istat
working
papers

A mixed-mode survey using mail and web questionnaires: the SCI case. First analysis

Alessandro Zeli

Istat
Istituto Nazionale
di Statistica

istat
working
papers

A mixed-mode survey using mail and web questionnaires: the SCI case. First analysis

Alessandro Zeli

Istat

Istat Working Papers

**A mixed-mode survey using mail and web questionnaires:
the SCI case. First analysis**

A mixed-mode survey using mail and web questionnaires: the SCI case. First analysis

Alessandro Zeli

Sommario

Questo documento illustra i risultati ottenuti nei primi anni di applicazione della modalità Web per la raccolta dei dati SCI (Rilevazione sul sistema dei conti delle imprese). Lo SCI è la principale indagine annuale condotta dall'Istat sui conti delle imprese e l'introduzione della modalità Web, che ha affiancato per tre anni (2003-2005) la modalità cartacea tradizionale, era volta a migliorare la qualità dei dati e a diminuire l'onere statistico per le imprese. Nel documento vengono analizzate le diverse dimensioni della qualità di indagine: tempestività, accuratezza, tasso di risposta, onere per le imprese, costi. Le evidenze dimostrano che l'introduzione della modalità Web ha garantito un incremento complessivo nella qualità dei dati.

Parole chiave: Web survey, procedure di indagine, qualità dei dati, tempestività.
JEL: C80

Abstract

The present paper describes the results obtained after some years of use of electronic questionnaires for the collection of data related to Istat Survey on enterprise accounting system(SCI). SCI is the main annual survey on enterprise accounting carried out by Istat. The introduction of electronic questionnaires, which was adopted in addition to the traditional paper mode for three years (from 2003 to 2005), aimed at improving data quality and lessening enterprise statistical burden. The paper focuses on different dimensions of quality such as: timeliness, accuracy, response rate, enterprise statistical burden and costs. The results shown in the document demonstrate that a notable improvement of survey data quality was achieved by means of electronic questionnaires.

Keywords: Web survey, survey process, data quality, timeliness.
JEL: C80

Index

1. Introduction

The impacts of new technologies on survey data collection quality may be evaluated in terms of efficiency, productivity, accuracy, production factor economic saving, improvement of organisation and increasing profitability.

From this point of view, the information and statistical data production may be compared to any other production process; the paper presents the case of SCI (survey on Accounting System of Enterprises); this one is an historical survey carried out by Istat concerning balance sheet values and other variables such as: employment, personnel costs, investment. SCI survey is also aimed at the estimation of the GDP. The step by step innovation may create difficulties upstream and downstream of the point (production phase) interested in the innovation (bottle neck). In the case of a statistical survey the most important steps are: data collection, data capture, data editing, validation and dissemination.

In 2002 specific innovations were introduced in SCI production process for the editing phase (automation and computerisation), the validation phase (creation of an *ad hoc* software aimed at enterprise profiling) and for dissemination in order to simplify this process too. Notwithstanding these improvements the survey timeliness did not register a substantial progress, because of traditional data collection carried out by means of paper questionnaires. Paper questionnaires data collection requires time consuming processes., The share of resources absorbed by data collection and data capture steps remained at an abnormally high level in comparison to the improvements attained for the editing, validation and dissemination steps. It also needed to cope with a decrease in available human resources assigned to the SCI survey making the resource constrains worse.

In order to solve these problems, a web questionnaire for the SCI survey was introduced from reference year 2003 in parallel to the usual paper questionnaire. The web questionnaire was also utilised in the following edition of the SCI survey: reference years 2004 and 2005.

Carring out a mixed-mode survey was a useful experience (demonstrated in this paper) and it had to be deeply understood in order to develop new behaviour patterns in communication strategy for firms. We are going to face in the next years change and challenges coming from the new methodologies of data collection that are implementing in more recent years in Istat: for instance the development of the "Enterprise portal" and XBRL Statistical Taxonomy Project. The experience coming from the introduction of electronic questionnaires in an important and complex survey in Istat can supplies some useful information on how to manage the new techniques too. The documentation of this experience can help to answer to the following question in the next future: is it a step by step strategy preferable? A mixed mode is the right approach? How firm's characteristics affect its tendency to accept new data collection technologies?

This paper is aimed to produce a first analysis and assessment of the experience gained until now and seeks to take the opportunity of studying a mixed-mode survey. There is not a lot of literature on parallel mixed-mode survey analysis, however they can be advised the following studies: M. Meckel, D. Walters, P. Baugh, (2005) and Biffignandi S., Fabrizi E. (2006).

The second paragraph contains a brief presentation of the relevant literature on this matter and describes the main statistical questionnaire collection methodologies, the mixed-mode survey and the criteria that the various survey modes should satisfy in order to achieve a better level of efficiency. The third paragraph introduces the SCI survey and presents the survey modes utilised for the questionnaires collection.

In the fourth paragraph we try to detect the influence of the main identification enterprise variables (size, economic activity) on the response modes: paper and web.

The fifth paragraph analyses the statistical results and the differences between the two survey modes according to various parameters such as: data collection timeliness, accuracy, time spent in filling in questionnaires and so on.

The last paragraph describes the possible explanations of the presented results and evaluations gathered from the experiences done.

2. The survey technologies and questionnaires management

The problem faced in the literature regarding statistical survey and questionnaires management are mainly related to the efficiency of the survey itself. Regarding the questionnaires sending and data collection steps the technique and survey mode play an important role in increasing the efficiency (Couper M.P. 2002). The efficiency of the survey mode can be measured by a set of characteristics such as:

- response rate;
- accuracy;
- consistency;
- timeliness;
- survey costs;
- easy data electronic storage.

Every researcher would like to see these characteristics matched as completely as possible by the survey techniques chosen. The Web supplies more and more powerful technologies and gives the opportunity to improve: accuracy, data collection timeliness, spares in survey management and data electronic transfer and storage as well. However, Internet technologies for the collection and transfer of information can not be applied to all possible cases and the sending and response paper mode still remains highly utilised; in certain cases it is essential for a successful survey.

Moreover, a scientific approach to the problem requires that the two survey modes have to be compared from a statistical point of view in order to evaluate the differences and the actual efficiency increase due to electronic mode.

Selecting a survey mode, it is necessary to make a comparison of advantages and disadvantages of a technology, for instance the possibility of access to electronic questionnaire via the Internet may cause problems such as a lack of full coverage of the target population (because of population units may have no connection to the web or technical difficulties to access to the global Internet), the questionnaire length may make the download difficult and the drawing up an on-line questionnaire problematic. The possibility of printing the questionnaire and the related instructions need to be assured.

The management of a statistical survey is in general unique. The choice may be a posted or faxed paper questionnaire (for the reasons explained above). An alternative is represented by an electronic questionnaire uploaded on a web site exclusively dedicated to the survey or via e-mail. The electronic questionnaire may be compiled on-line or off-line. Other survey techniques utilised in collecting enterprises data are CATI and, above all, CAPI, even if these techniques are quite expansive and don't imply a reduction of costs (Balestrino R., 1998, Balestrino R., Barcaroli G., 1998).

It is useful to have at your disposal studies and analyses on the mix-mode surveys results in order to have a better understanding and knowledge of the different techniques in order to make the better choice. In the literature, there are not a lot of studies that compare parallel mixed-mode directly, moreover the results are ambiguous.

A study (M. Meckel, D. Walters, P. Baugh, 2005) regarding the results of a mixed-mode survey (paper-web) concerned with 1,000 SME based in the North-West of England shows an increase in the response rate due to the utilisation of the web-mode; this is also confirmed by another analysis (Dillman et al., 2001) that issues "giving respondents a choice of response mode does not necessarily improve response rate".

In Meckel et al. (2005) two different response modes are correlated with certain enterprise characteristics (especially variables related to the tendency towards the ICT technologies). Obviously, the more technologically oriented the enterprises are, the higher the web-mode response will be. A survey carried out in a mixed mode (Biffignandi S., Fabrizi E. 2006) aimed to investigate the use of Internet and e-commerce for a sample of enterprises in the Bergamo district, represents an interesting analysis to consider in this paper. The enterprises which were contacted first by phone were invited to choose questionnaire completion either via the web or receiving it by traditional mail or

fax. The survey results show a clear tendency to web-mode (59 per cent), the remaining 40 per cent was distributed between fax (38 per cent) and mail (around 3 per cent).

The web-mode response tendency appears more intense for larger enterprises. As regards response timeliness the better results are shown by web-mode response; after the first call-back this represented 52 per cent of the final total web-mode respondents whilst the total fax respondent accounted to 45 per cent.

3. Main SCI survey methodologies and organisation features

3.1. Survey on economic and financial accounts of large enterprises (SCI)

In Italy a detailed survey on the economic and financial accounts of enterprises is carried out annually by ISTAT. This survey is intended to cover all enterprises operating in Italy with at least 100 persons employed (ISTAT 2007).

The survey collects data concerning profit-and-loss accounts and balance sheets according to the normative guidelines of the 4th EEC Directive scheme. Moreover, information regarding employment, investment, personnel costs and certain regional items is also collected. In particular the first section of the questionnaire contains data on profit and loss accounts and the second one on balance sheet data. Employment, broken down by main professional skill, is requested in section 3. Section 4 contains information concerning personnel costs and section 5 contains information regarding investment. In section 7 the main economic variables are broken down by region. A new section (section 8) is devoted to special surveys on enterprise behaviour such as: delocalization, globalization, social responsibility, innovation and so on. Data about KAU (Kind of Activity Units) are required for greater enterprises. The survey concerns all industrial and service sectors enterprises excluding financial services and the target population includes around 10,000 units.

The survey utilises the NACE classification Rev.1.1. The SCI survey concerns enterprises that operate in all economic sectors but the agriculture, finance and public administration.

3.2 Data collection methods

Since reference year 2003, data has been collected by a mixed mode (via the web with an electronic questionnaire and a paper questionnaire via ordinary mail) and several follow ups (mail and telephone) have been carried out. A mixed mode was chosen in order to decrease the statistical burden for enterprises, to gain rapidity, accuracy and reduce survey management costs. Moreover SCI survey concern only the biggest enterprises so Internet availability can be assumed to be present.

The decision to maintain also the paper questionnaire was taken in order to make the enterprises more confident with new electronic procedures and maintain the response rate. In the first year (2003) both the electronic and the paper questionnaires covered the whole SCI target population. In 2004 and 2005 units that utilised the web mode in previous years received only the electronic questionnaire in order to save money and consolidate the passage towards these new procedures. The response modes overlap decreased year after year.

Survey management requires the following steps (Zeli A., 2006): questionnaire preparation and delivery, data collection and the registration of respondents and data. These steps are planned in different ways according to the survey mode chosen by survey managers. In the case of the SCI survey both a paper mode and a web mode are available for enterprises from reference year 2003 to reference year 2005.

3.2.1 The paper mode

The paper mode needs a very accurate questionnaire preparation and the items required in the questionnaire have to be up-to-date with the current Italian legislation in terms, for instance, of finance reporting and worker status. The paper questionnaire is prepared in Excel format and transformed in a PDF file before being transferred to typography for printing, customizing and finally

posting. The questionnaire is posted together with instructions, a pre-paid reply envelopment, and a cover letter. The data collection is supported by a call-centre managed by the survey staff ; many reminders are carried out during the data collection period.

Completed questionnaires are registered as having arrived manually and are recorded on a magnetic tape in outsourcing. A consistency check is carried out before the outsourced recording. Then the questionnaires are returned to Istat, the data are uploaded on a working database for checks and submitted to editing procedures. There are two steps in this procedure where it is possible to achieve great costs and time processing reductions: printing and recording in outsourcing. Too many passages also imply a lack of accuracy (manual registration of questionnaires and uncontrolled outsourced recording).

3.2.2 The web mode

The preparation of electronic questionnaires is, of course, as accurate as the paper one. The electronic questionnaire is prepared utilising the Excel software; this software is widely adopted as a business tool and it is relatively user friendly for every response mode.

To avoid any possible influence on the answer through different ways of presenting the questionnaire the layout of both versions was kept as similar as possible. The enterprises opening the Excel file immediately recognise the electronic questionnaire and its different sections, a single Excel sheet is devoted to each questionnaire section and the sequence is unchanged. The questionnaire in Excel format is prepared in such as way as to contain many automatic controls driven by Excel formulas such as: main squares and cross-sections checks. This gives the respondent more confidence when filling in the questionnaire items. The check of outliers is not implemented because of the need of an auxiliary databases. When the Excel questionnaire is ready the automatic data capturing is organised.

A cover letter is posted to enterprises in order to communicate the possibility of downloading the electronic questionnaire and instructions for uploading the completed questionnaire. The Istat server utilised is a secure one and sending of e-mail is discouraged.

In order to increase security, a code number and a secret password is supplied to enterprises. The password generation process use a random logic; it generates a string of numbers and letters with no meaning in Italian (or in other known languages).

The respondents can download the Excel questionnaire from the Istat site and fill in it. When the questionnaire is completed it can be uploaded on the Istat site again. From the Istat site the questionnaires are addressed to the Economic Statistics Department server.

Each part of the SCI questionnaire is devoted to different data collection purposes (for instance data are collected in order to update the Business Register, to provide information on expenses for environmental protection and so on). A computerised data-capturing system stores the electronic questionnaire data in a storage server for the registration of the respondents and then each part of the questionnaire is redirected towards a specific working database.

The main flow of information is, of course, related to the economic accounts of enterprises. This data flow is automatically checked for list or identification mistakes (for instance code errors or eligibility conditions) and uploaded on the SCI survey working database for the data editing procedures. Every week a file containing the respondent codes is created in order to update the register of respondents, making it possible to monitor the response trend and manage the call back operations.

In the case of web mode, the survey costs decreased owing to a change in production process organisation and namely to the elimination of two outsourced process steps (printing and recording), as will be demonstrated in the following sections.

### 3.3 Mixed-mode survey results. Years 2003-2005

The SCI survey deals with the largest enterprises in Italy and does not face considerable technological or personnel skill problems. Enterprises generally appreciated and exploited the electronic questionnaire.

In April 2005 when the first mixed-mode data collection ended (SCI 2003 survey) a fast and explorative postal survey was carried out in order to collect information about non respondent and paper-mode respondent enterprises. In particular, the purpose of the fast survey was to attain a willingness of the respondents to fill in the electronic questionnaire in the next year.

Around 1,420 enterprises (out of 2,057 respondents) declared a willingness to fill an electronic questionnaire, they represented a share of 69.3 per cent,

Response rate (27.6 per cent) was not, as expected, very high. However, the results were quite encouraging (about 70 per cent positive responses) and it is possible that the respondent enterprises to the explorative survey and the enterprises available to use the web-mode to the SCI survey are the same. If the enterprises available to use the web-mode for the SCI 2004 survey are added to the enterprises that already utilised the web-mode for SCI 2003 survey (1,633 units) a possible web-mode respondent area of 3,000 enterprises can be predicted (that is a share of over 60 per cent of the total survey respondents). Notwithstanding that this target was not achieved in 2004 in 2005 SCI survey edition the web-mode respondents represented a share of 68.9 per cent of total respondents. The incidence of the web-mode responses has constantly augmented since 2003, as shown in Graph 3.1.

Graph 3.1 - SCI survey response percentage by response mode on total respondents - Years 2003-2005
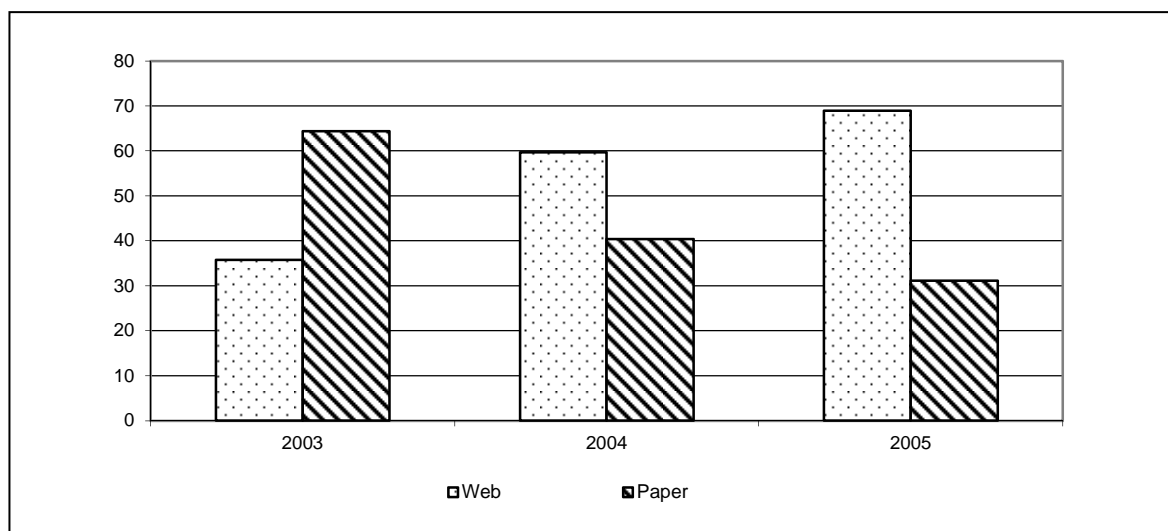
The web-mode increased from a share of 35.7 per cent in 2003 to a share of 68.9 per cent in 2005. It is proof of the increasing appreciation of enterprises for this response mode.

4. Respondents analysis by response mode

A preliminary analysis of respondents by response mode has been focused on the specific characteristics of the enterprises grouped by response mode, the analysis is important to detect some determinants for web mode response but also in order to detect some factor that can influence a better performance of web response (i.e to detect some self-selection bias factor)

Let us considered the identification data that are utilised for the target population statistical stratification:

- enterprise size in terms of person employed;
- enterprise economic activities;
- enterprise geographic area.

The analysis variable is represented by the web-mode responses share on total validated responses: that is web questionnaires plus paper ones. The web responses share was calculated for all

strata detected by the above mentioned stratification variables for 2005 SCI survey data. In particular 3 modalities were considered for the variable size: medium (from 100 to 249 persons employed), great (from 249 to 999 persons employed), large (1000 persons employed and over); 5 modalities were considered for the economic activity (mining and quarrying, manufacturing, electricity, gas and water supply, constructions, services) and, finally, 4 modalities were considered for the geographic area: North-West, North-East, Centre, South). The web mode response rates by strata are presented in Table 4.1.

Table 4.1 - Web mode response rates means and standard deviations by size, geographic area and economic activity - Year 2005

STRATIFICATION VARIABLE	Means	Standard Deviation
SIZE		
100-249 persons employed	68.4	15.3
250-999 persons employed	73.3	11.7
1000 persons employed and over	81.8	15.0
GEOGRAPHIC AREA		
North -West	70.3	10.3
North – East	71.1	15.4
Centre	68.3	14.0
South	71.5	19.5
ECONOMIC ACTIVITY		
Mining and quarrying	70.0	26.5
Manufacturing	71.2	6.1
Electricity, gas and water supply	79.1	18.1
Construction	75.1	19.8
Services	68.1	7.8

The data highlights great differences for the modality of size, with a highly evident and positive correlation between size and web mode response. The differences for the modality of geographic area are less evident; the maximum is registered for *South* (71.5 per cent) and the minimum for *Centre* (68.3 per cent). The scenery is more complex for economic activity; it registers a higher web response rate for the electricity, gas and water supply (79.1 per cent) and lower for mining and quarrying and services (70.0 and 68.1 per cent respectively). An ANOVA was carried out in order to understand the variables that influence the web response rate. The results are shown in Table 4.2:

Table 4.2 - ANOVA results on web response determinants

VARIABLES	F Value	Pr > F
Size	4,73**	0,0167
Economic activity	1,23	0,3203
Geographic area	0,85	0,4785
Size* Economic activity	2,44**	0,0492
Size* Geographic area	1,11	0,3808
MODEL		
R- square	F Value	Pr > F
0,568641	1,82*	0,0671

* Significance: p<.10 ** Significance: p<.05

As shown in Table 4.2 the analysed model does not have a very high F significance. In any case, it can be accepted within 10 per cent significance. The R-square is near to 57 per cent and is therefore satisfactory. The only variable that influences the web mode response rate is the enterprise size and the interaction between size and economic activity. This result is in accordance with other studies carried out on enterprises (e.g. Biffignandi S., Pratesi M., 2000) from which it is evident that the enterprise size significantly influences web-mode survey participation tendency. The other variables included in the ANOVA do not have a significant influence on web mode response rate.

This result is important in order to evaluate correctly the results shown in next section. If there is an *a priori* tendency of larger firms to be a respondent it may influence also the better performance of web mode with respect to paper one. Anyway there is not evidence of correlation between (for instance) the total number of detected errors and firms' size (correlation = 2 per cent) and between timeliness and firms' size (correlation=.2 per cent). Moreover given the nature of variables on which the tests are implemented, it can be quite confident that this problem does not matter.

5. Results

This section analyses both response modes, focusing on advantages and disadvantages in terms of the survey quality of electronic transmission in reference to the statistical survey efficiency criteria presented above. In particular:

- response rate;
- accuracy;
- timeliness;
- statistical burden for enterprises;
- organisation and survey costs for the Institute.

The analyses are carried out in reference to the 2005 SCI survey edition, the final year for which broader information is available and a comprehensive framework can be determined. The SCI questionnaire respondents are compared according to both response modes and efficiency measure categories presented above.

5.1 Rate of response

The survey response rate is one of the parameters that better measures efficiency of modes and techniques utilised in carrying out the survey. In the literature presented in the second section, evidence of improvements to this parameter by means of survey strategies based on web-modes are controversial.

In the SCI case-study it can be surely stated that the introduction of web-mode response possibility does not persuade, in any remarkable quantity, traditionally non respondent enterprises to fill in the questionnaires. The numbers presented in Table 5.1 confirm the substantial stability of the share of validated responses[1] on the total SCI survey target universe in the time period in which it was possible to choose between two response modes.

Table 5.1 - Validated response share on the total SCI survey target universe. Years 1998 - 2005

YEARS	1998	1999	2000	2001	2002	2003	2004	2005
Validated response share	44.9	46.7	46.3	48.1	46.3	48.4	47.0	49.7

---

[1] Only validated responses considered because this response category is not affected by events of mergers, demergers, enterprise death, response refusal and so on, and only the actual response-mode choice matter.

However, the web-mode is a factor that makes it easier to ensure a well-timed data collection; it increases both the timeliness of enterprise response and the receipt of the questionnaire on the Istat side. The increasing response velocity allows carrying out more call-backs and, indirectly, it encourages survey participation. So, the web mode may increase indirectly the response rate and this seems to be evident in the 2005 SCI survey edition.[2]

5.2 Accuracy

Data collected accuracy and consistency in respect to data definition required in questionnaires are analysed by means of two indicators:

- the error numbers detected in both respondents sub-populations (paper-mode respondents and web-mode respondents);
- the differences between the initial variables value (reported by respondent or raw data) and the value of the same variables after the editing process (validated data). The differences are calculated, of course, for the sub-population of web and paper mode respondents.

The revision and data editing process of the SCI survey raw data classifies the errors in three categories that can be listed, in descending order in terms of statistical and accounting importance, in the following way:

- error: occurs when a value cannot be corrected automatically (Replacement) and the procedure fails to go on to check other items;
- replacement: error of lesser importance, the incorrect value is replaced with another value;
- check pending: when a value has to be investigated interactively in order to accept or to correct it.

Values classified as *under check* or *check pending* are not necessarily errors but they can be outliers indicating anomalies in data entered in the questionnaire. The absolute number of errors may be misleading giving different numbers of two sub-populations, so the means relative to respondents by mode are presented.

An improvement in accuracy and quality of the raw data was expected because the filling in of the questionnaire is assisted by means of rules and formulas in the Excel sheets constituting the questionnaire (these rules and formulas help to sum the balance sheet items and to check the consistency between different items). The situation for the SCI survey 2005 grouped by two respondent sub-populations is presented in Table 5.2.

Table 5.2 - Types of errors detected by response mode. Absolute values and means - SCI survey 2005

	Number of questionnaires utilised in the analysis	Checks pending	Corrected errors	Replacements
Paper	1,424	1,271	6,750	6,964
Web	3,139	2,660	9,045	10,640
		AVERAGES ON TOTAL RESPONSES		
Paper		0.89	4.74	4.89
Web		0.85	2.88	3.39

The results in Table 4.2 show dramatic positive imbalances to electronic questionnaires. The average errors by questionnaire are in any case lesser for the web-mode questionnaires than for all other categories. Only the check pending category presents average data similar for both response modes but in any case this is the error category with less impact on the final estimates.

---

[2] In the 2005 survey edition, for instance, e-mail call-back has been launched.

Statistical tests are carried out in order to investigate if the absolute calculated differences were statistically significant. In particular two non–parametric tests have been carried out in order to verify if k population have the same mean: Kolmogorov-Smirnov and Kruskall-Wallis tests. The results are presented in Table 5.3. The results indicate significant differences for the means of the sub-population in the cases of errors and replacements categories. The null hypothesis is, instead, accepted in the case of the *check pending* category.

Table 5.3 - Results of statistical tests

	Kolmogorov-Smirnov		Kruskall-Wallis	
	CHECKS PENDING			
Test value	*KSa*	*0,528303*	*Chi-square*	*1,0706*
Probability	*Pr > KSa*	*0,9429*	*Pr > Chi-square*	*0,3008*
	ERRORS			
Test value	*KSa*	*7,856203*	*Chi-square*	*320,0017*
Probability	*Pr > KSa*	*<.0001*	*Pr > Chi-square*	*<.0001*
	REPLACEMENTS			
Test value	*KSa*	*6,494833*	*Chi-square*	*231,6015*
Probability	*Pr > KSa*	*<.0001*	*Pr > Chi-square*	*<.0001*

For the elaboration of the second indicator (differences between initial raw values and final validated values) it was necessary to compare the database of checked and validated data with the set of the initial data grouped by response mode. In the comparison process few variables were considered as being required in the SCI questionnaire, in particular the focus was on main economic variables:

- turnover;
- production value;
- labour cost;
- investment.

Table 5.4 presents comparison results by means of two indicators:

- the first one is the mean expressed as a percentage of the base value mean. In this case the base value mean is represented by raw data:

$$^{RM}I_1 = (M(\text{validated data}) - M(\text{raw data}))/M(\text{raw data})$$

- the second one is the difference between the mean of initial raw data and the mean of final data:

$$^{RM}I_2 = M(\text{validated data}) - M(\text{raw data})$$

Table 5.4 - Means differences between the raw and validated data by response mode - SCI survey 2005

	Percentage means differences on initial raw value	Differences between raw and validated averages
	TURNOVER	
Paper	-85,1	-761.348
Web	-54,5	-149.270
	PRODUCTION VALUE	
Paper	-87,0	-941.125
Web	-51,4	-140.942
	LABOUR COST	
Paper	-93,2	-241.146
Web	-68,8	-42.010
	INVESTMENT	
Paper	-81,0	-24.213
Web	21,7	1.145

The great differences between the values of the reported variables (or raw data) and the values of validated variables (assuming in general minus signs) are due to the presence of many measure errors (enterprises that fill in the questionnaire in Euro instead of thousands of Euro).

Data presented in Table 5.3 registers a better quality level for the web-mode questionnaires compared to paper questionnaires. The differences between raw data and validated data are lower for web questionnaires than for paper questionnaires, the data correction extent performed for web questionnaires was lower than the paper ones, and it implies a better degree of quality for web collected data.

5.3 Timeliness

Data collection timeliness is very important in order to set the efficiency of survey management. The possibility to have data closer and closer to the reference period is very important for the survey data user and even more for official statistics that have to match the timeliness according to the legal framework fixed by European Public Bodies.

One of the main reasons for implementing electronic questionnaires for the SCI survey was the reduction of the respondent reaction time. The opportunity to have questionnaire information in advance allows more confidence in scheduling and devoting more time to the steps of the qualitative data analysis. Respondent's reaction times were compared by means of the response percentage cumulate distribution on the total response at fixed time with reference to years 2003-2005 grouped by response mode. In this way the influence of different sub-populations numbers is eliminated. The results are presented in the following Graphs: 5.1, 5.2 and 5.3.

The analysis of Graphs 5.1 – 5.3 underlines that the web-mode respondents were quicker than paper mode respondents. During the period in which the mixed mode data collection was carried out the web-mode respondents were much quicker than the paper mode respondents, in particular in 2003 the 20 per cent response rate share was achieved almost two months before by the web-mode respondents in respect to the paper-mode ones and the 50 per cent around three weeks before. Regarding the web-mode response, in 2005 the 25 per cent response rate share was achieved around two weeks earlier than paper mode.
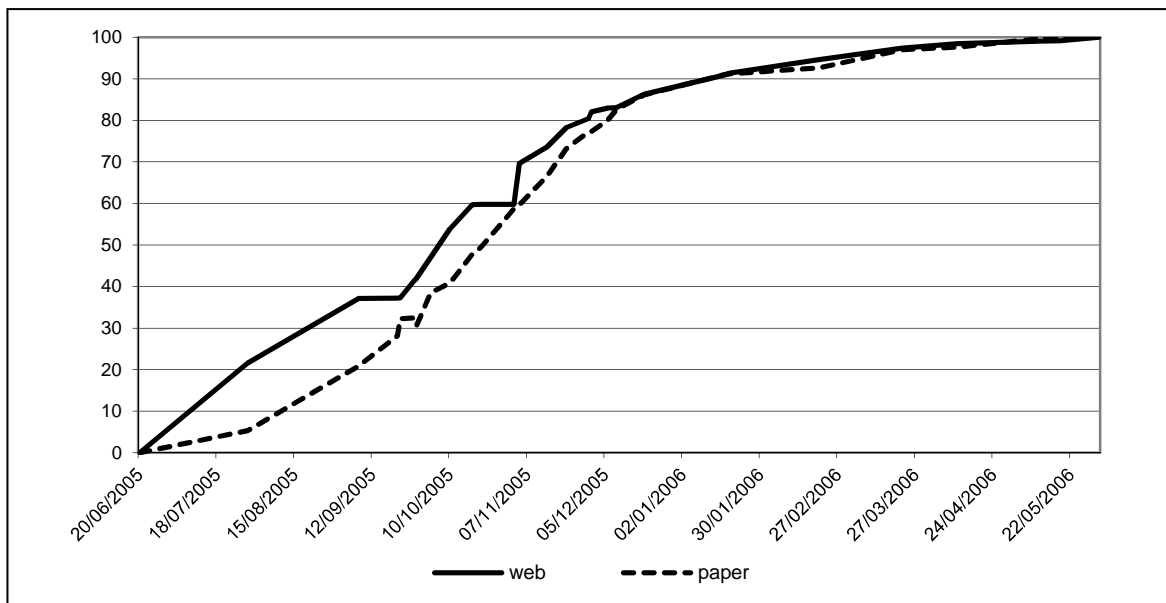
**Graph 5.1 - Time cumulate percentage of responses by mode - SCI 2003**
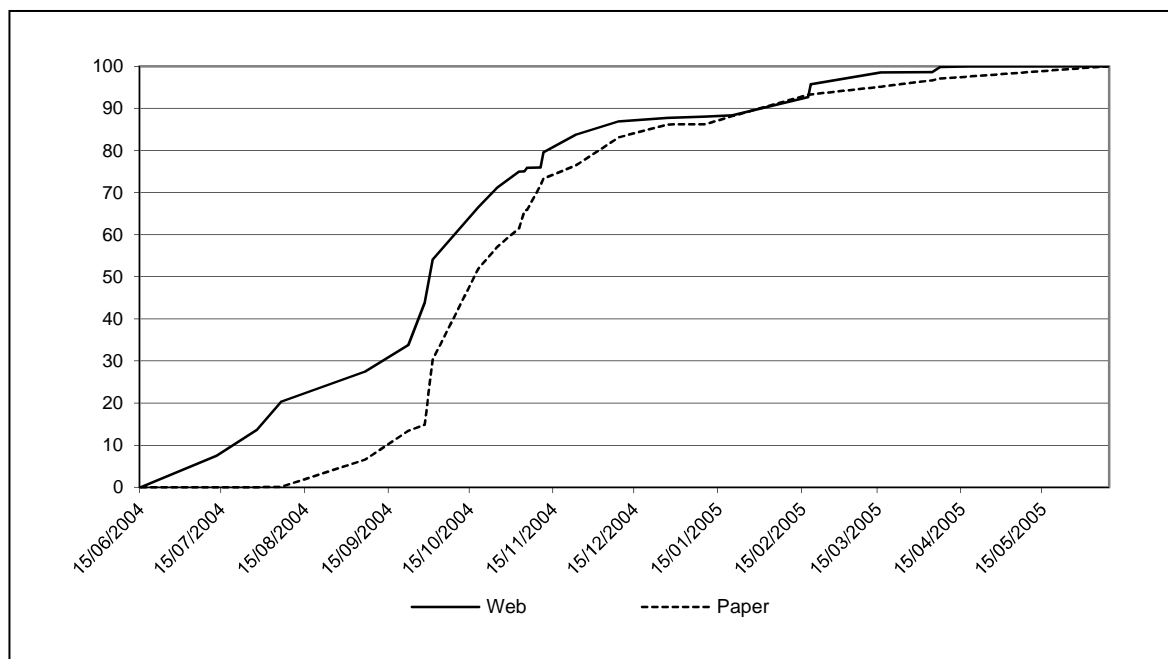
So, thanks to the electronic questionnaire, a decrease in response time was achieved[3] implying the need for a reduction in time devoted to data collection.

In fact, for the SCI survey 2003, 4.933 questionnaires were returned following the data collection step (30th March 2005). In the 2005 SCI survey edition the number of returned questionnaires was slightly greater (4.973) and was achieved almost two months before on 7th February 2007.[4]

Graph 5.2 - Time cumulate percentage of responses by mode - SCI 2004

Graph 5.3 - Time cumulate percentage of responses by mode - SCI 2005

---
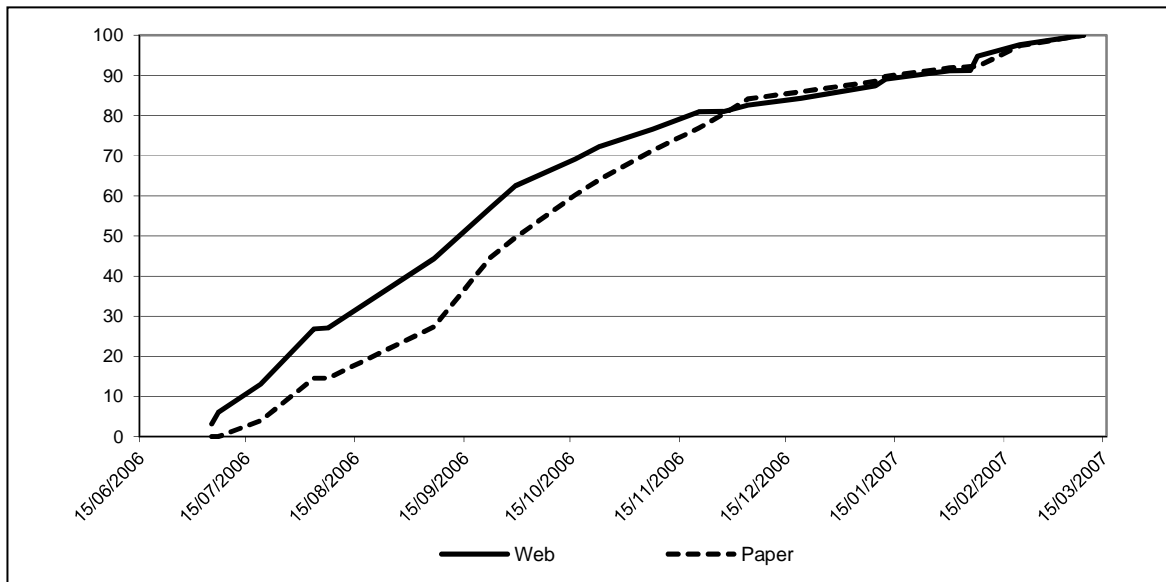
3   It confirms the results of other studies such as: Cobanoglu C., Moreo P. J, Warde B. (2001) and Tse A. C. B. (1998).
4   Although, as noted before, the SCI survey target population increases over the period 2003-2005 this does not explain the registered improvement in timeliness.

5.4 Statistical burden on enterprises

An important aspect to be considered in evaluating management efficiency for statistical survey is the statistical burden on respondents (in this case enterprises). In order to improve the timeliness, the accuracy and the rate of response is fundamental to make easier, as much as possible, the participation in the survey both in terms of costs and in terms of resources . The statistical burden on enterprises is measured by the time needed to fill in the questionnaire; this index considers only the direct utilization of resource but it does not evaluate other potential costs. However, it can be considered a good proxy of the total response burden.

It is necessary to highlight that this indicator, measured for the 2005 SCI survey, does not register an advantage for electronic questionnaires. The paper-mode respondents declared 6 hours and 40 minutes as the average time of compilation compared with 7 hours and 45 minutes as the average time of compilation[5] declared by web-mode respondents. These results may depend on the correlation between the response mode and enterprise size; in fact, the greater enterprises show the higher web-mode intensity and, of course, they have the more complex questionnaire and they spend more time to fill in. Moreover, it is necessary to keep in mind that all new technologies need time in order to be confident with and in order to train the personnel.

5.5 Organisation and costs

5.5.1 Organisation

The survey steps which registered the highest impacts of the electronic questionnaire were the printing/sending and the data recording on magnetic tapes. Both steps were outsourced so implying a strong and clear economic saving and therefore they will be investigated in the next paragraph; however it is necessary to highlight that many operations related to printing and recording were carried out in ISTAT and the subsequent data editing step is modified by electronic questionnaire.

As regards the aspects relative to questionnaire printing and sending, the human resources surplus was mainly due to the following organisational steps:

- print shop contacts;
- preparation and correction of paper questionnaires;
- repeated tests for personalisation print and set up.

The electronic questionnaire encouraged a time reduction in personalisation and sending. Now only the mailing of a notice to enterprises by means of ordinary mail is carried out. This enables the questionnaires to reach enterprises around mid June while in the previous survey editions the questionnaires were sent by the end of July and first responses arrived in September. The early return of questionnaires and the elimination of data recording on magnetic tape enabled (in June-July) the data collection and data editing steps to be done earlier and also allowed the follow-up phone calls to be carried out in October. .

The arrivals registration, the pre-check and the magnetic recording are the steps in which a surplus of human resources was greater. These are in the process of being eliminated and consist of such operations such as the manual operations related to arrivals registration, the paper questionnaires sorting, the manual pre-check, the questionnaires set up for the magnetic registration. The organisational and planning steps related to these activities were involved in a dramatic reduction too. Time gains in the data collection step are also going to have good effects on final data quality because this will enable the devotion of greater resources at check and data editing. It is also going to have a positive impact on job quality of survey personnel and encourage an improvement of professional skill. Good impacts are also foreseen with regards to the check and data editing steps. The

---

[5] Also, in this case non-parametric tests were performed in order to verify the statistical significance of the differences. The Kolmorov-Smirnov and Kruskall-Wallis tests refuse the H0 (null hypothesis = the means are equal).

utilisation of the electronic questionnaire implies (please see above how this is carried out) lesser errors because of interactive checks included in the electronic sheet (Excel). Moreover the elimination of paper questionnaire recordings on magnetic tapes dramatically reduces the errors related to replication and typing. All this implies fewer errors and checks in the following validation steps.

5.5.2 Costs

The economic savings due to the electronic questionnaire is valuable with reference to the Institute outsourced services. In particular questionnaire printing, personalisation and mailing and moreover the data recording on magnetic tape. The savings related to the questionnaire set up originate from the elimination of printing and sending steps.

The costs of the questionnaire set up step consisted of printing, mailing and also other costs related to the post-free return envelopment. These costs had to be multiplied for all sending carried out during the data collection period: first sending and one call-back with the questionnaire attached done before the web-mode implementation; this implies a duplication of these kinds of costs. The web-mode collected questionnaires were already on a server so the outsourced magnetic recording of paper questionnaires (and the related cost) was eliminated. It has to be highlighted that the electronic questionnaire management implies the sending of a new notice in order to inform enterprises of the survey launch, so the costs for this mailing have to be added to total costs. We must also add the costs for the software development and for the Excel set up of the questionnaire. Anyway, software related costs have a very long amortization period. The costs for web site and server set up are not considered because of the very little incidence of the SCI survey share on the total Istat web-managed surveys (now over thirty); the SCI share for these costs is not very significant. It is also necessary to consider the surplus, in terms of human resources redirected from low level operations to higher productivity tasks. The costs described above represent different shares on the total; Graph 5.4 presents total cost distributions.

**Graph 5.4 - Costs share distribution on total costs**
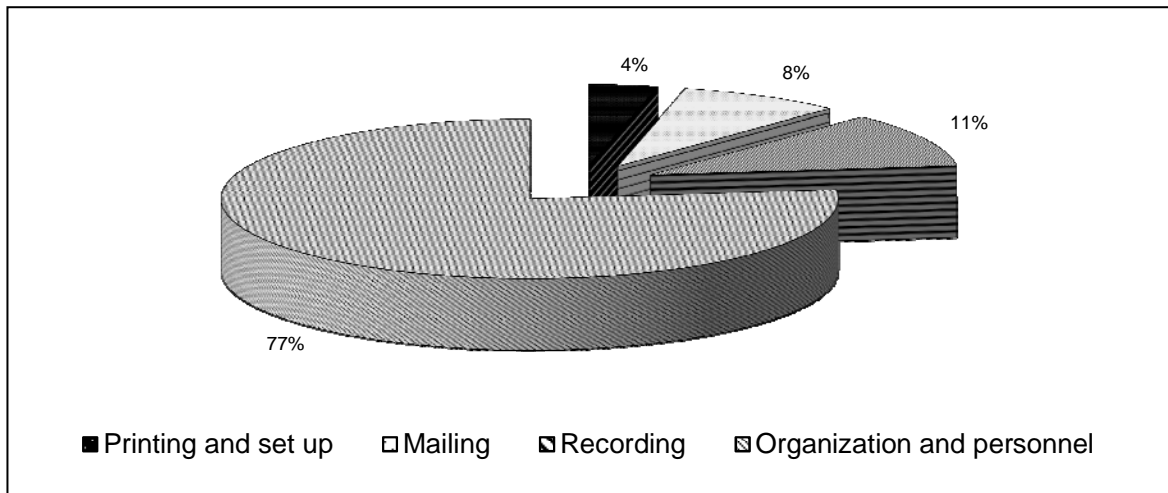
The sum of above mentioned expense items gives the total saving, in particular:
- printing, set up and mailing related to first sending and questionnaire attached call-back;
- new personnel organisation;
- free-tax return envelopment;
- recording of magnetic tape;
- new mailing costs;
- software implementation and electronic questionnaire set up.

Now SCI is a fully implemented web survey and we obtained great savings. An estimation is supplied in Graph 5.5.

Graph 5.5 - Year savings estimation by expense items under the full web managed SCI survey *(percentage values)*
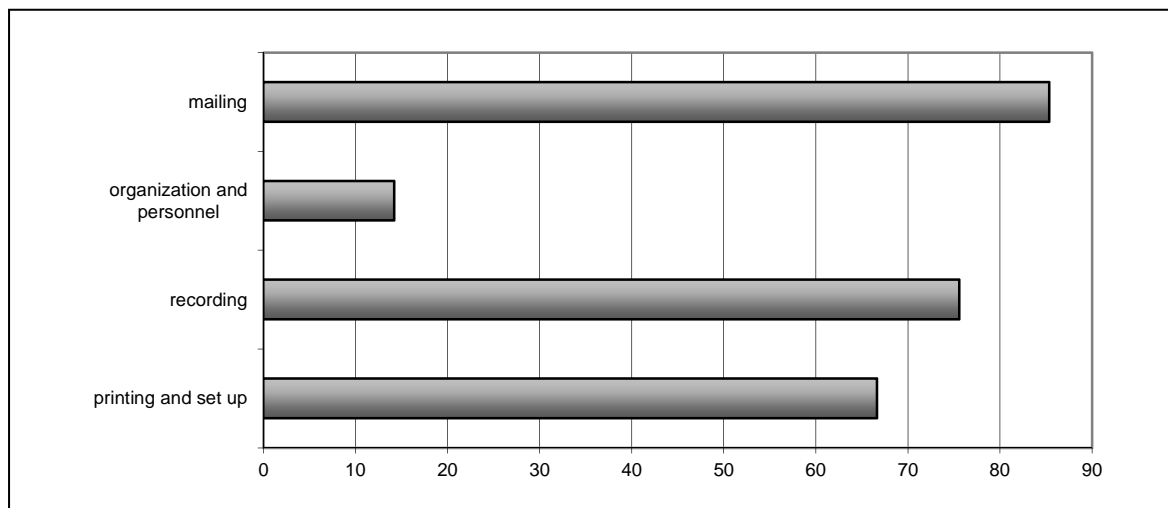
Graph 5.5 presents the costs percentage reduction on the total costs for each expense component for the fully implemented web mode survey and for all enterprises included in the SCI target population (i.e. no more paper questionnaires).[6]


6. Conclusions

The experience gained in these three years of SCI survey management regarding web mode questionnaires can be ranked as very positive. Almost all parameters considered in order to detect the management efficiency and the data quality are clearly in favour of the web mode questionnaire compared with the paper mode. These facts confirm the results obtained by other studies, that data collection via the web is more efficient if some conditions are verified, such as: the Internet diffusion in particular target populations, the implementation of standards oriented towards web mode survey participation maximisation and, at the same time, costs minimisation for respondents (Crawford S., McCabe S., Couper M., Boyd C., 2002).

Graph 6.1 synthesizes the effects on the quality of electronic questionnaire adoption according to the main aspects addressed in the present paper. The presented values are percentage change estimations in respect to the initial situation, the estimations are elaborated by means of the same indicators utilised in paragraph 4 in order to detect the impact on costs of web questionnaires and therefore:

- cost savings on total costs;
- enterprises burden in terms of time of questionnaire compilation;
- timeliness calculated as percentage difference of arrived questionnaires at the same date;
- accuracy calculated as percentage distance between row data and validated values of main economic variables;
- response rate in terms of percentage of respondent enterprises on total concerned enterprises.

---

[6] Since 2006 SCI survey is carried out in web mode only.

Graph 6.1 - Percentage quality changes in respect to initial situation according to utilised criteria *(percentage values)*
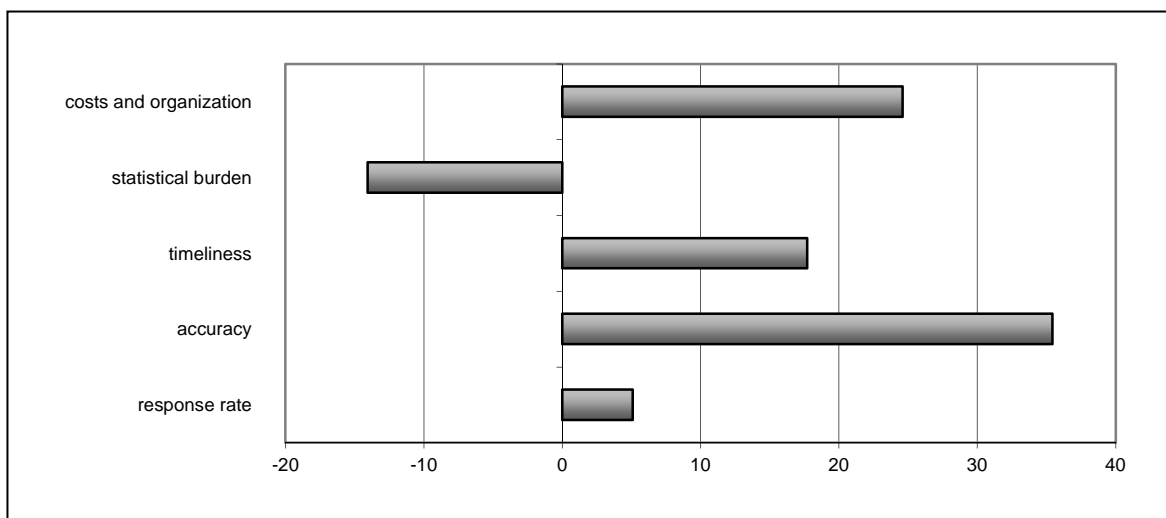
The data shown in Graph 6.1 highlight some controversial aspects such as for instance: the response rate does not seem influenced by the electronic questionnaire and the enterprises statistical burden, in terms of time necessary for compilation. While the response rate in 2005 appears to have risen with regards to enterprises burden further work is necessary in order to make the electronic questionnaire more user friendly (Dillman *et al*, 1999). Last but not least: improvements of utilised technology will be necessary.

Savings obtained by web mode questionnaires are now an achieved goal; in fact the costs of web mode survey do not increase in a linear way as happened for a large part of postal surveys. The great part of survey costs via web is constant and they depend only on questionnaire length and survey complexity. The results in terms of survey costs savings are in line with the conclusions of other surveys (Crawford S., McCabe S., Couper M., Boyd C., 2002), where the costs comparison is even more unfavourable towards paper mode survey, costing approximately 222% more then web survey.

In order to list the reasons for the SCI survey web mode achievements, the following aspects must be highlighted:

1. enterprise awareness of SCI electronic questionnaires channelled by explorative postal survey on the enterprises willingness to fill in an electronic questionnaire carried out before the SCI 2004 survey launch;
2. the utilisation of simple and very widespread software;
3. the simple graphic interface and respondent-friendly design (very similar to the old paper questionnaire);
4. safety and confidentiality granted both in data transmission and in data treatment.

In 2006 having the electronic share went over the threshold of two-thirds of total respondents in the previous year it was decided to abandon the paper mode. From 2006 onwards the respondent can utilise the paper questionnaire only on request, this cause a dramatic crash in paper mode respondents share (reaching the value of 1 per cent in 2010) as shown in Graph. 6.2.

Graph 6.2 - Paper mode respondents percentage on total respondent - SCI surevy - Years 2006-2010
(percentage values)
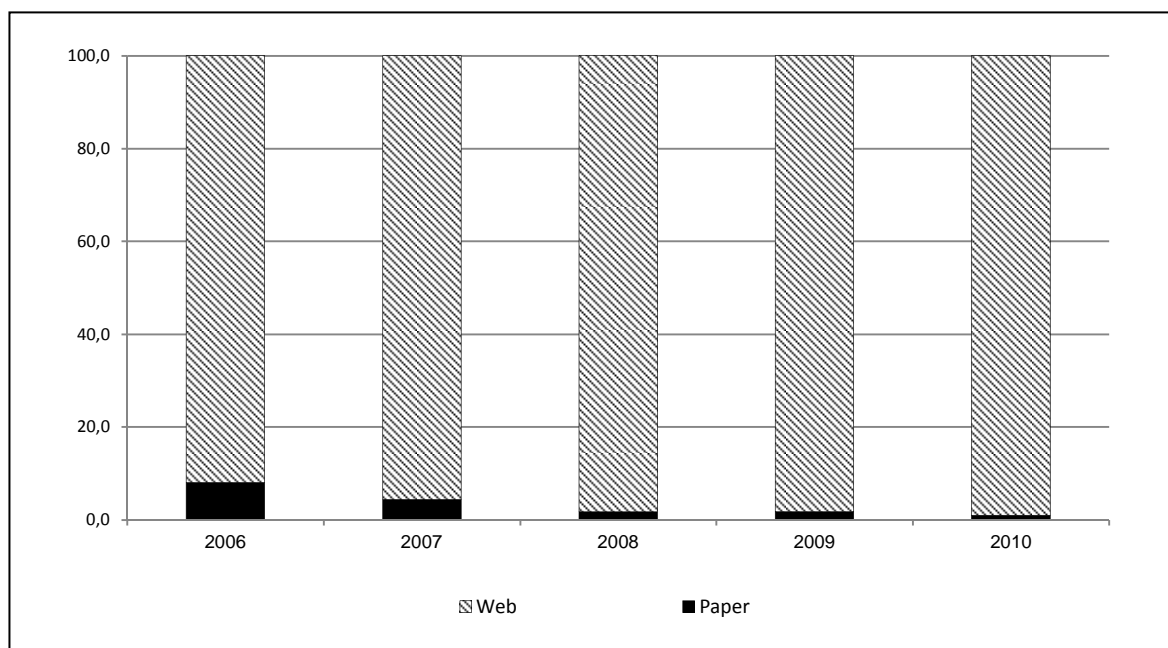
These results support the decision of abandon of the paper questionnaire and clearly confirm that firms were ready to pass to electronic mode.The analysis of variables that determine the choice of web mode response as an alternative to paper mode response could permit a more focused management of this response mode and incline the paper mode enterprises toward electronic questionnaires. It is necessary to point out the requirement of a further exploration into this matter in order to highlight the different respondent behaviours.

References

Balestrino, R. 1998. Data Capturing on the Web Proceedings of the Seminar NTTS'98, Sorrento.

Balestrino R., Barcaroli, G. 1998. *The introduction of CASIC Technologies in an Institute Producing Official Statistics* Documenti Istat, n. 3.

Biffignandi S., Pratesi M. 2000. *Le indagini via Internet sulle imprese: aspetti metodologici e un'analisi dei rispondenti* in Tecnologie informatiche e fonti amministrative nella formazione dei dati economici, F.Angeli, Milano.

Biffignandi S., Fabrizi E. 2006. Mixing web and paper based data collection modes in a survey on manufacturing firms Proceedings of Q2006 European Conference on Quality in Survey Statistics, Cardiff 24 - 26 April.

Cobanoglu C., Moreo P. J, Warde B. 2001. A comparison of mail, fax and web-based survey methods. *International Journal of Market Research*, Vol. 43, No. 4.

Couper M.P. 2002. New Technologies and Survey Data Collection: Challenges and Opportunities International Conference on Improving Survey – Copenhagen.

Crawford S., McCabe S., Couper M., Boyd C. 2002. From Mail to Web: Improving Response Rate and Data Collection Efficiencies International Conference on Improving Survey – Copenhagen.

Dillman D.A, Tortora R.D., Bowker D. 1999. Priciples for costructing WEB survey http://survey.sesrc.wsu.edu/dillman/papers/websurveyppr.pdf.

Dillman D.A, Phelps G., Tortora R.D., Swift K., Korhell J. Berck J. 2001. Response Rate and Measurement differences in Mixed Mode Surveys http://survey.sesrc.wsu.edu/dillman/papers/mixed%20mode%20ppr%20_with%20Gallup_%20POQ.pdf.

Meckel M., Walters D., Baugh P. 2005. Mixed-mode Survey Using Mail and Web Questionnaires *The Electronic Journal of Business Research Methodology*.

ISTAT 2007. *Conti economici delle imprese – Anno 2003* Collana Informazioni, n.8.

ONS 2005. Web versus Paper Questionnaires: A Design and Functionality Comparison. UN Conference of European Statisticians – Ginevra

Tse A. C. B. 1998. Comparing Response Rate, Response Speed and Response Quality of Two Methods of Sending Questionnaires: E-mail vs. Mail. *International Journal of Market Research*, Vol. 40, No. 4.

Zeli A. 2006. Fasi di allestimento, invio e raccolta dei questionari per la Rilevazione sul Sistema dei Conti Economici delle Imprese (SCI). Sviluppo e primi risultati dell'introduzione del questionario elettronico *Mimeo - Istat*.

Informazioni per gli autori

La collana è aperta ad autori dell'Istat e del Sistema statistico nazionale, e ad altri studiosi che abbiano partecipato ad attività promosse dal Sistan (convegni, seminari, gruppi di lavoro, ecc.). Da gennaio 2011 essa sostituirà Documenti Istat e Contributi Istat.

Coloro che desiderano pubblicare sulla nuova collana dovranno sottoporre il proprio contributo alla redazione degli Istat Working Papers inviandolo per posta elettronica all'indirizzo iwp@istat.it. Il saggio deve essere redatto seguendo gli standard editoriali previsti, corredato di un sommario in italiano e in inglese; deve, altresì, essere accompagnato da una dichiarazione di paternità dell'opera. Per la stesura del testo occorre seguire le indicazioni presenti nel foglio di stile, con le citazioni e i riferimenti bibliografici redatti secondo il protocollo internazionale "Autore-Data" del *Chicago Manual of Style*.

Per gli autori Istat, la sottomissione dei lavori deve essere accompagnata da una mail del proprio dirigente di Servizio/Struttura, che ne assicura la presa visione. Per gli autori degli altri enti del Sistan la trasmissione avviene attraverso il responsabile dell'ufficio di statistica, che ne prende visione. Per tutti gli altri autori, esterni all'Istat e al Sistan, non è necessaria alcuna presa visione. Tutti i lavori saranno sottoposti al Comitato di redazione, che valuterà la significatività del lavoro per il progresso dell'attività statistica istituzionale. La pubblicazione sarà disponibile su formato digitale e sarà consultabile on line.

Gli articoli pubblicati impegnano esclusivamente gli autori, le opinioni espresse non implicano alcuna responsabilità da parte dell'Istat. Si autorizza la riproduzione a fini non commerciali e con citazione della fonte.