ISTAT

LAUNCHING AND IMPLEMENTING THE JOB VACANCY STATISTICS

Authors
Ciro Baldi[1], Diego Bellisai[1], Stefania Fivizzani[1],
Annalisa Lucarelli[1] and Marina Sorrentino[1]

[1] ISTAT - Servizio Statistiche congiunturali sull'occupazione e sui redditi

AUTHORSHIP

This paper is the result of joint work. However, Sections 1 and 9 have been jointly written by Ciro Baldi, Diego Bellisai and Marina Sorrentino; Section 2 has been written by Marina Sorrentino; Sections 3 and 4 have been written by Stefania Fivizzani; Section 5 has been written by Diego Bellisai; Sections 6, 7 and 8.2 have been written by Ciro Baldi; and Section 8.1 has been written by Annalisa Lucarelli.

ACKNOWLEDGEMENT

The authors would like to sincerely thank Gian Paolo Oneto, Leonello Tronti and Roberto Gismondi for their advice, suggestions, and comments.

ABSTRACT

Istat has been running a quarterly business survey on job vacancies since the 3rd quarter of 2003. The survey is designed as to produce quarterly estimates of job vacancies and occupied posts by Nace Rev. 1.1 sections, as required by the gentlemen's agreement between Eurostat and the EU Member States on job vacancy statistics.

This report describes the first four waves of the survey.

In particular, the following aspects are considered: the sample design, the target population and its stratification, the sample allocation and drawing procedures; the questionnaire, its pre-test and the main consistency checks; the techniques used to contact the sample units and to collect data; the information flows and the survey database; the check and editing procedures; the estimator used to gross up the sample values to the target population and its variance.

Furthermore, some indicators of data quality are calculated: the incidence of consistency checks violations on the collected data, the percentage relative sampling errors of the estimates of job vacancies, occupied posts and job vacancy rates and the efficiency of the sampling design and estimator.

CONTENTS

1. Introduction
2. Sample design
3. Questionnaire design
4. Contact and data collection techniques
4.1 A first assessment of data quality with respect to the survey technique
5. Information flows
5.1 Enterprises' data management
6. Check and editing procedures
6.1 Outlier identification on job vacancies
6.2 Imputation of job vacancies item non responses
7. Estimation procedure
8. Indicators of data quality
8.1 Violations of consistency rules
8.2 Evaluation of the estimates quality
9. Conclusions
References
Annex 1 – Quarterly job vacancy survey questionnaire

List of tables

Table 1 - Contacted enterprises by contact mode (III quarter 2003 – II quarter 2004)
Table 2 - Contacted enterprises by requested mode of data transmission (III quarter 2003 – II quarter 2004)
Table 3 - Respondents by mode of data transmission (III quarter 2003 – II quarter 2004)
Table 4 - Consistency checks violations (share of total respondents, III quarter 2003 – II quarter 2004)
Table 5 - Sampling errors on job vacancies and occupied posts (estimates of percentage relative sampling errors, III quarter 2003 – II quarter 2004)
Table 6 - Deft on job vacancies and occupied posts (III quarter 2003 – II quarter 2004)

1. Introduction

Istat has been running a quarterly business survey on job vacancies[2] since the 3rd quarter of 2003. The survey is designed as to produce quarterly estimates of job vacancies and occupied posts by Nace Rev. 1.1 sections, as required by the gentlemen's agreement between Eurostat and the EU Member States on job vacancy statistics.

This report describes the design, the operational details and results of the first four waves of the survey, providing the estimates of job vacancies and job vacancy rates obtained.

In particular, Section 2 describes the sample design, and presents the target population, its stratification, the sample allocation and drawing procedures.

The questionnaire, its pre-test and the main consistency checks run on the various questions are described in Section 3.

Section 4 discusses the characteristics of the techniques used to contact the sample units and to collect data, the reasons behind their choice, and the relevance of each of them in the first four survey waves.

The information flows, the characteristics of the survey database, and the procedures to update the sample enterprises' identification details and keep track of all their changes over time are presented in Section 5.

The procedures to check and edit the sample data that were used in the first four waves (including mainly the identification of outliers and the imputation of item non responses on job vacancies) are then discussed in Section 6.

The estimator used to gross up the sample values to the target population and its variance are described in Section 7.

Section 8 presents some indicators of data quality: the incidence of consistency checks violations on the collected data, the percentage relative sampling errors of the estimates of job vacancies, occupied posts and job vacancy rates and the efficiency of the sampling design and estimator.

Section 9 concludes the report.

2. Sample design

The observation unit of the job vacancy survey is the enterprise, as defined in the Council Regulation (EEC) No 696/93 of 15 March 1993 on the statistical units for the observation and analysis of the production system in the Community.

[2] It is worth mentioning that the same survey also serves to collect data on hours worked and paid.

The survey sample is composed of around 7,800 enterprises, selected among those of a target population defined as the set of enterprises with at least 10 persons employed, whose economic activity is classified in NACE Rev. 1.1 sections C to K and which were active in the reference quarter.

As a frame, the most recent release of Istat statistical business register, ASIA, that is that relative to 2001 for the first 4 waves of the survey, is considered.

The sample design is a one stage stratified one.

The economic activity of the enterprise, its size in terms of the number of persons employed, and the geographical area of its head office are the stratification variables, since they are correlated with those of interest and available for all enterprises in the target population. On the whole, 115 strata are defined, on the basis of the following classes:

- 12 economic activity classes: NACE Rev. 1.1 sections C and E to K, and 4 classes identified in NACE Rev. 1.1 section D according to a Pavitt-like classification;
- 4 size classes: 10 to 19 persons employed, 20 to 99, 100 to 499 and at least 500;
- 3 geographical areas: North, Centre, South (including Sardinia and Sicily).

Contiguous strata containing less than 100 enterprises employing less than 500 people have been collapsed.

All enterprises with at least 500 persons employed (around 1,200) are selected with probability one from the 34 strata concerned.

The remaining 6,600 enterprises with less than 500 persons employed are selected with inclusion probabilities strictly smaller than one from a frame population of around 189,000 enterprises (that is with an average sampling fraction around 3.5%), with a simple random sampling without replacement scheme in each of the 81 strata concerned.

The 6,600 sample units employing less than 500 persons have been allocated in the 81 strata with the Bethel method, that is by minimising a linear cost function subject to the constraint of given coefficient of variation for the interest variables in the study domains (including the NACE sections C-K and the entire population).

The sample allocation procedure requires preliminary estimates of means and standard deviations of the target variables in each stratum.

The only source of these preliminary estimates for similarly defined job vacancies for a sufficiently large subset of the target population was the pilot survey on the same variables, run by Istat in spring 2002 (see the final report for Eurostat Contract n. 200032100024, Istat, 2002). However, not only the respondents of this survey were a relatively small number (just above 1,900),

but the target population did not include enterprises with 10-19 persons employed or those classified in NACE Rev. 1.1 section J.

Therefore, the sample allocation procedure has been run considering only occupied posts as target variable.

However, we have recently revised the target population definition, as well as the sample design and allocation on the basis of the data collected in the first 5 quarterly survey waves. The new design is being applied starting from the wave for the 4th quarter 2004 and considers also the job vacancy rate as an interest variable in the sample allocation procedure.

The simple random sampling scheme within each stratum was carried out with a technique based on permanent random numbers and designed to consider also a negative coordination objective with respect to a set of annual business surveys. To this aim, eight possible samples were drawn and the actual sample was selected as the one for which the response burden due to the overlapping with specific annual business surveys was minimum.

A set of enterprises was drawn larger than the theoretical sample size, so as to create a reserve pool to be utilised to contain the negative effects of unit non responses and business demography. In the first wave of the survey (that is, for the 3rd quarter of 2003) the reserve pool was quite large (around 9,800 enterprises), but it became rapidly smaller, even if it was updated in the second and third waves (when it included, respectively, around 2,050 and 700 enterprises).

3. Questionnaire design

The first section of the questionnaire (see Annex 1) contains the data identifying the sample unit and is filled in by Istat in advance with the information contained in the register, but includes fields to allow the respondent to modify such information if outdated or incorrect.

It also includes a subsection with the contact data of the person responsible for filling in the questionnaire, in order to facilitate the contact and response in following waves.

The second section aims at collecting information on the number of employees working in the enterprise in the last day of the previous quarter and in the last day of the reference quarter, and on the number of employees who have started and stopped working in the enterprise during the reference quarter. These data are required separately for manual and non manual workers, a disaggregation that was not required in the pilot survey questionnaire, while information about managers is excluded. The data on employees working in the enterprise in the last day of the reference quarter is the basis for the estimates on occupied posts transmitted to Eurostat to be used as the denominator of the job vacancy rate.

Job vacancies are the subject of the third section of the questionnaire.

Whilst in the pilot survey there was a single direct question aimed at measuring job vacancies, in the quarterly survey we have introduced two preliminary questions before the one measuring this variable, with the purpose to help focus the respondent's attention on the conditions defining the variable.

Thus, the first question of the section asks whether in the last day of the reference quarter the enterprise was looking for people to hire from outside the enterprise itself.

If the answer is yes, a further question is asked as a second filter, that is whether in the last day of the reference quarter the enterprise had already taken active steps to find and hire suitable candidates (for example by advertising on mass media or internet, requests to public or private employment agencies, word of mouth with other employers, collection of job applications, interviews of prospective candidates, etc.).

Only if also the answer to this second question is positive, we ask for how many posts the enterprise had already taken active steps to find and hire suitable candidates and was ready to take further such steps, if needed.

Job vacancies are split into those for manual and non manual workers and exclude those for managerial positions.

The section includes two further questions, concerning immediately available and hard-to-fill job vacancies.

The final section of the questionnaire aims at measuring hours actually worked as normal time and overtime, hours paid by the employer but not worked, hours not worked but paid for by the Wage Supplementation Fund ("Cassa Integrazione Guadagni") allowing for temporary layoffs , and hours not worked for labour disputes.

To contain the response burden and facilitate the communication between Istat and respondent, there is a final question asking the enterprises to indicate whether they would prefer the web, fax or e-mail mode to communicate with Istat for this survey.

As already indicated, the questionnaire designed for the quarterly survey differs from the one used in the pilot version. Hence, it was tested via a small number of face-to-face interviews carried out by Istat statisticians who had worked both on the pilot survey and the design of the quarterly survey questionnaire.

The new structure for the questions on job vacancies was found to be more easily comprehensible for enterprises than the one utilised in the pilot survey.

Furthermore, on the whole, the questionnaire was considered comparatively simpler and the collection of the information asked was said to require generally a manageable amount of time.

However, some of the interviewed respondents indicated that they would have at their disposal the information asked in the questionnaire only 10-15 days after the end of the reference quarter. If this circumstance was common to a large number of enterprises, it could be reasonable to delay of 7-10 days after the end of the reference quarter the starting date of the data collection phase.

Enterprises can transmit the data by filling in a questionnaire in a specially designed internet site, via e-mail or via CATI. In the wave for the 1st quarter of 2004, also the fax option was given. All these techniques and their different roles in the various waves will be described in the following Section.

Consistency checks were designed for the CATI and web-based versions of the questionnaire.

In both cases, a violation of consistency checks is signalled via a warning message, explaining it. However, the remaining part of the questionnaire can be filled in also if the data which caused the violation are not modified, so as to avoid a negative effect of the violation of consistency checks on the response rates.

The consistency checks implemented in the CATI version of the questionnaire are mostly based on comparisons between answers supplied by a given enterprise in a single wave to different questions of the questionnaire (that is they are within record checks).

Consistency checks on employment aim at verifying that:

- figures which should be subsets of others actually satisfy these constraints (for example, part-time non manual workers in a given enterprise and at a given date cannot be more than all non manual workers in the same enterprise and in the same date);
- the geographical distribution of manual and non manual workers in the last day of quarter t is consistent with the aggregate figures also asked in the questionnaire;
- data on employees working in the enterprise in the last day of quarters t and (t-1) are consistent with those on people whose job contracts in the enterprise started and ended in quarter t.

Moreover, there is a check based on data supplied by the same enterprise in two subsequent quarters: data on employment in the last day of the quarter (t-1) collected in the questionnaire for quarter t for a given enterprise are compared during the CATI with the answer to the same question supplied by the same enterprise in the questionnaire on quarter (t-1), if available.

Furthermore, both in the employment and in the job vacancy section, there are routine checks, aiming at verifying that the sequences of answer to different questions are internally consistent. For example, enterprises that answered that they were not looking for people to hire from outside the enterprise at the end of quarter t, or that they had not already taken active steps to find and hire suitable candidates, should not answer to the question asking for how many posts the enterprise had

already taken active steps to find and hire suitable candidates at the end of the same quarter and were ready to take further such steps, if needed.

In the job vacancy section, there are also checks to verify that:

- if an enterprise had already taken active steps to find and hire suitable candidates at the end of quarter t, then there was at least a post for which these active steps had been taken (that is a job vacancy);
- the number of immediately available job vacancies for a given job position was not larger than the total number of job vacancies for that job position;
- the number of hard-to-fill job vacancies for a given job position was not larger than the total number of job vacancies for that job position.

Finally, special attention is paid during the CATI to the questions in the job vacancy section if the interviewed enterprise is classified in labour-contracting activities (corresponding to the Ateco 2002 code 74.50.2, and included in the Nace Rev. 1.1 class 74.50) and in the considered quarter is reporting high turnover rates. This peculiar treatment has been designed to cope with the high turnover and hence, presumably, also high job vacancy rates characterising these activities, and the large size of some of the enterprises belonging to this activity class.

Further checks are implemented in the last section of the questionnaire, on worked hours.

To prevent the discouragement of web-based and e-mail respondents, the consistency checks implemented in these versions of the questionnaire represent a very small subset of those in the CATI version: the only checks included in employment section are those verifying that employment variables which should be subsets of others actually satisfy these constraints; the job vacancy section has no consistency check; and there is only one further check in the section on worked hours.

4. Contact and data collection techniques

The techniques used to contact the sample units and to collect data changed during the first four waves of the quarterly survey, because they had to be adapted to the effective availability of resources dedicated to the survey itself. In the first two waves (for the last two quarters of 2003) the data collection was entirely run with the CATI technique. The wave for the 1st quarter of 2004 was carried out with a mix of techniques, including the submission via fax of paper questionnaires as the main one, as planned in the original programme of the survey. Afterwards, since the wave for the 2nd quarter of 2004, we reverted to a mix of techniques including the CATI mode as the main one, but also the web-based questionnaire submission. This mix of techniques is the one we are still

using and that we plan to use at least in the next few years, due to the problems met with the third wave and given the resources currently dedicated to the survey.

According to the original plan on the survey techniques, the first two waves would have been entirely run via the CATI technique, while each of the subsequent ones would have been run via CATI only for the enterprises which entered the sample in that wave or the previous one. The remaining enterprises (which should have been the large majority in each wave) could have responded using one of a set of possible methods, including fax, web and e-mail.

The original choice to use only the CATI technique in the first two waves was made to exploit its advantages in this survey, and in particular its suitability to: correctly identify the sample unit and, inside it, the person responsible for actually supplying the data; check the adequate comprehension by the respondents of the survey questionnaire; verify the willingness to respond of the sample units and, if required, implement strategies to increase it.

In the first wave, the survey material (including the questionnaire, the letter announcing and briefly explaining the survey and the guide for a correct answering to the questionnaire) was sent by mail to the enterprises included in the sample and the reserve pool. However, in all the following waves, starting from the second one, the material was sent by mail only to the enterprises which had not already supplied a fax number or an e-mail address. The number of contacted enterprises (including the theoretical sample and the reserve pool) was 17,600 in the first wave, but substantially lower, around 9,850, in the second one. All enterprises, including those that would have transmitted the data via CATI, were required to answer to the questions on their own, that is, in the case of a CATI transmission, before the interview. The rationale is that the data asked in the questionnaire in general cannot be calculated in the short time of a CATI interview.

In each of the four waves considered in this report (from the 3rd quarter of 2003 to the 2nd of 2004) the CATI was run by an external private provider (COS-Infotec is the name of the company) purposely selected by Istat via a nationwide procurement procedure.

The CATI interviewers were trained in specific sessions before each wave (to train those who were not present in preceding waves, and to underline specific aspects for those already trained). These training sessions concerned the survey, its aims, the enterprises involved, the questions asked, the consistency checks, the behaviour during the interviews and more technical aspects, such as the software implementing the CATI version of the questionnaire, the way to navigate inside it and to manage the consistency checks. They were held both by Istat statisticians and (for the more technical aspects)by the people responsible of the CATI interviews in the company providing the service.

Istat researchers were available along the entire CATI phase for any clarification or further information that might have been useful to the CATI interviewers and they supervised and monitored consistently the data collection. Furthermore, Istat researchers answered to telephone calls by enterprises (to the numbers indicated in the letter sent together with the questionnaire) and supplied to them all the clarifications and information required (mostly on the questions asked in the questionnaire, how to submit the data, the length of the period during which a given enterprise would have been involved in the survey).

Besides the training of the interviewers, another critical aspect of a CATI data collection technique concerns the telephone numbers of the sample and reserve pool enterprises. The availability of the correct phone numbers of the people inside the sample and reserve pool enterprises who are responsible for answering to the questionnaire has a great impact both on unit response rates and timeliness. However, this condition in general can be fulfilled only for enterprises which were respondents in previous waves of the survey, not necessarily via CATI, provided that, if respondents by other means, they supplied the name and telephone number of the reference person in the first page of previous quarters' questionnaires. On the contrary, when an enterprise first enters the pool of those that should respond to the survey (and, in particular, in the first wave of the survey, all the enterprises to which the survey material was sent), the telephone number of the person responsible for answering to the questionnaire is not available, because not present in the register. Therefore a telephone number of these enterprises has to be found from other sources.

For this survey, the first source of telephone numbers for enterprises which have never previously responded is the latest release of the SEAT database of business phone numbers available in Istat (generally updated to a later date than that to which the register refers to).

The link between the register (and thus also the sample and reserve pool drawn from it) and the SEAT database is the taxpayer's identification code. Not all the enterprises in the register have one or more corresponding records in the SEAT database. We found that we could not match at least one third of the enterprises of the sample and the reserve pool of the survey with any record in the SEAT database. Furthermore, for each enterprise included in the SEAT database there can be from one to a few hundreds telephone numbers. During the first survey wave, we tested on a small subsample of a few tens of enterprises that a telephone number from which we could obtain that of the head of the human resources office (the person we believe should be the best informed one in the enterprise on the information asked in the questionnaire) could be found by calling the first few of the numbers listed for each enterprise in the SEAT database. Hence we decided to select, for each enterprise in the sample and reserve pool that could be matched with at least one record in the

SEAT database, the first few telephone numbers listed in the database (initially 10, in later waves 5-7).

These are the telephone numbers which were provided together with the other identification details to the supplier of the CATI services for each enterprise to be contacted.

The company which ran the CATI searched for the telephone numbers of the sample and reserve pool enterprises which could not be matched with those in the SEAT database and for more accurate telephone numbers for the enterprises for which those supplied by Istat did not allow to contact the reference persons. This further search was made on other publicly available telephone number databases.

As anticipated, in the wave for the 1st quarter of 2004 (that is the 3rd wave of the survey) the respondents were allowed to choose among a variety of data transmission methods: web-based data submission, e-mail, CATI for 997 enterprises which had never responded before to the survey, and fax. This last option was coupled with the utilisation of manual data entry service provided by a private company selected for the purpose by Istat. Letters of recall were sent by fax and e-mail to the enterprises which had not already responded within a given date, and 1,000 of them also received follow-up phone calls. Furthermore, a toll free phone number was activated to support enterprises in responding to the questionnaire and in the data transmission. Two e-mail addresses and a fax server number were also dedicated to the survey and employed exclusively for it. They are currently used by Istat to send questionnaires and the accompanying material, but can also be used by the enterprises to transmit data or to ask for information.

However, the management and coordination of the various phases of the data entry proved to be very complex. In particular, it was very difficult to adequately control the data quality both in the transmission by the respondents and in the data entry phases. Due to the lack of human resources available, the check of the number of data entry errors by visual comparison of the paper questionnaires and the recorded data could be run only on a sample of the entered data.

These circumstances, and the fact that the first two CATI waves had ensured excellent results, both in terms of number of responding enterprises and the quality of the collected data (see Section 8.1 below), led to the decision to revert to the CATI technique of data transmission for all the enterprises which do not respond via web or e-mail.

Hence, since the wave for the 2nd quarter of 2004 the survey has been carried out with a hybrid data collection technique, including CATI, web and e-mail. The last two methods are dedicated especially to the enterprises which explicitly chose them (in the response to the last question of the questionnaire) and to those which have used the web mode at least once. The CATI mode concerns

all enterprises which are required to respond by no more than two quarters and to all the other enterprises in the sample and reserve pool. To all the enterprises which should respond via web or e-mail and which have not responded by a certain date, letters of recall are sent via fax and e-mail.

The web-based data submission is allowed by a purposely built web site, whose address is http://indata.istat.it/vela, where the enterprises can find, beside the questionnaire to fill in (in Adobe Acrobat PDF format), information on the survey and the guide to answering the questionnaire. The access to the pages where the enterprises can fill in the questionnaire and transmit the data is possible only via a login and a password authentication. The login and a temporary password are supplied by Istat, together with the questionnaire. For security and privacy reasons, each enterprise needs to register and change the temporary password before accessing for the first time the web-based data transmission pages.

In the e-mail data transmission, Istat sends via e-mail a questionnaire (in Adobe Acrobat PDF format) where the identification details of the enterprise are already present. The enterprise fills it in with the remaining data (and with changes to the identification details, if needed) and sends it directly to the survey database (see also Section 5).

No evidence of particular difficulties has arisen with respect to the e-mail and web data transmission. The areas in which relatively more frequent problems have been met are the following:

- forgotten or lost password for the web-based data transmission (in such cases the enterprises need to call Istat that cancels their previous registration and allows them to re-register with the original temporary password);
- problems in viewing the web-based version of the questionnaire due to different or outdated software releases (for what concerns both browsers and Acrobat Reader) or stringent security policies applied to protect some enterprises' computer network;
- errors in implementing the data submission procedure by the enterprises, resulting in no data being transmitted.

4.1 A first assessment of data quality with respect to the survey technique

The following three tables supply information on the relevance in the four waves of the techniques adopted to contact the sample and reserve pool enterprises, those in which enterprises where asked to respond and, finally, those actually used by enterprises to transmit the data.

Table 1 - Contacted enterprises by contact mode (III quarter 2003 – II quarter 2004)

Quarter	Fax	Email	Mail	Total
III 2003	0	0	17600	17600
IV 2003	4016	2676	3162	9854
I 2004	4542	2726	1170	8438
II 2004	4252	2766	930	7948

Table 2 - Contacted enterprises by requested mode of data transmission (III quarter 2003 – II quarter 2004)

Quarter	CATI	Web	Unspecified	Total
III 2003	17600	0	0	17600
IV 2003	9854	0	0	9854
I 2004	997	0	7441	8438
II 2004	6114	1834	0	7948

Table 3 - Respondents by mode of data transmission (III quarter 2003 – II quarter 2004)

Quarter	CATI	Web	Other	Total
III 2003	8120	0	0	8120
IV 2003	7493	0	0	7493
I 2004	520	1850	3837	6207
II 2004	6179	774	0	6953

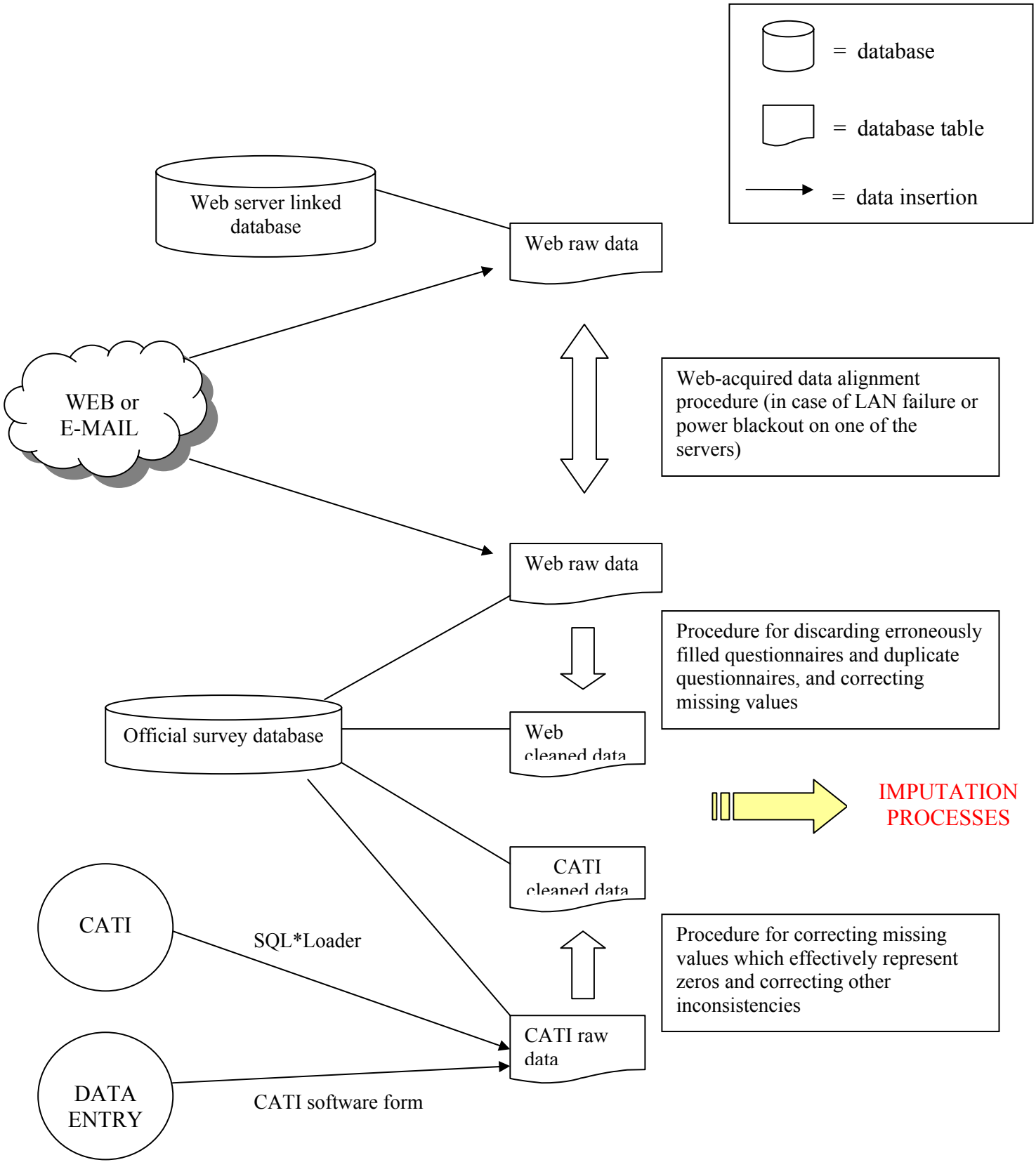
The contact mode composition varied according to the percentage of enterprises which were contacted for the first time in each wave. As already indicated, the mail mode is used only for the enterprises for which no other contact detail is yet available: hence its relevance decreased with the increase in the share of enterprises which had already responded. In the last two waves considered, the fax mode was used to contact more than 50% of enterprises, while the e-mail was used in around one third of cases.

The evolution of the data transmission modes mix can be grasped from Table 2: the CATI mode was chosen for all enterprises in the first two waves, only for around 12% of them in the third, but for more than 75% of them in the fourth. As already indicated, in the third wave, all enterprises which were not contacted via CATI could choose the data transmission mode. Finally, in the fourth wave, the enterprises which had responded via web (or by e-mail) in the previous wave were asked to respond via the same mode.

The parallel evolution can be assessed in terms of respondents from Table 3. The web-based and e-mail respondents were 30% of all respondents in the first wave when this option was available. However, they declined substantially in the following wave (but increased again in the subsequent waves, not considered in this report), possibly due to problems of forgotten or lost passwords.

The number of web and e-mail respondents in the 1st quarter of 2004 is larger than that of the enterprises asked to use this data transmission technique, because enterprises are suggested a data transmission technique, but not prevented to use the other available ones.

5. Information flows



As already mentioned in the previous paragraph, in the survey there are three main sources from which data are collected. In order of importance, they are:

1. CATI interviews, supplied on the basis of an established record structure
2. web-based questionnaire submission
3. e-mail submitted questionnaires.

There exists also a residual number of questionnaires which are received via the survey dedicated fax server and which are inserted into the survey database by data entry operators.

The survey database consists in an Oracle 9i DBMS based on a Linux Red Hat 9 operative system.

Data arriving from CATI interviews are loaded, via the Oracle SQL*Loader utility, on a daily basis into a table designed to contain raw data. This loading procedure fires a trigger PL/SQL procedure (i.e. a procedure triggered by an operation on the database objects, such as insertion, update, deletion), which copies the raw data into a new table and performs a first clean-up of the data on the final table.

This mainly consists in correcting some inconsistencies, such as

- setting missing values to zero in the cases in which this is supposed to be a typing mistake of the CATI interviewer. For example, missing hours actually worked for employees are set to zero if the average number of employees in the quarter is zero, both for the total number of employees and for part-time ones. This happens for each job position and also for the other hourly variables, including strike hours.

  This clean-up procedure is used also to correct missing values for the inflows and outflows of employees (hirings and separations), which by inspection are identified as zero values, and also to correct missing values for numbers of employees in a particular geographical area, when they are patently identified as actual zeros.
- setting the total number of employees equal to the declared total number of employees plus the declared number of part-time employees in the cases where the number of part-time employees is declared to be larger than the total number of employees. This correction accounts for a possible misunderstanding of the question by the firm's reference person, who, instead of giving the total number of employees, gave the number of full-time employees.
- setting the number of employees vacancies to zero when the number of employees at the end of the quarter is zero, while the firm declared to be recruiting and leaving the number of vacancies missing.

The next source of data collection in order of importance is the web questionnaire. As mentioned above, all firms taking part in the survey have been provided with a username (coinciding with the ASIA archive identification code) and a randomly generated temporary password with which they can connect to the survey's web site. On this site, they have the possibility of filling in questionnaires for all past quarters. The questionnaires are personalized in that the firm's identification code, name, address, economic activity code, description of the economic activity are pre-typed fields.

In the future, new web features are planned to be added which will allow the firm to consult the data it provided in all the previous quarters as well as selected vacancies and hourly indexes regarding their economic activity area.

After being filled in, the questionnaire is submitted by clicking on the "Submit" button, which launches a PHP script sending an e-mail containing the typed data in a particular format to the survey's official e-mail address. After few minutes, the Reader module of the Cardiff TeleForm software reads the e-mail and automatically insert the data in a table (different from that containing CATI acquired data) of the survey database. For the sake of safety (LAN interruption or power blackout), two parallel copies are made into two different databases managed by two different servers. After being inserted into the table, a PL/SQL procedure is called which moves to a discarded questionnaires table those which were erroneously left completely empty and, in case of multiple submissions by a firm for the same quarter, all but the most recent questionnaire.

The resulting "good" data are inserted into a new table and undergo the same clean-up procedure which applies to the CATI acquired data.

The e-mail source is quite analogous to the web, since a personalized questionnaire is sent on a quarterly basis to firms interested, which can fill it in and submit it by e-mail. Exactly as for the web, data are automatically inserted into the same table containing questionnaires filled in our web site and undergo the same clean-up treatment.

Finally, data received by fax are first manually inserted via the same software used by CATI operators. Should the data not pass the consistency checks implemented in the software, the firm is interviewed in order to confirm the data they typed in the questionnaire. Subsequently, these data are provided together with the CATI collected interviews and undergo the same process as the CATI loaded interviews.

5.1 Enterprises' data management

As far as enterprises' identification data are concerned, the questionnaire includes the possibility of communicating variations in the company name, address, taxpayer's identification number and economic activity.

In order to guarantee an effective management of the enterprises' identification data, these are kept in a specific table which also contains other information, such as fax number and e-mail address (when available), activity status (active, wound-up, out of the sample target, merged into a different company, broken up into two or more new companies and so on), sample stratum and expected date of entrance/exit from the sample.

At the moment there exist database PL/SQL procedures undergoing further development, which help managing these changes in identification codes and activity status, which are usually either hand-typed as written communications arrive or collected from the interviews loaded into the database. Moreover, information on enterprises' historical data are maintained via triggers which fire every time a change in the data table is made and which keep track of the kind of variation (update, insertion or deletion), of the date of variation and of whom inserted this variation.

Moreover, specific PL/SQL procedures are being developed which help to systematically correct typos in enterprises' identification data and identify possible inconsistencies in fax numbers and e-mail addresses of the reference persons.

As a matter of fact, this table is also used to create, via PL/SQL procedures, up-to-date lists of firms with a specific record structure, which are utilized for customizing the questionnaire and the letters to be sent to the enterprises (via fax or e-mail) in a neighbourhood of the last day of the reference quarter.

Finally, another PL/SQL procedure is being developed which automatically controls which enterprises have not answered yet and sends them a letter of recall by e-mail or fax where available.

6. Check and editing procedures

The check and editing procedure implemented so far are: the check on the number of violations of the editing rules built in the CATI and the web questionnaire; the check on outliers; the treatment of missing data..

While the results of the first procedure are described in the section dedicated to indicators of data quality (Section 8.1), this section presents the methods of outlier identification and imputation of item non responses.

Outlier and imputation procedures concern only data on job vacancies and they are carried out separately for manual and non manual workers. The absence of editing procedures on employment is justified on the basis that enterprises have less difficulties in reporting the number of employees and of the presence of tighter consistency checks in the data entry phase (see Section 3).

6.1 Outlier identification on job vacancies

The method of outlier identification can be considered a modified version of the Hidiroglou and Berthelot (1986), in the following HB, method (see Gismondi, 2000, 2002).

The HB approach is based on the comparison of the ratio between the number of job vacancies at time t and the same variable at time t-1 with the median value of this ratio. The procedure employed here is, instead, based on the distribution of the variable level and not on its quarter to quarter change because at this point in the development of the survey there is not enough knowledge on the relationship between the variable and its lagged value. Moreover since the target parameter is the level of job vacancies and not the change, the analysis of the cross section distribution can be more appropriate.

The outlier identification procedure is carried out only on the subset of firms that account for the 75% of employment (in terms of manual workers or non manual workers). The exclusion of the remaining firms is justified on the basis of a cost-benefit criterion, because the influence of an anomalous value on the final estimates is smaller the smaller is the firm.

The procedure is carried out within groups defined by the section of economic activity (or aggregation of sections where the number of cases is too small) and job position (manual and non manual workers).

In formulas the method is based on the following transformation of the job vacancy ratio of each firm:

$$[1] \quad y_{thi}=1+100*jvr_{thi}$$

where jvr_{thi} is the job vacancy ratio at time t, for the occupation h (h=1,2 that is manual or non manual workers) for the firm i that is:

$$[2] \quad jvr_{thi} = jv_{thi}/e_{thi}$$

where jv is the number of job vacancies and e is the number of employees.

The addition of 1 to the job vacancy ratio (expressed in percentage) prevents the median of the variable to be zero. A zero median is a very common result since a large part of firms declares to have no job vacancies.

The following transformation of y_{thi} assumes values not less than 1:

$$[3]\ s_{t,h,i} = \begin{cases} q_{0.50} / y_{t,h,i} & if \quad 0 < y_{t,h,i} < q_{0.50} \\ y_{t,h,i} / q_{0.50} & if \quad y_{t,h,i} \geq q_{0.50} \end{cases},$$

where $q_{0.50}$ is the median of the distribution of y_{thi}. The values jvr_{thi} whose transformations s_{thi} are higher than a threshold are classified as *potential outliers*. The threshold is defined as:

$$[4]\ A_{sup} = q_{0.50} + c_{sup} d_{sup},$$

where $d_{sup} = q_{0.75} - q_{0.50}$ is the interquartile difference and c_{sup} is an arbitrary parameter whose value has to be discretionarily chosen. On the basis of previous experience on other surveys this value has been fixed to 5. Finally, the potential outliers are classified as outliers if the job vacancy ratio is higher than 25%.

This methodology identifies about 20 to 30 outliers each quarter.

Once identified as an outlier the jv_{thi} is set as missing value and undergoes the procedure of imputation described in the following paragraph.

6.2 Imputation of job vacancies item non responses

Where the data is missing (included the cases of outliers whose values have been set to missing), an imputation procedure based on the method of the mean value is used.

In each imputation group (the same groups of the outlier identification procedure) the mean values of the number of job vacancies and the number of employees are calculated. The imputed value is calculated as the product of the ratio between these two values and the number of employees of the firm.

The number of imputed values is quite low: for the 3rd quarter of 2003, only 84 (44 for non manual workers and 40 for manual workers) out of 8120 respondents; for the 4th quarter of 2003, 69 cases (38 for non manual workers and 31 for manual workers) out of 7493 respondents; for the 1st quarter of 2004 85 values (41 for non manual workers and 44 for manual workers) out of 6207 respondents have been imputed; for the 2nd quarter of 2004 13 values (7 for non manual workers and 6 for manual workers) out of 6953 respondents.

While this methodology is quite simple, it disregards the use of longitudinal information and tends to reduce the variance of the number of job vacancies. From the wave for the 3rd quarter of 2004 onward, thanks to the availability of the information referring to one year earlier, it will be possible to consider also the use of longitudinal information.

7. Estimation procedure

The number of job vacancies and the number of employees at the end of the quarter are estimated through a calibration estimator[3]. This kind of estimators is widely used in ISTAT surveys because of its good properties. In short, the main advantage derives from the reduction of the biases originating from unit non responses and register coverage deficiencies (if at the moment of estimation a frame more updated than the one used to draw the sample is available). A second advantage from the use of this kind of estimator is the efficiency gain which is increasing with the correlation between auxiliary and target variables. Third, the resultant estimates are coherent with known population totals. Fourth and last, this estimator is asymptotically design unbiased.

As a frame for grossing up to, the most up to date release of the ASIA register available is used, that is ASIA 2001 for the last two quarters of 2003, and ASIA 2002 for the first two quarters of 2004.

The considered auxiliary variable is the number of employees excluding managers. It is estimated by correcting the number of employees, as in ASIA, with a coefficient, derived from the OROS survey, that represents the share of employees excluding managers on the total number of employees. The reduction coefficient is calculated by cell defined by the section of economic activity and a size class. The rationale for this reduction is the increase in the correlation between the target variable and the auxiliary one.

To better define the calibration estimator let Y be the total of a target variable, that is:

$$Y = \sum_{k \in U} y_k \ ,$$

where U is the target population. The parameter estimator can be written as:

$$\hat{Y} = \sum_{k \in S} y_k d_k \gamma_{ks} = \sum_{k \in S} y_k w_{ks} \ ,$$

where S is the sample of respondents, d_k is the sampling weight, that is the inverse of the inclusion probability of the unit k, γ_{ks} is the correction factor of the sampling weight, and w_{ks} is the final weight. The vector of the final weights for the respondent units is obtained as the solution of a constrained optimisation problem, where the function to be minimized is:

$$E\left[\sum_{s} G(w_{ks}, d_k)\right],$$

where $G(w_{ks},d_k)$ represents the distance between the w_{ks}s and the d_ks.

The constraint can be written as:

[3] The estimator is built in the software Genesees used to calculate both the grossing up weights and the sampling errors.

$$\sum_{k \in S} w_{ks} x_k = X$$

where X is the known total of the auxiliary variable x (see Istat, 2003).

Here, the truncated logarithmic distance function is used, due to its preferable properties with respect to other distance functions. In fact, the Euclidean distance function can produce negative weights and the (non truncated) logarithmic distance function can produce too high weights.

Being part of the family of the calibration estimators, the truncated logarithmic based one, converges asymptotically to the generalised regression estimator and, as mentioned above, it is approximately unbiased and consistent. Moreover its sampling variance converges to that of the generalised regression estimator[4].

8. Indicators of data quality

8.1 Violations of consistency rules

As already mentioned, to avoid a negative effect of the violation of consistency rules on the response rates, none of the consistency rules implemented in the CATI software or in the web-based and e-mail questionnaires prevents the respondent from confirming the figures even if they violate the constraints.

In order to monitor the data quality, a procedure to count the consistency checks violations has been set up. Table 4 shows the violations, in terms of percentage of the total number of respondents, for the main rules. We have classified the rules according to the variables: consistency checks on the stock of employees, on the relation between the stocks and flows of employees and on job vacancies. In the table a violation in a specific group happens if a record has violated at least one of the rules of the group.

The first result is that the number of violations on the rules concerning job vacancies and variables related to the stock of employees (by itself), with the exception of the wave for the 1st quarter of 2004 is almost insignificant.

The maximum number of violations is registered on the consistency check on employment stocks and flows. It requires that the total number of employees at the beginning of the quarter plus the entering ones minus the exiting ones must be equal to the number of employees at the end of the quarter. Firms run into difficulties in measuring the number of employees that began or ceased to work during the reference quarter. This difficulty had already been identified during the pilot

[4] Genesees uses this property and the Woodruff linearisation to calculate the sampling errors (ISTAT, 2004). These are the sampling errors estimated in the job vacancy survey.

survey. However, the CATI technique contributes to contain the number of violations, as it is apparent in the 1st quarter of 2004. It is worth reminding that in this quarter, there were only 520 CATI respondents, hence the relatively high violation rates correspond to a small number of records. Furthermore, the percentage of violations for non CATI respondents is lower than that for CATI respondents in the 2nd quarter of 2004, when non CATI respondents whose data violated various consistency checks were recalled by CATI interviewers in order to verify and, if possible, modify the inconsistent transmitted data. As expected, the percentage of violations for fax respondents is the highest. As elsewhere mentioned, this circumstance has contributed to the decision to dismiss the fax method of response.

Nevertheless, many firms continue to use it. Thus, starting with the wave for the 2nd quarter of 2004, faxes are forwarded to the supplier of the CATI services, that verifies the data and eventually calls the respondents if the data violate consistency checks.

Table 4 – Consistency checks violations (share of total respondents, III quarter 2003 – II quarter 2004)

Consistency check violations	III quarter 2003	IV quarter 2003	I quarter 2004			II quarter 2004	
	CATI	CATI	CATI	WEB	FAX	CATI	WEB
Consistency checks on employment stocks							
non manual workers	0.2	0.5	1.2	5.5	9.1	0.4	0.6
manual workers	0.4	0.3	1.5	3.2	6.4	0.2	0.5
Consistency checks between employment stocks and flows							
non manual workers	2.4	2.8	10.2	17.8	18.5	2.2	1.6
manual workers	2.0	2.8	11.0	16.7	24.3	2.2	1.3
Consistency check on job vacancies							
non manual workers	0.2	0.3	0.2	0.4	0.7	0.0	0.1
manual workers	0.2	0.2	0.0	0.4	0.8	0.0	0.0

8.2 Evaluation of the estimates quality

On the basis of the procedures described above, provisional estimates of the number of job vacancies, the number of employees and the rate of job vacancies referred to the first four waves have been calculated. These estimates, not yet made public at national level, have been sent to Eurostat to be used in the first estimates at EU level.

Table 5 reports the estimates of the percentage relative sampling errors on the number of employees and on the number of job vacancies.

Table 5 – Sampling errors on job vacancies and occupied posts (estimates of percentage relative sampling errors, III quarter 2003 – II quarter 2004)

Economic activity	Job vacancies				Occupied posts			
	III quarter 2003	IV quarter 2003	I quarter 2004	II quarter 2004	III quarter 2003	IV quarter 2003	I quarter 2004	II quarter 2004
C - Mining and quarrying	19.9	31.2	28.4	28.7	1.8	2.0	2.7	1.5
D - Manufacturing	5.8	6.2	6.2	7.2	2.9	1.6	1.5	2.3
E - Electricity, gas and water supply	4.8	2.7	8.5	8.5	0.4	0.4	0.7	0.7
F - Construction	15.6	15.5	21.0	15.6	2.1	3.1	3.4	2.3
G - Wholesale and retail trade; repair of motor vehicles, motorcycles and personal and household goods	8.2	11.6	19.8	14.7	2.8	3.1	6.0	4.0
H - Hotels and restaurants	25.1	46.3	26.3	19.0	2.4	3.4	3.1	3.3
I - Transport, storage and communication	9.2	10.1	16.8	9.3	2.6	1.6	2.4	1.5
J - Financial intermediation	5.6	6.1	5.7	6.8	1.8	1.9	1.0	1.2
K - Real estate, renting and business activities	18.6	18.6	20.3	24.2	4.0	4.7	4.6	4.3
TOTAL	**4.7**	**5.9**	**6.1**	**6.9**	**1.5**	**1.1**	**1.3**	**1.3**

Table 6 – Deft[1] on job vacancies and occupied posts (III quarter 2003 – II quarter 2004)

Economic activity	III quarter 2003		IV quarter 2003		I quarter 2004		II quarter 2004	
	job vacancies	occupied posts	job vacancies	occupied posts	job vacancies	occupied posts	job vacancies	occupied posts
C - Mining and quarrying	0.98	0.18	1.48	0.19	0.42	0.04	1.10	0.02
D - Manufacturing	0.46	0.31	0.40	0.17	0.43	0.17	0.31	0.25
E - Electricity, gas and water supply	0.11	0.01	0.03	0.01	0.39	0.02	0.39	0.02
F - Construction	1.11	0.27	0.66	0.38	0.71	0.30	0.56	0.26
G - Wholesale and retail trade; repair of motor vehicles, motorcycles and personal and household goods	0.23	0.11	0.29	0.11	0.34	0.28	0.35	0.17
H - Hotels and restaurants	0.42	0.11	0.73	0.14	0.91	0.10	0.47	0.13
I - Transport, storage and communication	0.15	0.05	0.23	0.03	0.19	0.04	0.22	0.02
J - Financial intermediation	0.20	0.08	0.26	0.08	0.27	0.05	0.29	0.06
K - Real estate, renting and business activities	0.43	0.22	0.31	0.28	0.39	0.27	0.40	0.26
TOTAL	**0.35**	**0.14**	**0.30**	**0.10**	**0.34**	**0.11**	**0.37**	**0.09**

NOTE: [1] The deft is the square root of the ratio between the variance estimate under the adopted sampling design and estimator and the variance estimate in a hypothetical sampling strategy, consisting in the simple random sampling of a sample of the same size of the actual one and a Horvitz-Thompson estimator.

Further evidence on the efficiency of the adopted sampling design and estimator can be based on the calculation of the deft (that is, the square root of the ratio between the variance estimate under the adopted sampling design and estimator and the variance estimate in a hypothetical sampling strategy, consisting in the simple random sampling of a sample of the same size of the

actual one and a Horvitz-Thompson estimator) concerning the job vacancies and occupied posts (see Table 6). The results show that the adopted sampling design and estimator are remarkably more efficient of a hypothetical sampling strategy consisting in a simple random sampling of a sample of the same size of the actual one and a Horvitz-Thompson estimator.

This result seems to be quite robust because is equally valid on the two variables and, with the exception of three cases, on all the estimation domains.

In the next months, as the information available will increase, the analysis carried out so far to evaluate the estimates quality will be further developed. Furthermore, we will study alternative procedures of missing data imputation and/or estimation. In particular, the possibility of using information from other sources (i.e. the large enterprises survey and the OROS survey) on the employment variables will be thoroughly examined. The study on estimation methods should also evaluate if alternative estimators provide a better variance-bias trade-off. To better validate the aggregate figures the estimates on the number of employees (levels and changes) produced on the basis of the job vacancy survey will be compared more systematically with similar estimates coming from the Labour Force Survey and the OROS survey.

Finally in order to increase the timeliness of data release, automatic data check and editing procedures will be developed. The objective is to release the data on job vacancies and number of employees 90 days after the end of the reference quarter by 2006.

9. Conclusions

The report describes design, running and results of the first four waves (for the last two quarters of 2003 and the first two of 2004) of Istat quarterly business survey on job vacancies.

In particular, the sample design, questionnaire, consistency checks, survey techniques, database, check and editing procedures, estimator, some indicators of data quality (both for sample data and for estimates) have been discussed.

As already mentioned, the target population definition, and the sample design and allocation have been recently revised with respect to those employed in the here described first four waves, on the basis of the data collected in these waves. However, we do not intend to modify the sample size.

We also consider that the design phase is complete and has led to satisfactory solutions for what concerns the questionnaire, the consistency checks implemented in its various versions, the mix of survey techniques (with the exception of those towards the largest size enterprises), and the database structure.

However, it is worth noting that many aspects of the survey methodology (among which, the check and editing and the grossing up to the population procedures, but also the data collection

techniques for the largest enterprises) are still considered not fully stabilised. In fact, we are currently working at improving a number of aspects of the survey and of the data produced, and expect that also the data thus far transmitted to Eurostat may be revised in the future when the design and the implementation of these new methodologies will be completed.

References

Gismondi R. (2000), "Metodi per il trattamento dei dati anomali nelle indagini longitudinali finalizzate alla stima di variazioni", *Contributi*, n. 8, Istat, Rome

Gismondi R. (2002), "Confronti tra metodi per l'individuazione di osservazioni anomale in indagini longitudinali: proposte teoriche e verifiche empiriche", *Rivista di statistica ufficiale*, 1, pp. 25-60, Istat, Rome

Hidiroglou M.A., Berthelot J.M. (1986), "Statistical Editing and Imputation for Periodic Business Surveys", *Survey Methodology*, 12, Statistics Canada, Ottawa, pp. 73-84

Istat (2002), *Pilot Job Vacancy Survey – Final Report*, Contract n. 200032100024, Istat, Rome

Istat (2003), *Genesees v. 3.0 – funzione riponderazione - manuale utente e aspetti metodologici*, Istat, Rome

Istat (2004), *Genesees v. 3.0 – stime ed errori campionari - manuale utente e aspetti metodologici*, Istat, Rome

Annex 1 – Quarterly job vacancy survey questionnaire

ISTAT
ISTITUTO NAZIONALE DI STATISTICA

Enterprise code |__||__||__||__||__||__||__||__||__|
ATECO (Code of economic activity)|__||__||__||__||__|

SURVEY ON JOB VACANCIES AND WORKED HOURS
(CODE PSN ISTAT 01381)
Reference quarter (dates)

SECTION A – ENTERPRISE

Please verify the following data, and modify and/or complete them if needed.

Corporate name
|__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__|
|__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__|

Address
|__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__|
|__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__|

City/Town
|__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__|
|__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__|

Postcode **Province** **Taxpayer's identification code**

|__||__||__||__||__| |__||__| |__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__|
|__||__||__||__||__| |__||__| |__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__|

Description of the main economic activity

Full name of the person filling in the questionnaire
|__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__|

Position and office of the person filling in the questionnaire
|__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__|

Telephone number of the person filling in the questionnaire
|__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__|

Fax number of the person filling in the questionnaire
|__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__|

E-mail address of the person filling in the questionnaire
|__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__||__|

SECTION B – EMPLOYEES

1. How many employees (managers excluded) there were on … (last day of the previous quarter)?

	Non manual workers		Manual workers	
	Total	**Of whom with a part-time contract**	**Total**	**Of whom with a part-time contract**
Employees (managers excluded)	**1.1**|__|__|__|__|__|__|	**1.2**|__|__|__|__|__|__|	**1.5**|__|__|__|__|__|__|	**1.6**|__|__|__|__|__|__|
→*of whom with a fixed term contract*	**1.3**|__|__|__|__|__|__|	**1.4**|__|__|__|__|__|__|	**1.7**|__|__|__|__|__|__|	**1.8**|__|__|__|__|__|__|

2. How many employees (managers excluded) there were on … (last day of the reference quarter)?

	Non manual workers		Manual workers	
	Total	**Of whom with a part-time contract**	**Total**	**Of whom with a part-time contract**
Employees (managers excluded)	**2.1**|__|__|__|__|__|__|	**2.2**|__|__|__|__|__|__|	**2.5**|__|__|__|__|__|__|	**2.6**|__|__|__|__|__|__|
→*of whom with a fixed term contract*	**2.3**|__|__|__|__|__|__|	**2.4**|__|__|__|__|__|__|	**2.7**|__|__|__|__|__|__|	**2.8**|__|__|__|__|__|__|

3. On … (last day of the reference quarter) did all the employees work in only one building?

YES	NO
1 ☐	2 ☐

If not please go to question 5, else please go to question 3bis.

3bis. On … (last day of the reference quarter) did all the employees work in only one of the following geographical areas: North (Piemonte, Valle d'Aosta, Lombardia, Liguria, Trentino-Alto Adige, Veneto, Friuli-Venezia Giulia, Emilia-Romagna), **Centre** (Toscana, Umbria, Marche, Lazio), **South** (Abruzzo, Molise, Campania, Puglia, Basilicata, Calabria, Sicilia, Sardegna)**?**

YES	NO
1 ☐	2 ☐

If not please go to question 5, else please go to question 4.

4. How many employees (managers excluded) worked in the following three geographical areas on … (last day of the reference quarter)?

NOTE: If an employee was working abroad, please include her/him in the geographical area of the office which was paying her/him.

Geographical area		
	Non manual workers	**Manual workers**
4.1 North (Piemonte, Valle d'Aosta, Lombardia, Liguria, Trentino-Alto Adige, Veneto, Friuli-Venezia Giulia, Emilia-Romagna)	**4.1.1**|__|__|__|__|__|__|	**4.1.2**|__|__|__|__|__|__|
4.2 Centre (Toscana, Umbria, Marche, Lazio)	**4.2.1**|__|__|__|__|__|__|	**4.2.2**|__|__|__|__|__|__|
4.3 South (Abruzzo, Molise, Campania, Puglia, Basilicata, Calabria, Sicilia, Sardegna)	**4.3.1**|__|__|__|__|__|__|	**4.3.2**|__|__|__|__|__|__|

5. In the period between … (last day of the previous quarter) and … (last day of the reference quarter):

	Non manual workers	**Manual workers**
5.1 how many workers (managers excluded) started to work as employees of the enterprise?	**5.1.1**|__|__|__|__|__|	**5.1.2**|__|__|__|__|__|
5.2 how many workers (managers excluded) finished to work as employees of the enterprise?	**5.2.1**|__|__|__|__|__|	**5.2.2**|__|__|__|__|__|

SECTION C – JOB VACANCIES

NB: In answering to the questions in this section, please consider only the search for candidates for non managerial positions.

6. On … (last day of the reference quarter) were you looking for people to hire from outside the enterprise?

YES	**NO**
☐	☐

If not please go to section D, else please continue to the following question.

7. On … (last day of the reference quarter) had you already taken active steps to find and hire suitable candidates (for example advertising on mass media or internet, requests to public or private employment agencies, word of mouth with other employers, collection of job applications, interviews of prospective candidates, etc.)**?**

YES	**NO**
☐	☐

If not please go to section D, else please continue to the following question.

8. On … (last day of the reference quarter) for how many posts had you already taken active steps to find and hire suitable candidates and were ready to take further such steps, if needed?

Non manual workers	Manual workers
8.1|__|__|__|__|__|	8.2|__|__|__|__|__|

9. In how many of these posts a suitable candidate could have started to work at once?

Non manual workers	Manual workers
9.1|__|__|__|__|__|	9.2|__|__|__|__|__|

10. For how many of the posts in question 8 a particular difficulty in finding a suitable candidate has been experienced, or is thought likely to be experienced?

Non manual workers	Manual workers
10.1|__|__|__|__|__|	10.2|__|__|__|__|__|

SECTION D –WORKED AND PAID HOURS

All the enterprises which did not answer to the ISTAT OCC.1 questionnaire in all the three months of the reference quarter have to answer to this section.

11. How many hours have been actually worked as normal time, actually worked as overtime, and paid by the employer but not worked by non manual workers (managers excluded) in the quarter …?

	All non manual workers	Of whom with a part-time contract
Hours actually worked as normal time	**11.1** |__|__|__|__|__|__|__|__|	**11.2** |__|__|__|__|__|__|__|__|
Hours actually worked as overtime	**11.3** |__|__|__|__|__|__|__|__|	**11.4** |__|__|__|__|__|__|__|__|
Hours paid by the employer but not worked	**11.5** |__|__|__|__|__|__|__|__|	**11.6** |__|__|__|__|__|__|__|__|

12. How many hours have been actually worked as normal time, actually worked as overtime, and paid by the employer but not worked by manual workers in the quarter …?

	All manual workers	Of whom with a part-time contract
Hours actually worked as normal time	**12.1** |__|__|__|__|__|__|__|__|	**12.2** |__|__|__|__|__|__|__|__|
Hours actually worked as overtime	**12.3** |__|__|__|__|__|__|__|__|	**12.4** |__|__|__|__|__|__|__|__|
Hours paid by the employer but not worked	**12.5** |__|__|__|__|__|__|__|__|	**12.6** |__|__|__|__|__|__|__|__|

13. How many ordinary or exceptional Wage Supplementation Fund ("Cassa Integrazione Guadagni") hours were actually used by employees, and how many hours were not worked by employees (managers excluded) because of labour disputes originated and non originated from employer-employee relations in the quarter …?

	Non manual workers	Manual workers
Ordinary Wage Supplementation Fund ("Cassa Integrazione Guadagni") hours	**13.1** |__|__|__|__|__|__|__|__|	**13.2** |__|__|__|__|__|__|__|__|
Exceptional Wage Supplementation Fund ("Cassa Integrazione Guadagni") hours	**13.3** |__|__|__|__|__|__|__|__|	**13.4** |__|__|__|__|__|__|__|__|
Hours not worked because of labour disputes originated from employer-employee relations	**13.5** |__|__|__|__|__|__|__|__|	**13.6** |__|__|__|__|__|__|__|__|
Hours not worked because of labour disputes non originated from employer-employee relations	**13.7** |__|__|__|__|__|__|__|__|	**13.8** |__|__|__|__|__|__|__|__|

14. If you were to fill in and send the survey questionnaire by yourself, how would you prefer to send it?

14.1 Fax	☐
14.2 Web	☐
14.3 E-mail	☐

OBSERVATIONS BY THE PERSON FILLING IN THE QUESTIONNAIRE: