a broader set of information than just time spent completing the questionnaire. (2012J.Jones).

On the other hand, the concept of time spent to provide responses needs to be specified as well, making clear what it is intended to include. In particular, it should be decided whether considering all the following activities:

- time to understand what is being required;
- time to retrieve data to be provided;
- time to complete the questionnaire;
- time to respond to re-contacts, *e.g.* during the data editing phase.

Istat decided that **the burden each survey places on respondent businesses will be measured only in terms of: *i*) time to retrieve information to be provided; *ii*) time to completely fill in the questionnaire**. It was considered that tracking time spent on the remaining activities in the list would have been too complex and time consuming.

The identification of the 'perceived burden' is even more complex as it could depend on a number of aspects, such as the survey design, the respondent's characteristics and other external factors, that might require gathering additional data directly from the respondents.

In designing the set of indicators to measure burden, Istat sought a solution imposing two constraints:

- do not cause further burden to respondents in order to collect information on burden;
- limit internal investments to set up IT procedure to estimate burden in different surveys.

The first constraint compelled to define a set of indicators by relying on a strategy of maximum exploitation of available sources and minimal request of information from respondents.

Following the second constraint, it has been decided that **data concerning burden will be collected and processed for surveys already residing in the Portal environment or as soon as they migrate to it, so as to implement and set up a generalised software procedure**.

Besides what already mentioned, the burden can be considered from two different perspectives:

- burden by survey (BBS), *i.e.* the burden the single survey places on the involved businesses;
- burden by business (BBB), *i.e.* the total amount of burden generated by all the surveys a business is involved in.

Istat decided to measure burden from both points of view, expressing BBS in terms of total time spent to fulfill the requested survey task, and BBB in terms of persistency.

The choice not to produce burden indicators for business surveys that still are not migrated to the new environment is offset by means of the persistency indicators, which will allow to have an overall view of burden imposed on all businesses surveyed by Istat.

The last aspect which has to be stressed before describing the proposed indicators is that they are aimed at representing the evolution of the phenomenon of burden more than a precise estimate of a statistical entity. This is because the indicators are thought for all kinds of surveys (both short term and structural ones), that involve different types of sample designs and different treatments of changes that occur to the respondent units during the

same year. Hence, the chosen criteria for selecting the set of businesses to be considered in each period to measure burden will not take into account a series of events (for instance events depending on the business demography, as well as late responses of businesses providing data after the given deadline) which, on the contrary, are relevant for the survey's results. This is to assure the comparability of the several surveys burden indicators to be evaluated year by year.

4. Burden indicators by survey (BBS)

As already mentioned, the burden that each survey lays on respondent businesses will be measured in terms of time to complete the questionnaire and time to retrieve information needed. In particular, burden will be considered for all the respondent businesses (businesses who submitted the filled-in questionnaire). In details:

- **time to fill in the questionnaire** will be quantified processing the paradata[5] the data collection system automatically records, indeed it was decided to exploit sources of information automatically generated by the software system (paradata) and minimise the request of data to businesses. As known, paradata are a rich informative source for monitoring the data collection process (M.P. Couper, 1998), as they store automatically all actions performed by respondents while navigating the data collection environment;
- **time to retrieve information to provide data** and other information will be asked to respondents in an *ad hoc* section to be added to each surveys questionnaire as they migrate to the Business Portal system (the '*Burden section*' is shown in Figure 4.1).

There is a very important difference between the two sources used: paradata are exhaustive, which means they are available on all businesses who submitted the filled-in questionnaire, while data coming from the *Burden section* of the questionnaires are partial as answering was not compulsory. On the other hand, it has been confirmed (see Section 5) that the set of businesses which fills in this section is not particularly characterised against the complete set of respondent businesses.

5 The paradata of a data set or survey are data about the process by which the data were collected. Example paradata topics about a survey include the times of day interviews were conducted, how long filling in the questionnaire took, whether questionnaire was completed with a single access to the system or in different times, etc. Thus there are paradata about each 'observation' in the survey. The analysis of these data are useful to asses the costs and management of a survey, so as to identify possible improvements of the questionnaire design.

Figure 4.1 - The *Burden section*

1. Report how many people were involved in providing information:

☐☐

2. Indicate the time you spent to retrieve the information necessary to fill in the questionnaire, selecting one of the following time classes

Please check only one answer

1	2	3	4	5	6	7
Up to 15 minutes	*From 15 minutes to half an hour*	*More than half an hour and up to 1 hour*	*More than 1 hour and up to 3 hours*	*More than 3 hours and up to 10 hours*	*More than 10 hours and up to 30 hours*	*More than 30 hours*
☐	☐	☐	☐	☐	☐	☐

The upper classes won't be presented for very short questionnaires

3. Indicate whether and which difficulties you encountered in filling in the questionnaire:

More than an answer is possible

a) *Technical and IT problems* ☐

b) *Difficulties in navigating through the questionnaire sections* ☐

c) *Too many questions* ☐

d) *Insufficiently clear questions* ☐

e) *Unclear or not exhaustive answer options* ☐

Not pertinent for very short questionnaires

f) *Not enough information supporting the understanding of questions* ☐

g) *Excessive distance between information requested and information available in the business* ☐

h) ***No difficulties encountered*** ☐

4. Please provide any comments and/or suggestions regarding the filling in of the questionnaire:

Source: Istat - Study on the burden measurement system

4.1 Time spent to fill in the questionnaire

As already mentioned, paradata will be used to estimate the time spent to fill in the questionnaire. The strength of using this source is that it is absolutely objective and precise. It allows to quantify the actual and net time each respondent dwells on each of the questionnaire sections, even if the compilation takes place in different sessions. As a matter of fact, with paradata the start and end times of each compilation session are recorded (date, hour, minute, second). By processing these timestamps, it is possible to measure the duration of each session and, in case the questionnaire has been filled during different sessions, the summed duration of all sessions represents the total actual time spent to fill in the questionnaire. As we see, this calculation is based on primary data and absolutely objective, while compilation times elicited directly from respondents could be affected by subjective evaluation and prone to perception bias.

By processing paradata it is also possible to get additional and very interesting information, for instance: the gross amount of time the respondent spent connected to the web questionnaire (from the start date and time of the first session to the end date and time of the last session), whether the respondent did it in one single session or in several sessions, whether he/she made it in a single day or across several days.

The indicator of burden will be processed for respondent businesses, namely those that submitted a completed questionnaire. This does not imply any further consideration for structural surveys (SBS), while a clarification has to be made on short term surveys (STS): only businesses who provided data for all the periods of the reference year are considered as respondents (12 months for monthly surveys, 4 quarters for quarterly surveys).

So the indicators, for SBS and STS surveys respectively, are as follows:

SBS surveys

(1) $\overline{\boldsymbol{ACT}}$ (Average Compilation Time) = $\sum_{i=1}^{n}(CT_i)/n$

where

CT_i = compilation time, expressed in minutes, of the *'i'* respondent business

n = total number of respondent businesses

STS surveys

(2) $\overline{\boldsymbol{ACT}} = \frac{\sum_{j=1}^{m}(\sum_{i=1}^{n}(CT_{ij})/n)}{m}$

where

CT_{ij} = compilation time, expressed in minutes, of the *'i'* respondent business in the *'j'* period

n = total number of businesses respondent for **all** the periods of the reference year (constant value for all the periods)

m = total number of periods (periods of the reference year for which the respondents provided data, which means 12 months for monthly surveys, 4 quarters for quarterly surveys)

As it can be seen from the formulas, this indicator is a simple mean, but it provides survey managers with a starting point for further analysis, for instance to highlight whether there are significant differences depending on the businesses' characteristics (in terms of dimension or turnover or economic activity sector).

4.2 Time spent retrieving required information and number of people involved in the task

The answer given by respondents to the second question of the *Burden section* will be used to estimate time spent to retrieve information. This datum is surely subjective and, if the respondent perceives the task of providing data for statistical survey as a burden, it could be overestimated. On the other hand, the experience in conducting business survey clearly shows that within the 'Response process' (Edwards and Cantor, 1991; Sudman *et al.*, 2000; Willimack and Nichols, 2001) the step of 'retrieval information' is particularly heavy for businesses because it often implies finding and analysing data from corporate databases and/or paper filing systems. It follows that this information must absolutely be taken into account in estimating the respondent burden.

As it can be seen in Figure 4.1, the question in the *Burden section* (*time you spent to retrieve information necessary to fill in the questionnaire*) is structured in classes (class 1: up to 15 minutes; class 2: from 15 minutes to half an hour,etc.). An approximation of the time spent on average will be

elaborated, taking the central value of each class as reference time for the final calculus. The last class will not be considered for this calculus because defining the central value would need not available information. In addition, this class is selected by a very low percentage of respondents (see Section 5).

In synthesis, it will be expressed through the following indicator:

(3) $\overline{\boldsymbol{ART}}$ (Average Retrieval Time) = $\sum_{i=1}^{n}(RT_i)/n$

where

RT_i = retrieval time of the *'i'* respondent business

n = total number of businesses which provided this information in the *Burden section*

Another aspect was considered relevant to estimate burden: the number of people involved in providing information. It is worth considering that the fact of involving more than one person may have several explanations: providing the requested data requires different professional skills, so that different experts within the business must be contacted or, particularly for STS surveys, the person who fills the questionnaire could not be the same in all data collection periods during the year. Anyway, collecting such an information can be helpful for the survey manager to conduct further analysis.

The indicator will be expressed through the following formula:

(4) $\overline{\boldsymbol{PI}}$ (Persons Involved) = $\sum_{i=1}^{n}(PI_i)/n$

Where

PI_i = number of persons involved in providing information of the *'i'* respondent business

n = total number of businesses which answered to this question in the *Burden section*

As already said, the *Burden section* was added to questionnaires as surveys migrated to the new GX system. It must be specified that for STS surveys, it was decided to request the *Burden section* only in the last period (last month or last quarter) of the collection year, specifying that the information requested was the average time spent to retrieve information to fulfill the task in the different periods of the year.

4.3 Possible burden caused by difficulties encountered in providing information

With the third question of the *Burden section* (*Indicate whether and which difficulties you encountered in filling in the questionnaire*) some aspects related to other possible causes of burden pertaining to survey design are investigated: questionnaire length, questionnaire design (in terms of clarity of questions, support information, question options), usability of the electronic questionnaire (ease of navigation and functions supported), IT problems, etc. The set of response options of this question is surely not exhaustive and respondent's attitude towards collaboration is not being investigated, but this would have required a wider set of questions which in turn would have caused more burden on respondents. Certainly this question does not allow to single out the specific problem/difficulty (if the option '*unclear and not exhaustive questions*' is selected, it is not known which question it is referred to), but it can be viewed as evidence of some problems occurring, especially in the context of surveys recently migrated to the new IT system, or every time the questionnaire has undergone deep changes.

The indicator will be simply expressed through a table showing the frequency distribution of businesses per number of problems encountered. This information allows to monitor the evolution of the phenomenon, with the aim of reducing the percentage of respondents encountering higher numbers of difficulties.

Table 4.1 – Respondents

Number of difficulties declared by the respondent	*Number of respondent businesses*	*% of respondent businesses*
0		
1		
2		
3		

Source: Istat - Study on the burden measurement system

4.4 Summary report on burden by survey (BBS)

For each survey the defined set of indicators will be systematically calculated.

Besides the values of the indicators, some further information on the survey, on the questionnaire and on the actions undertaken to reduce the respondents' burden can be of interest for an overall assessment of the response burden by survey.

First of all, the raw[6] number of respondents (respondent businesses) to the survey provides an idea of the survey size and the 'incidence' of the burden on the businesses population.

Secondly the raw number of respondent units is reported because the burden is measured at business level, but a business, if articulated in different units, would provide information for all of them, so knowing the number of units pertaining to a single business helps to explain the obtained value of the burden indicator.

Finally it can be interesting to know what actions have been already implemented to contain the burden, in particular if the sample was selected limiting the overlapping with other surveys.

The following figure summarises the information to be provided/collected for each survey.

6 'Raw' meaning that businesses responding after the deadline, for example, are not considered.

Figure 4.2 - Summary report

Summary report on response Burden

(1) $\overline{ACT}$ = Average Compilation Time in minutes

(2) $\overline{ART}$ = Average Retrieval Time

(3) $\overline{PI}$ = Average number of Persons Involved in providing information

(4) Frequency distribution of businesses per number of problems encountered

Raw Number of respondent businesses _____ Raw Number of respondent units _____

Use of sampling coordination function: _____

Source: Istat - Study on the burden measurement system

In addition, another information can be analysed to explain burden, even if it will not be included in the summary report because mostly relevant for internal experts: the level of complexity of the questionnaire. Istat has defined a simple indicator to measure in a standard and comparable way the complexity of its questionnaires. The indicator of complexity (CI)[7] takes into account the number and type of questions included in the questionnaire, the difference between the minimum and the maximum number of answers that should be provided and the presence of routing rules. These different factors are then summarised on a qualitative scale (easy, medium, difficult). The value of the CI can be used as a support to compare the indicators on burden across different surveys: *e.g.* if two surveys have the same level of CI and very different values for the burden indicators the situation calls for further investigation. The difference in topics between the two surveys can, for example, justify the variation, but it is also possible that the survey with higher level of burden would benefit from some improvements in its design to reduce it.

7 The concepts considered for this index are: number of questions, scores of questions and paths to fill in the questionnaire. Different scores are assigned according to the characteristics of the questions (open-ended, closed-ended, multiple choice, etc.). The different paths of the questionnaires are analysed in order to identify the shortest and the longest path.
Questions scores corresponding to the shortest and to the longest path are calculated.
The Index is the arithmetic average between these two scores.

5. Results of burden indicators in ICT, LES and Turnover surveys

It is necessary to mention that those reported in this document are the first results obtained through the described methodology, which could be further assessed after a deeper analysis. In particular more insights and reflections are foreseen for short terms surveys to validate the methodologic choices defined for them through the analysis of the first year results.

5.1 First results of burden indicators on ICT survey

Istat conducts survey on the use of information and communication technologies (ICT) on the basis of EC Regulation on Statistics on Information society involving active enterprises in industry and services with 10 or more employees. Since the year 2014 this survey has been managed through the new data collection system integrated into the Business Statistical Portal. Concurrently, major/significant innovations were introduced, regarding the overall design of the questionnaire, the way inconsistencies are highlighted in the electronic form when rules violations are triggered and the tools for monitoring the survey progress. Moreover, it was decided to add a new section at the end of the questionnaire to measure *respondent burden* in terms of classes of time necessary to find information (*retrieval time*) and about number of persons involved in providing the requested information.

In the rest of the present paragraph, results of the analysis on both the *Burden section* and the compilation time are presented.

Response rate of ICT survey in 2015 was about 61%. Table 5.1.1 reports the percentage distribution of respondents who only answered the survey and respondents who gave an answer also to the *Burden section*, showing a very good data representativeness.

Table 5.1.1 - Respondents ICT2015: *Burden section* by size class (absolute value and percentages)(a)

Size class	Total	Respondent to ICT and not to burden	Respondent both to ICT and burden
		percentages	
10-19	9,146	0.70	99.30
20-49	3,675	0.65	99.35
50-99	1,982	0.50	99.50
100-249	2,109	0.95	99.05
250+	2,509	0.88	99.12
Total	19,421	0.72	99.28

Source: Istat – ICT survey: study on the burden measurement system
(a) Enterprises are considered as respondents to the *Burden section* if they answered at least one question between retrieval time and persons involved.

Table 5.1.2 shows the retrieval of information time classes by size classes of enterprises. As we see, there is a clear-cut concentration of enterprises of all sizes in the first three classes of retrieval time and, moreover, it is evident that a larger amount of time was needed by companies with at least 50 persons employed (*i.e.* more than 60% of large enterprises needs a time between 30 minutes and 3 hours to find information). The extra time required by larger enterprises is justified by the fact that a more complex organisation implies more intensive use of ICT, which results in a longer path of the questionnaire (different paths are due to the responses to the filter questions).

Table 5.1.2 - Respondents ICT2015: time to retrieve information by size class (percentages)(a)

Size class	Time to retrieve information					
	Up to 30 minutes	More than half an hour and up to 1 hour	More than 1 hour and up to 3 hours	More than 3 hours and up to 10 hours	More than 10 hours and up to 30 hours	More than 30 hours
10-19	41.11	41.00	15.77	1.58	0.29	0.25
20-49	36.64	41.95	18.36	2.53	0.28	0.25
50-99	32.49	39.86	22.01	4.73	0.51	0.41
100-249	25.11	38.36	28.52	6.05	0.91	1.06
250+	15.41	35.94	34.33	10.85	1.82	1.65
Total	34.33	40.12	20.68	3.76	0.57	0.54

Source: Istat – ICT survey: study on the burden measurement system
(a) Missing: 210.

Similarly, with increasing firm size also the number of people involved in compiling increases due to the greater variety of skills required to answer questions of different topics investigated by ICT survey (Table 5.1.3). Compared with an average of about 1.7 people needed by enterprises with fewer than 20 persons employed, filling in the model takes on average 3.5 persons to the larger enterprises with almost 250 persons employed. The more general *Average number of Persons Involved in providing information* ($\overline{PI}$) is 2.17 persons per enterprise.

Table 5.1.3 - Respondents ICT2015: number of persons Involved in providing information by size class (percentages)(a)

Size class	1(b)	2	3	4	5	6+	Average number of Persons Involved in providing information
10-19	47.40	37.64	11.66	2.25	0.65	0.41	1.73
20-49	38.79	37.30	16.65	4.44	1.94	0.87	1.98
50-99	30.46	36.46	19.04	7.15	4.49	2.40	2.30
100-249	20.15	34.04	24.67	9.98	6.20	4.96	2.71
250+	11.40	27.86	24.84	12.19	10.43	13.28	3.53
Total	36.51	35.82	16.44	5.26	3.13	2.83	2.17

Source: Istat – ICT survey: study on the burden measurement system
(a) Missing: 686.
(b) In this category enterprises answered 'zero persons involved' are included.

Using central value of each time classes to retrieve information (even if the classes do not have the same width), it is possible to calculate the *Average Retrieval Time* ($\overline{ART}$) as reported in Table 5.1.4. Standard deviation shows high dispersion of the data around the mean.

Table 5.1.4 - Average Retrieval Time by size class (a)

Size class	Missing	N	Mean *(minutes)*	Standard deviation
10-19	97	9,026	53.29	83.19
20-49	42	3,624	59.73	88.62
50-99	15	1,959	74.07	114.71
100-249	26	2,061	90.74	139.07
250+	30	2,438	125.86	182.58
Total	210	19,108	69.94	114.20

Source: Istat – ICT survey: study on the burden measurement system
(a) Central values of first 5 time classes shown in Table 5.1.2 have been used.

Paradata are also useful in measuring respondent burden in terms of time spent to complete the questionnaire, so that it is possible to calculate the *Average Compilation Time* ($\overline{ACT}$) (2015, Masselli *et al.*; 2014, Nuccitelli *et al.*). In Table 5.1.5 standard deviation shows high dispersion of data and an average time of about 47 minutes increasing passing from small enterprises (42 min) to large ones (65 min). Data show a certain direct relation between size classes and net time needed to complete the questionnaire.

Table 5.1.5 - Average Compilation Time in minutes by size class

Size class	N	Mean (*minutes*)	Standard deviation
10-19	8,962	41.95	50.75
20-49	3,614	43.30	56.64
50-99	1,952	46.47	64.40
100-249	2,064	54.11	75.37
250+	2,465	64.68	95.44
Total	19,057	46.92	64.01

Source: Istat – ICT survey: study on the burden measurement system

In Tables 5.1.6 and 5.1.7 are presented results from a question included in the *Burden section* only during the first wave of the survey hosted by the new data collection system (2014). The question was about difficulties encountered by respondents in filling in the questionnaire. It was asked about *IT difficulties* caused by the new system, *conceptual difficulties* related to lack of clarity or of supporting information, *ease of data availability* and on length of the questionnaire. About 1 out of 2 respondents was experiencing no difficulties (49%). Later on, we tried to solve the problem of conceptual difficulties improving language used in the questions, using more effective FAQ and instruction for filling questionnaire uploaded in web site of survey and adding more tooltips in the web model.

Table 5.1.6 - Respondents to ICT2014 survey and to difficulties question, by size class (absolute value and percentages)

Size class	Total	Respondent to ICT and not to difficulties	Respondent both to ICT and difficulties
		Percentages	
10-19	1,831	5.74	94.26
20-49	9,021	6.52	93.48
50-99	3,329	7.08	92.92
100-249	2,308	9.56	90.44
250+	2,061	10.49	89.51
Total	18,550	7.00	93.00

Source: Istat - ICT survey: study on the burden measurement system

As it is shown in Table 5.1.7, 1 enterprise out of 3 indicates not more than 2 difficulties encountered in compiling the questionnaire.

Table 5.1.7 - Respondents ICT2014 by number of difficulties encountered in compiling the questionnaire, by size class (percentages)

Size class	Number of difficulties				
	0/missing	1	2	3	4+
10-19	50.13	21.44	14.75	7.87	5.81
20-49	56.32	20.04	13.25	6.58	3.81
50-99	60.80	19.21	11.89	5.77	2.33
100-249	65.05	18.79	10.54	3.39	2.24
250+	68.93	17.76	8.45	3.16	1.69
Total	56.24	20.22	12.96	6.38	4.20

Source: Istat - ICT survey: study on the burden measurement system

In general, for all sizes of enterprise the highest difficulty was related to the length of the questionnaire (21%), also conceptual difficulties were big obstacles to fill in the questionnaire (26,5%); finally, very few indicated difficulties encountered in the new data collection tool (9%) (Table 5.1.8).

Table 5.1.8 - Respondents ICT2014 by type of difficulties encountered in compiling the questionnaire, by size class (percentages)

Size class	Technical and IT problems	Difficulties in navigating through the questionnaire sections	Too many questions	Insufficiently clear questions	Unclear or not exhaustive answer options	Not enough information supporting the understanding of questions	Excessive distance between information requested and information available in the business	No difficulties encountered
10-19	9.33	1.52	23.74	22.30	11.66	17.83	14.54	44.39
20-49	8.44	1.20	20.46	17.87	9.28	13.97	11.93	49.80
50-99	7.57	0.92	18.78	13.29	8.01	10.29	11.60	53.71
100-249	5.57	1.47	17.86	10.32	7.97	7.65	9.34	55.49
250+	3.94	0.95	15.86	8.80	7.50	5.20	9.36	58.45
Total	7.94	1.32	21.04	17.64	9.95	13.72	12.59	49.24

Source: Istat - ICT survey: study on the burden measurement system

Figure 5.1 reports the summary report on response burden for ICT survey.

Figure 5.1 - Summary report for ICT survey year 2015

(1) $\overline{ACT}$ = Average Compilation Time in minutes: 46.92

(2) $\overline{ART}$ = Average Retrieval Time : 69.94

(3*) $\overline{PI}$ = Average number of Persons Involved in providing information:2.17

(4*) Frequency distribution of enterprises per number of problems encountered

Number of difficulties declared by the respondent	*0/missing*	*1*	*2*	*3*	*4*	*5*	*6*	*7*
% of respondent enterprises	56.24	20.22	12.96	6.38	2.73	1.04	0.23	0.20
Number of respondent enterprises	10,432	3,751	2,405	1,183	506	192	43	38

- Number of respondent businesses used[7]: *19,421*

- Coordination function used in selection of the sample: *negative coordination*

**Note that for indicators 3 and 4 data from ICT year 2014 are used*

Source: Istat - ICT survey: study on the burden measurement system

5.2 First results of burden indicators on LES survey

Since the '80s of the last century, Istat has conducted the monthly survey on employment, working hours, wages and labour costs in large enterprises (LES). Starting from 2012 a section about job vacancy has been included in the questionnaire at the end of every quarter (in March, June, September and December). This survey on the *large enterprises* contributes along with other two quarterly surveys (one concerns job vacancy and hours worked – Vela, the other one concerns Employment Remuneration and Social Security Contributions – Oros) to determine indicators on the input and labour costs in all enterprises with employees. The values obtained through the integration of these three different statistical surveys are sent to Eurostat in compliance with the following European regulations:

1. STS-term statistics (no. 1165/98) and subsequent amendments and additions to the number of persons employed, hours worked and gross and salary wages.
2. Labour Cost Index (n. 450/2003).
3. Regulation on job vacancies of the European Parliament and of the Council (no. 453/2008) and on the Commission's implementing regulations (no. 1062/2008 and no. 19/2009).

The LES survey refers to enterprises with more than an average of 500 employees at the base year 2010[8]. These are about 1,250 and are monitored heavily to minimise the non-response rate, that has been registered to be about 17% for preliminary estimates and only 3.5%[9] for the final estimates released in April 2016.

It is worth reporting that, in order to renew the panel for the base year 2015, a set of 380 new enterprises started to be surveyed. As these enterprises do not contribute to the current results, they are monitored more lightly: at April 2016 the non-response rate on this set of units has been registered to be about 40%.

8 This base refers to 2015 year of survey, but it changes periodically.

9 Monthly reminders (by e-mail and fax) and intensive follow-ups by phone are addressed to not responding LES units. In 2015 once a year a warning with penalty (registered letter with return receipt) was sent to firms that had not answered to LES for two or more months.

In 2015 this survey started to be managed with the new data collection system integrated in the Business Statistical Portal and the questionnaire was developed in GX system.

As ICT survey, with the migration to the Business Statistical Portal some significant innovations were introduced. They regarded:

- number of questions: some added and others removed from previously 2014 edition;
- overall design of the questionnaire;
- the way inconsistencies are highlighted in the web form when rules violations are triggered.

Furthermore, the already mentioned *Burden section* was added in the questionnaire (section K). It's important to underline that answering to this section was not compulsory and no alerts appeared on the screen if the section was not filled.

In the following, results of the analysis on both the compilation time and the *Burden section* are presented.

As already said, response rate of LES survey in 2015 is different depending on whether the enterprises belong to panel of the base 2010 (1,250 enterprises) or to the renewed panel for new base 2015 (not already considered for published data, 380 units).

In addition, as only the subset of businesses who provided data for 12 months is considered to calculate the burden indicators (see Paragraph 4.1), the response rate for 2015 is 88% (1,422 enterprises as respondents over 1,630 enterprises in total).

The *Average Compilation Time* ($\overline{ACT}$) is almost 43 minutes, more or less the same time as standard deviation.

Table 5.2.1 - Average Compilation Time in minutes

	N	Mean (*minutes*)	Standard deviation
Total	1,422	42.6	43.7

Source: Istat - LES survey: study on the burden measurement system

The respondents to section K of questionnaire of December were 997, while the non respondent units were 362. In this case the number of units that were supposed to respond were 1,765: in fact this section is referred to each respondent unit (KAU) (Table 5.2.2).

Table 5.2.2 - Respondents LES 2015: *Burden section*

	Number of units	Respondent to LES December and not to burden	Respondent both to LES December and burden
Total	1,765	362	997

Source: Istat - LES survey: study on the burden measurement system

Table 5.2.3 shows the time to retrieve information by size classes. As we see, nearly half of respondents needed more than 1 hour and up to 3 hours, while the 31% needed more than half an hour and up to 1 hour to fill the questionnaire.

Table 5.2.3 - Respondents LES2015: time to retrieve information by size class (percentages)(a)

	Time to retrieve information				
	Up to 30 minutes	More than half an hour and up to 1 hour	More than 1 hour and up to 3 hours	More than 3 hours and up to 10 hours	More than 10 hours and up to 30 hours
Total	12.1	31.0	44.4	12.0	0.4

Source: Istat - LES survey: study on the burden measurement system
(a) Missing: 367.

The *Average number of Persons Involved in providing information* ($\overline{PI}$) is 1,5 persons per enterprise (Table 5.2.4).

Table 5.2.4 - Respondents LES2015: number of persons Involved in providing information by size class (percentages) (a)

	1[(b)]	2	3	4	5	6+	Average number of Persons Involved in providing information
Total	64.1	24.0	9.4	2.5	0	0	1.5

Source: Istat - LES survey: study on the burden measurement system
(a) Missing: 367.
(b) In this category enterprises answered 'zero persons involved' are included.

Finally the *Average Retrieval Time* ($\overline{ART}$) is reported in Table 5.2.5. Standard deviation shows high dispersion of the data around the mean.

Table 5.2.5 - Average Retrieval Time

	Missing	N	Mean *(minutes)*	Standard deviation
Total	367	997	120.8	128.7

Source: Istat - LES survey: study on the burden measurement system

As it is shown in Table 5.2.6, less 10% of enterprises encountered 2 or more difficulties compiling the questionnaire.

Table 5.2.6 - Respondents LES 2015 by number of difficulties encountered in compiling the questionnaire (percentages)

	Number of difficulties				
	0/missing	1	2	3	4+
Total	67.03	23.27	7.56	1.69	0.44

Source: Istat - LES survey: study on the burden measurement system

The summary report on response burden for LES survey is shown in Figure 5.2.

Figure 5.2 - Summary report for LES survey year 2015

(1) $\overline{ACT}$ = Average Compilation Time in minutes: 42.6

(2) $\overline{ART}$ = Average Retrieval Time : 120.8

(3) $\overline{PI}$ = Average number of Persons Involved in providing information: 1.5

(4) Frequency distribution of enterprises per number of problems encountered

Number of difficulties declared by the respondent	*0/missing*	*1*	*2*	*3*	*4+*
% of respondent enterprises	67.03	23.27	7.56	1.69	0.44
Number of respondent enterprises	911	316	103	23	6

- Number of respondent businesses used: 1,422

- Coordination function used in selection of the sample: *Census survey*

Source: Istat - LES survey: study on the burden measurement system

5.3 First results of burden indicators on Industrial Turnover and new Orders survey

The monthly turnover index is designed to measure the performance of industrial sales over time, limited to mining and manufacturing activities. The orders index measures the dynamic of the value of new orders received by enterprises from clients each month. This second indicator is based on the information provided only by a sub-group of enterprises participating in the survey, in particular by the industrial sectors which usually work on commissioned orders.

The reporting unit for both variables is the enterprise; however, if a firm's turnover/orders refers to more than one economic activity (at three-digit level of NACE), data are collected separately for each kind of activity unit (KAU).

Turnover is defined as the total value of all the invoices issued during the month, for sales in the domestic or non-domestic market (divided into Euro and non-Euro areas), net of VAT invoiced to clients and any discounts or rebates shown in the invoice, before expenses (shipping, packaging, etc.) or other duties.

The Industrial Turnover Indices are sent to Eurostat in compliance with the European regulation no. 1165/98 on STS-term statistics and subsequent amendments and additions.

Orders include all the new orders – in term of value – received and accepted during the month. The information is disaggregated according to whether the orders come from domestic or non-domestic clients.

The survey refers to enterprises with more than 20 employees[10], the sample is selected from the Statistical Business Register with a cut off criteria[11], the index is a fixed index with 2010 as base year[12]. The sample consists of about 8,900 companies that are monitored carefully to ensure a low rate of non response, that was near to 10% for preliminary monthly estimates and 4%[13] for the final estimates released in November.

10 For particular sectors characterised by small enterprises the size could be lower.

11 The sample is extracted to cover for all sectors -defined at 3 digit level of NACE - more or less the 70% in terms of turnover of each sector.

12 See note n.6

13 Monthly reminders by e-mail and intensive follow-ups by phone are made to non-respondent units. A legal annual warning (with a financial penalty) is sent to firms that didn't answered for at least two months in a year.

At the beginning of 2016, a set of 2,000 enterprises has been added to the sample to prepare the next rebasing (base year=2015); as these enterprises are 'new' and not very skilled with the survey, the rate of non response of this particular subset of enterprises was higher[14].

In 2015 the survey was redesigned as a Computer Assisted Web Interviewing (CAWI) survey and the questionnaire is available through the new data collection system integrated into the Business Statistical Portal (GX).

As for the other two surveys previously analysed, with the introduction of the Business Statistical Portal some significant innovations were introduced:

- reduction of number of questions: the split of foreign orders into euro and non-euro area was removed;
- introduction of some automatic checks to highlight inconsistencies during the filling of the electronic form.

Also the turnover questionnaire has a new section to measure *respondent burden*. The filling of this section was not mandatory for the 2015 and no alert appeared if the section was not filled.

In the following, the results of the analysis on the compilation time and on the *Burden section* are presented.

As seen previously, survey's response rate in 2015 is different according to the different purposes of the analysis: if the enterprises belong to the sample referred to the 2010 base (6,500 enterprises) or to the new set of enterprises extracted for the 2015 rebasing and not already considered in the calculation of the index currently disseminated.

In addition, as only the subset of businesses which provided data for 12 months is considered to calculate the burden indicators (see Paragraph 4.1), the response rate for 2015 is 90% (6,252 respondents out of 6,928 sampling enterprises).

The *Average Compilation Time* ($\overline{ACT}$) is equal to 5.4 minutes, the standard deviation to 6.2.

14 These enterprises don't contribute to the index calculation before the introduction of 2015 as base year.

Table 5.3.1 - Compilation time in minutes - $\overline{ACT}$

	N	Mean *(minutes)*	Standard deviation
Total	6,252	5.4	6.2

Source: Istat - Turnover and new Orders survey: study on the burden measurement system

The respondents to the *Burden section* in December were 1,468, while the non respondent units were 5,059. In this case the number of units that were supposed to respond were 6,527 because this section is referred to each respondent unit (KAU) (Table 5.3.2).

Table 5.3.2 - Respondents Turnover 2015: *Burden section*

	Number of units	Respondent to LES December and not to burden	Respondent both to LES December and burden
Total	6,527	5,059	1,468

Source: Istat - Turnover and new Orders survey: study on the burden measurement system

Table 5.3.3 shows the time to retrieve information by size classes; the 66.8% of respondents needs less than half an hour, while the 24.5% needs more than half an hour but less than 1 hour.

Table 5.3.3 - Respondents Turnover 2015: time to retrieve information by size class (percentages)(a)

	Time to retrieve information		
	Up to 30 minutes	More than half an hour and up to 1 hour	More than 1 hour
Total	66.8	24.5	8.8

Source: Istat - Turnover and new Orders survey: study on the burden measurement system
(a) Missing=5,059.

The *Average number of Persons Involved in providing information* ($\overline{PI}$) is 1.3 persons per enterprise (Table 5.3.4), the 95.5% of enterprises involving a single person in filling in the questionnaire.

Table 5.3.4 - Respondents Turnover 2015: number of persons Involved in providing information (percentages)

	1[(b)]	2	3	4	5	6+	Average number of Persons Involved in providing information
Total	95.5	3.8	0.6	0.1	0	0	1.3

Source: Istat - Turnover and new Orders survey: study on the burden measurement system
(a) Missing: 5,059.
(b) In this category enterprises answered 'zero persons involved' are included.

Finally the *Average Retrieval Time* ($\overline{ART}$) is reported in Table 5.3.5. Standard deviation shows very low dispersion of the data around the mean.

Table 5.3.5 - Time to retrieve information (a)

	Missing	N	Mean *(minutes)*	Standard deviation
Total	5,059	1,468	31.6	30.25

Source: Istat - Turnover and new Orders survey: study on the burden measurement system
(a)Central values of time classes shown in Table 1 have been used.

As it is shown in Table 5.3.6, 1 less than 2% of enterprises encountered 2 or more difficulties compiling the questionnaire.

Table 5.3.6 - Respondents Turnover survey by number of difficulties encountered in compiling the questionnaire (percentages)

	Number of difficulties				
	0/missing	1	2	3	4+
Total	81.85	19.27	0.31	0.05	0.05

Source: Istat - Turnover and new Orders survey: study on the burden measurement system

The summary report on response burden for Turnover survey is shown below.

Figure 5.3 - Summary report for Turnover survey year 2015

(1) $\overline{ACT}$ = Average Compilation Time in minutes: 5.4

(2) $\overline{ART}$ = Average Retrieval Time : 31.6

(3) $\overline{PI}$ = Average number of Persons Involved in providing information: 1.3

(4) Frequency distribution of enterprises per number of problems encountered

Number of difficulties declared by the respondent	*0/missing*	*1*	*2*	*3*	*4+*
% of respondent enterprises	81.85	19.27	0.31	0.05	0.05
Number of respondent enterprises	5,343	1,258	20	3	3

- Number of respondent businesses used: 6,252

- Coordination function used in selection of the sample: *the sample is selected through cut-off criterion*

Source: Istat - Turnover and new Orders survey: study on the burden measurement system

6. Automatic Procedure for Burden Indicators by Survey (BBS)

A generalised software application for the production of the indicators described in the previous section has been implemented.

This procedure runs for all the surveys in the Business Portal architecture, using the GX software for the electronic questionnaire (the possibility of expanding it also for questionnaires developed with other software systems is under study).

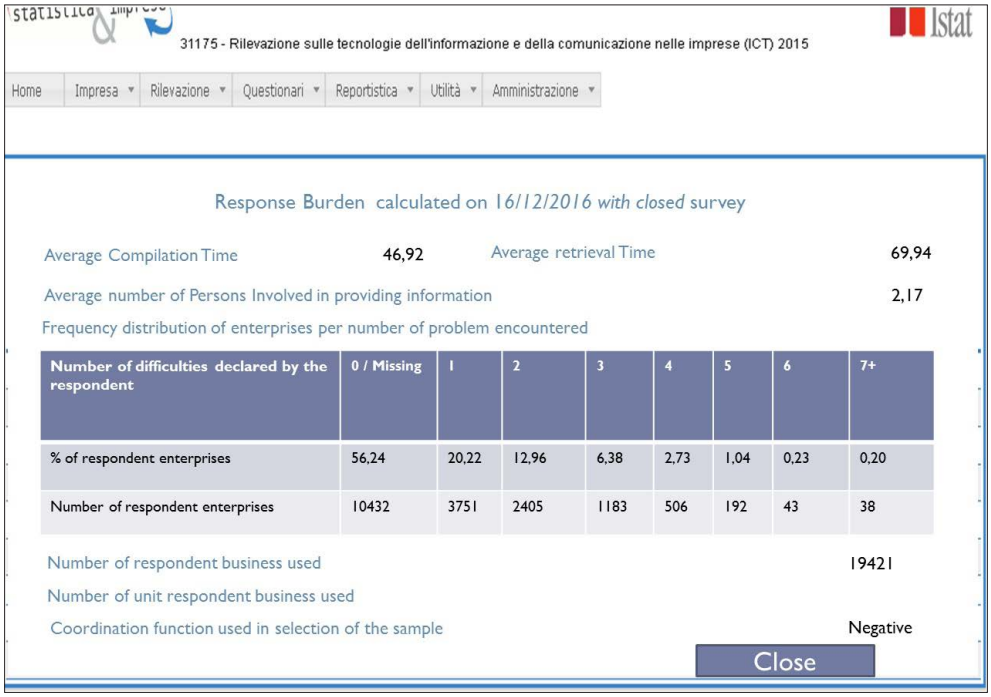
In practice, each time a new survey starts, the responsible of the survey will define:

- the type of the *Burden section* to be used for his survey: as described in Section 4, the general structure of this section is customised according the characteristics of the survey questionnaires (short, medium or long);
- the use of sampling coordination function;
- the reference of RDBMS where the *Burden section* microdata are stored;
- the year of reference of SBR ASIA archive to retrieve information regarding size class of employees or Economic Activity sector of the businesses.

Further information to be provided by the responsible of the survey is the starting and ending date of the survey to be considered, in order to select the corresponding paradata to be stored in the RDBMS.

In the Business Statistical Portal Section 'Online Report', the procedure provides a new report to be selected in the list of those which can be generated: Summary report on response burden.

Figure 6.1 - Response burden report for the ICT survey

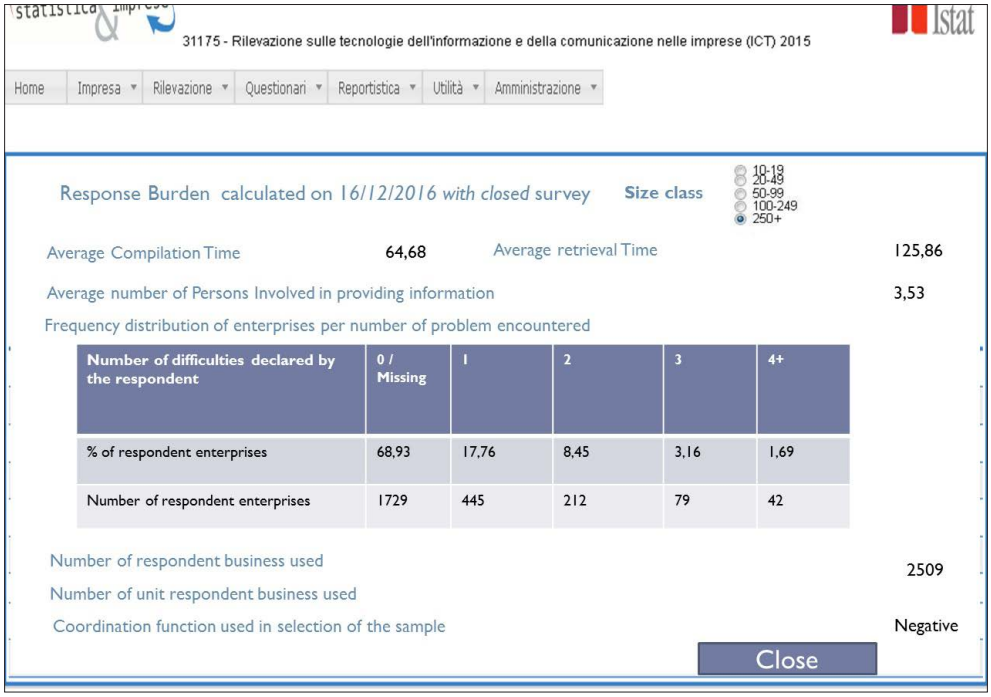


The number of columns regarding 'Number of difficulties declared by the respondent' will be different depending of the type of the *Burden section*.

During the data collection phase, the person in charge can run the reports and the software application processes both the paradata and the results obtained on the *Burden section* to generate the standard burden indicators by survey. However, the results will be stored in RDBMS only when the survey is definitively closed (the person in charge will close the calculation with the appropriate button).

The report produces also results by size class of employees, using the information already stored.

Figure 6.2 - Response burden report for the ICT survey, by size class of employees



7. Burden indicators in terms of persistency (BBB)

The persistency indicators aim at measuring burden from the point of view of a single enterprise potentially involved in several Istat surveys in a given time period. The main reason why Istat has decided to calculate the persistency indicators is that multiple requests addressed to the same enterprise may reduce the quality of the response given.

In the last years, many National Statistical Institutes have been increasingly using sampling coordination techniques to reduce the overlap between samples of different surveys. Nevertheless, large enterprises are more likely to be selected into survey samples, causing a 'persistent' burden to these statistical units. Indeed, they are often selected with probability 1 from different surveys or from the same survey in consecutive time periods.

In this report we produce two groups of persistency indicators referring to planned and actual burden for year 2015. The first group of indicators is based on the number of surveys a unit is selected for, while the second one considers the number of questionnaires the enterprise fills in. The calculation of the persistency indicators does not imply further burden on the enterprises, because it is based on information already available in Istat databases.

The indicators we present in this report are based on the following quantities:

S1, S2, ... , Sn = number of enterprises selected into (exactly) 1, 2, … , *n* surveys during year *t*

R1, R2, ... , Rn = number of enterprises that have filled in (exactly) the questionnaires of 1, 2, … , *n* surveys during year *t*.

From the quantities above we have calculated the following indicators:

$S'_k = \frac{\sum_{i=k}^{n} S_i}{N} \cdot 100$ = percentage of active enterprises selected into at least *k* surveys during year *t*

$R'_k = \frac{\sum_{i=k}^{m} R_i}{N} \cdot 100$ = percentage of active enterprises responding to at least *k* surveys during year *t*,

where N is the total number of active enterprises according to Asia, the Italian Statistical Business Register (SBR)[15], n and M are the maximum number of surveys an enterprise is involved in during year t[16].

Table 7.1 - Enterprises selected into at least *k* surveys during year 2015

Number of surveys (k)	1	2	3	4	5	6	7
% active enterprises selected into k surveys	4.96	1.20	0.56	0.21	0.32	0.14	0.09

Source: Istat - Study on the burden measurement system

Table 7.2 - Enterprises responding to at least *k* surveys during year 2015

Number of surveys (k)	1	2	3	4	5	6	7
% active enterprises responding to k surveys	2.37	0.66	0.32	0.19	0.12	0.08	0.05

Source: Istat - Study on the burden measurement system

The indicators of persistency have been finally calculated by size classes (in terms of number of persons employed) and economic activity, exploiting the information available from the SBR Asia[17].

Table 7.3 - Enterprises selected into surveys during year 2015, by size classes

Size classes	0 - 9	10 - 19	20 - 49	50 -249	≥250	Total
% active enterprises selected into surveys	2.49	42.77	67.34	93.33	100	4.960

Source: Istat - Study on the burden measurement system

Table 7.4 - Enterprises responding into surveys during year 2015, by size classes

Size classes	0 - 9	10 - 19	20 - 49	50 -249	≥250	Total
% active enterprises responding to surveys	0.82	22.31	43.43	72.67	96.07	2.37

Source: Istat - Study on the burden measurement system

15 Calculation is based on the updated SBR Asia from which the sampling frame has been taken (2013).

16 n and m are the maximum number of surveys an enterprises involved, in the sense of planned and actual burden, respectively.

17 About SBR Asia year of reference, see footnote above.

Table 7.3 shows that the share of enterprises that are selected into several surveys rapidly increases with the number of persons employed. In particular, it emerged that 2.5% of microenterprises (<10 persons employed) are involved in surveys, while this percentage grows to 49.5% for small enterprises (10-49 p. e.) and 94.5% for medium and large enterprises (≥50 p. e.). The share is 100% for large enterprises (≥250 p. e.) because in most Istat business surveys they are enumerated. We can also observe from Table 7.4 that unit response rate of larger enterprises is very high.

Table 7.5 - Enterprises selected into surveys during year 2015, by economic activity

Economic activity	Industry	Construction	Services	Total
% active enterprises selected into surveys	15.30	3.36	3.86	4.96

Source: Istat - Study on the burden measurement system

Table 7.6 - Enterprises responding to surveys during year 2015, by economic activity

Economic activity	Industry	Construction	Services	Total
% active enterprises responding to surveys	8.29	1.54	1.74	2.37

Source: Istat - Study on the burden measurement system

Tables 7.5 and 7.6 finally show that Industry has the largest share of enterprises selected and responding to Istat surveys.

8. Conclusions and perspectives

The activities carried out in 2016 by the Istat Working Group on the measurement of burden for surveys on businesses have been focussed on the development and testing of a set of standard indicators (and metadata) on burden to be calculated in a systematic way.

So far, the indicators 'by survey' have been tested on three business surveys with different characteristics ('ICT survey', 'Monthly survey on employment and labour cost in large enterprises' and 'Industrial turnover and orders') and the results are reported in Section 5.

The test has been useful not only to assess the proposed indicators, but also to define the requirements for the software application that should automatically calculate them. The proposed strategy is to add a standard *Burden section* to all surveys that migrate to the business portal IT environment, so as to implement a generalised software application that process both the paradata and the responses collected in the *Burden section* and calculate the standard burden indicators by survey (BBS). An example of this procedure has already been implemented for the GX data collection system, as reported in Section 6.

In perspective, it is planned that the burden indicators by survey will be stored in the Istat official system for the documentation of reference metadata and quality indicators, named SIDI/SIQual. This database should automatically be updated with these indicators in order to disseminate them to users through their integration into the National Quality Reports[18] that are produced through SIDI/SIQual and disseminated on the website starting from June 2018. These enhancements are still in progress as SIDI/SIQual system is also starting to be redesigned.

Furthermore, the Working group tested the Burden indicators in terms of persistency for year 2015 and the results are reported in Section 7. A software procedure should be developed to calculate annually and automatically also this set of indicators. In this case the source to derive the indicators will be the db built to manage the outcome of data collection of all the businesses surveys.

18 "Schede standard di qualità" in Italian are available at: https://www.istat.it/it/metodi-e-strumenti/strumenti-per-la-qualit%C3%A0/schede-standard-di-qualit%C3%A0.

The last task of the Working Group on the measurement of burden for surveys on businesses was to make suggestions on the indicators interpretation in order to identify strategies for reducing the burden and hopefully improving the quality of data produced by the surveys.

From the test carried out, it resulted quite clear that the specific expertise of the survey managers is necessary to interpret the indicators and identify the possible areas of improvement. The survey managers also know, for example, if there are alternative sources (*e.g.* administrative data, 'big data', web scraping, etc.) that can be exploited to reduce the burden. For example, following this study on the ICT survey, some reflections were made about the need to reduce respond burden not only acting more on development of the questions and of their explanation inside the questionnaire, but also trying to experiment other data source to have information similar to collected ones by survey. In fact in year 2017 web scraping and machine learning techniques were used to produce alternative experimental estimates on three variables collected by ICT survey (enterprises offering in their web sites web ordering, job application functionality and link to social media).

However, often the implementation of improvement actions implies the involvement and coordination of different surveys. This is the case, obviously, of the application of negative coordination of samples among different surveys in order to reduce the burden in terms of persistency.

Another example is the reduction of redundancies in information asked on different surveys: with the introduction of the Istat Business Statistical Portal the demographic information on the enterprises are now managed and updated directly by the Portal and not asked separately by all the surveys[19], other information could be managed in the same way.

It is also well known that the 'subjective' burden can be reduced improving the data collection tool. Some examples of functionalities that are deemed useful and are not implemented at the moment are the possibility to print the draft version of the questionnaire, or the consultation of questionnaires of previous editions.

19 In the Business Portal environment, each respondent, before filling in a survey questionnaire, can check and update, if necessary, enterprise' stored demographic data. In this way, this information is available for all the surveys where enterprise is involved in.

Another functionality which could be improved in the Business Portal concerns the feedback of statistical data to the businesses. In particular, relevant data for the business are going to be provided together with some functionalities connected to them, like the elaboration of 'Graphical presentation' according to the business features (economic sector, dimension and territorial location).

In conclusion, this work has to be analysed taking into account the transitional period during which it has been implemented. Major changes occurred to the organisational architecture of the Institute, so that implementation of such a system of indicators and the proposal for a full dissemination should be tailored to the new scenario. Nevertheless, the underlying concepts, both from a methodological and IT points of view, have been defined according to the generalised criteria as requested from the modernisation process all the NSI are going through.

Hence the proposed system can represent anyway a milestone and the cost for its adaptation so as to use it systematically should be low enough.

References

Couper, M.P. 1998. "Measuring survey quality in a CASIC environment". In *Proceedings of the Survey Research Methods Section*, American Statistical Association: 41-49.

Edwards, W.S., and D. Cantor. 1991. "Towards a Response Model in Establishment Surveys". In Biemer, P.P., R.M. Groves, L. Lyberg, N. Mathiowetz, and S. Sudman (Eds.). *Measurement Errors in Surveys*. Hoboken, NJ, U.S.: Wiley & Sons.

European Commission, Eurostat, Directors of Methodology – DIME. 2014. "Memobust Handbook - Handbook on Methodology of Modern Business Statistics". *Methodological handbooks*. Collaboration in Research and Methodology for Official Statistics – CROS: Eurostat.

Fazio, N., M. Murgia, and A. Nunnari. 2013. "The Business Statistical Portal: a new way of organising and managing data collection process for business surveys in Istat". In *Proceedings of the UNECE Seminar on Data Collection*. Geneva, Switzerland, 25th - 27th September 2013.

Italian National Institute of Statistics - Istat. 2016. *Istat's Modernisation Programme*. Roma: Istat.
https://www.istat.it/it/files//2011/04/IstatsModernistionProgramme_EN.pdf.

Masselli, M., A. Nuccitelli, and A. Laureti Palma. 2018. "The behaviour of respondents while filling in a web questionnaire: the case of the Italian Business R&D Survey. *Istat working papers*, n. 24/2016. Roma, Italy: Istat.

Nuccitelli, A., A. Nunnari, A. Nurra, and N.R. Fazio. 2014. "Towards a systematic use of paradata: the case of the Italian ICT survey". Presentation at the *7th Internet Survey Methodology Workshop*. Free University of Bozen-Bolzano, Italy.

Sudman, S., D.K. Willimack, E. Nichols, and T.L. Mesenbourg. 2000. "Exploratory Research at the U.S. Census Bureau on the Survey Response Process in Large Companies". In Proceedings of the *Second International Conference on Establishment Surveys*. American Statistical Association: 327-337.

Willimack, D.K., and E. Nichols. 2001. “Building an Alternative Response Process Model for Business Surveys”. In Proceedings of the *Annual Meeting of the American Statistical Association*.