DOCUMENTI ISTAT

n. 5/2010

Quality assessment and reporting in a short-term business survey based on administrative data

M. C. Congia e F. M. Rapiti

Istat

DOCUMENTI ISTAT

n. 5/2010

Quality assessment and reporting in a short-term business survey based on administrative data

M. C. Congia(*) e F. M. Rapiti(*)

(*) ISTAT – Servizio Statistiche congiunturali sull'occupazione e sui redditi

Quality assessment and reporting in a short-term business survey based on administrative data

Maria Carla Congia, Istat, Servizio Statistiche congiunturali sull'occupazione e sui redditi
Fabio Massimo Rapiti, Istat, Servizio Statistiche congiunturali sull'occupazione e sui redditi

Abstract: Information about the quality of statistics has always been essential both for producers and users. In recent years, at national and international level, it has become more and more important to develop a standard approach to statistical quality assessment. In this paper we present the experience of the Italian Oros survey in adopting standard quality indicators and reporting. The survey estimates quarterly changes in gross wage and total labour cost for firms in the private sector using administrative sources combined with data drawn from the Monthly Survey on Labour Input variables in Large Firms. However, statistics based on the extensive use of administrative data imply additional problems in thinking up and building quality indicators. In fact, it is well known that the use of administrative data for statistical purposes requires to solve unusual quality problems and explore new conceptual and methodological approaches on quality assessing. In a traditional survey non-sampling errors can mainly be faced ex–ante while with administrative data mainly ex-post. This involves that many indicators conceived for traditional statistical surveys cannot be used and new ones must be created. Moreover, the timeliness of quarterly statistics implies further problems in choosing which quality indicators have to be monitored on a short term basis. The Oros survey has been implemented to keep control on the overall quality and efforts have been made to assess and measure quality through appropriate indicators. At first, the calculation of standard quality indicators (SQIs) suggested by Istat central quality experts has been implemented. Then, an Istat quality report (scheda di qualità SIDI) has been tested. Moreover, the Oros survey also satisfies the Labour Cost Index (LCI) Regulation, which implies the delivery to Eurostat of an annual Quality Report. The Italian LCI Quality Reports have been produced since 2004, strictly following the Eurostat approach to statistical quality. The Oros experience confirms that the quality reporting of surveys based on administrative data is difficult to standardise but some advancements can be made carefully choosing the proper indicators and the frequency of quality assessment, taking into account both users needs and production process monitoring requirements.

Keywords: administrative data – short term statistics – quality indicators

Contents

1. Introduction

This paper deals with the Italian experience in assessing quality in a short-term survey based on administrative data. The focus is on the specific problems and quality issues arising using a massive quantity of highly disaggregated administrative data with release timeliness constraints. The consequences on quality indicators and quality reporting are discussed.

After a brief description of the main features of the Oros survey in paragraph 2, the quality issues in using administrative data and the peculiarities of Oros data quality assessment are discussed in paragraph 3. In paragraph 4 the quality indicators from the Oros processing view are presented. A short review of the Oros quality reporting is given in paragraph 5. Some final remarks are presented in paragraph 6.

2. Main features of the Oros survey

Since 2003 the Oros[1] survey estimates quarterly changes in gross wage, other labour costs and total labour cost per full time equivalent (FTE) for all firms in the private sector using administrative sources combined with data drawn from the Monthly Survey on Labour Input variables in Large Firms (hereafter Large Enterprises Survey - LES). The short-term survey was designed to fill a crucial gap in the Italian statistics and to meet the EU Regulations on Short Term Statistics (STS Regulation n.1165/98) and Labour Cost Index (LCI Regulation n.450/2003). The latter requires the quarterly production and transmission of an hourly labour cost index[2]. Until 2002, the Italian National Statistical Institute has collected this information with a monthly frequency, to the limited extent of firms with 500 or more employees, through the LES. Then, the Oros survey was planned to extend the coverage to all business size classes. The huge number of small-size enterprises and the extremely dynamic nature of the Italian firms would have implied the design of a too onerous sample survey, with a considerable impact on the statistical burden on enterprises. The use of the employers' social contribution declarations of the Italian National Social Security Institute (INPS) was the best alternative.

The survey exploits the INPS data for small and medium enterprises estimates, integrated with the data coming from Istat monthly census LES for enterprises with 500 or more employees (Baldi et al., 2004). The target sectors are sections from C to K of the European classification Nace Rev.1.1. Each quarter provisional indicators on the interest variables are released with a delay of about 70 days from the reference period. A revision of the preliminary data follows with a delay of 15 months.

Two different INPS archives are used, which contain respectively employment and wages information and structural details on the administrative units. The first source is the archive of the monthly social contribution declarations (DM10 forms). It refers to the electronic forms that all firms with at least one employee have to transmit to INPS within the 30th day from the end of the reference month. Given the Oros release timeliness, Istat asked INPS to transmit these electronic information as soon as it is uploaded on the central database, implying the availability of data completely raw, not subjected to any administrative check procedure. This set of information, delivered to Istat about 35 days after the end of the reference quarter, is used to produce the provisional estimates. That "provisional population" is extremely large and covers about 97-98% of the entire population which is available with a delay of about 12 months and used to produce the final estimates. In the current situation, each quarter about 1.3 million employers are considered, covering about 10 million employees. The other INPS source is the Administrative Register (AR) which contains structural information on the single administrative unit. It is downloaded at the end of each quarter and needs some treatment to be used for statistical purposes.

---

[1] Oros stands for Occupazione (Employment), Retribuzioni (Wages), Oneri Sociali (Other labour cost).
[2] The hourly LCI is compiled integrating Oros wages and labour costs data at the numerator with an hours worked estimation based on other Istat statistical sources.

3. Quality issues in using administrative data and peculiarities of Oros data quality assessment

It is well known that the use of administrative data for statistical purposes requires to tackle unusual quality problems and explore new conceptual and methodological approaches on quality assessing. In fact, in a traditional survey non-sampling errors can mainly be faced ex–ante while with administrative data mainly ex-post. This involves that many indicators conceived for traditional statistical surveys cannot be used and new ones must be created. In the last decade National Statistical Institutes have accumulated more experience on this topic and also from a theoretical point of view few step ahead have been done (Eurostat 2003, ONS 2007, Thomas 2005, Wallgren and Wallgren 2007). But at the moment, a shared theory of accuracy assessment for statistics based on administrative data is still missing and statisticians seems to lag behind in building it (Platek and Sarndal, 2001). An interesting and promising quality framework has been proposed very recently (Daas et al. 2008) but still a lot of work must be done. Therefore, the Oros approach in data quality assessment is mainly based on a pragmatic view derived from the experience and a deep knowledge of the source.

To understand the peculiarities and the main quality issues in the Oros survey is useful to recall shortly the administrative data exploiting strategy Istat chose to face the release timeliness constraint.

After preliminary studies INPS contribution declarations have been considered to be suitable for Oros purposes although the information contained inside was extremely detailed/disaggregated and the administrative metadata were rather fragmented. As INPS could not aggregate in the very strict time scheduled the DM10 data in the format required for Oros purposes, Istat decided the acquisition of the whole data source. In other words, Istat has been obliged to capture the extremely disaggregated raw micro data, in the original format they are transmitted by firms and without any check by INPS. This constraint has became an opportunity considering that it allows a more direct control on the aggregation/translation process and also a lot of detailed information available for other different statistical purposes. On the other hand, it implies a very complex preliminary phase of checks and computation inside the single DM10 to get to the target variables at micro level.

Moreover, the correct exploitation of the huge quantity of administrative data entails coping with the very frequent changes in the basic INPS metadata which have an high impact on the correct translation of the target variables. Those continuous changes depend on the fact that in Italy a large part of labour, occupational and industrial policies have been taking the form of rebates in social contributions and enterprises have to use the DM10 declaration to take advantage of them. This implies that the process of retrieval of statistical target variables is heavily affected because continuously new components (administrative codes) of labour cost have to be included, while other information have to be excluded because they are not relevant for the statistical purposes.

The combination of a huge quantity of very detailed unchecked microdata in input with frequent changes in administrative metadata and a very short time schedule to produce the final quarterly output has lead to a complex and peculiar quality strategy along the whole process (Congia, Pacini, Tuzi 2008).

The main unusual quality problems of the Oros production process have been faced planning ad hoc instruments and specific check phases (Figure 1):
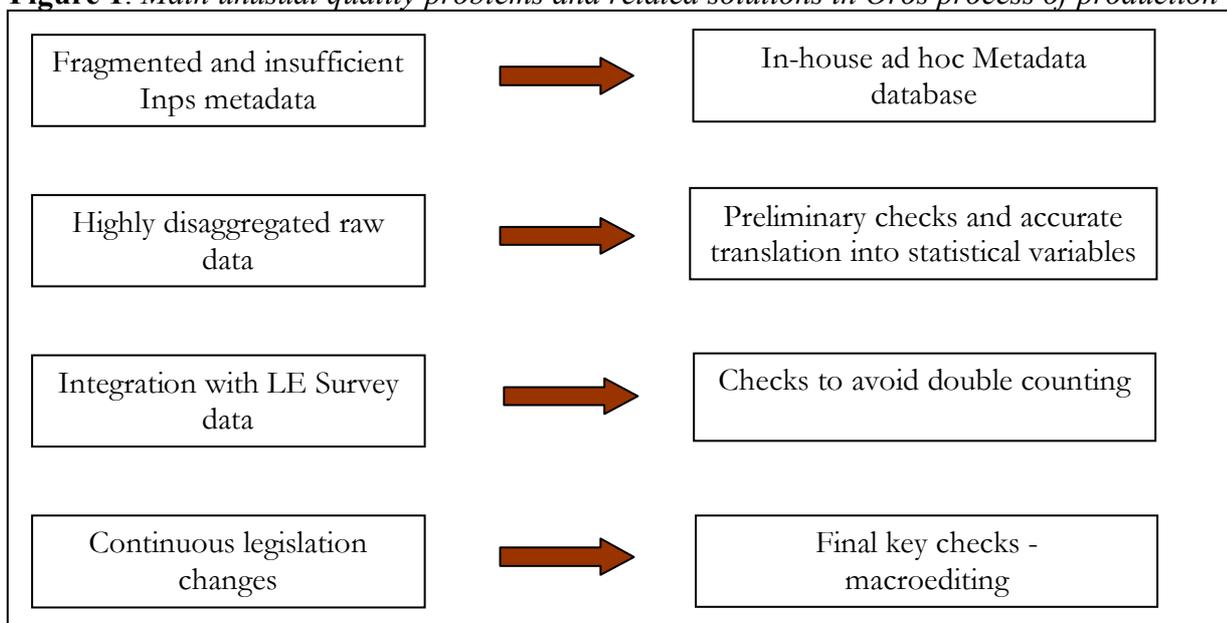
- The availability of fragmented and insufficient INPS metadata has been overcome with the implementation of an in-house ad hoc metadata database collecting laws, regulations and other technical aspects regarding social security contribution (Banca Dati Normativa su retribuzione e contribuzione - BDN). To guarantee a stable retrieval of target statistical variables the BDN has to be quarterly updated[3].
- To assure the quality of the aggregation and translation procedure to obtain the target variables, the monthly declarations go through a complex preliminary check procedure aimed at investigating and possibly correcting errors on administrative micro codes, record duplications, incoherencies with current legislation, etc. Furthermore, other labour costs such as employer injuries insurance premiums (INAIL) and termination of employment relationship allowance (TFR), not recorded in the DM10 declaration, have to be estimated at micro level using other

---

[3] To keep track of and understand changes in concepts, definitions, etc.

sources of information. Afterwards the micro administrative codes are used to aggregate and obtain the target variables.

- The combination between INPS and LES data, realized at micro level, implies specific procedure to avoid units (and employment) double counting and to harmonize the statistical variables.
- To check if residual influential errors or legislation changes (implying administrative micro codes changes) not correctly reported in the previous steps are still present in the data, final key checks on macrodata are carried out.

Figure 1: *Main unusual quality problems and related solutions in Oros process of production*

Fragmented and insufficient Inps metadata	→	In-house ad hoc Metadata database
Highly disaggregated raw data	→	Preliminary checks and accurate translation into statistical variables
Integration with LE Survey data	→	Checks to avoid double counting
Continuous legislation changes	→	Final key checks - macroediting

Those unusual quality issues imply a specific data quality assessment. Generally, before starting to use a particular set of administrative data, statisticians verify if they are suitable for the statistical purposes proposed, in other words, if all statistical requirements are met. Obviously, before using INPS data for the Oros survey all requirement (concept, definitions, classifications, coverage, etc.) have been deeply verified (Baldi et al. 2001). But in the Oros case, given the combination of:

1) frequent changes in law, regulations - and therefore in administrative metadata,
2) short-term frequency and the timeliness of the outputs,

some requirements have to be continuously verified and this adds a further level of difficulty in assessing the overall quality.

Thus, quality checks have been implemented along all Oros production process, trying to identify and develop appropriate quality indicators specific for administrative data to measures, for example:

- efficient and stable data capturing;
- completeness and consistency of metadata;
- stable and correct translation/retrieval of target statistical variables;
- correct integration among sources (INPS and LES).

All these quality indicators are described in details in the fourth paragraph.

The Oros survey has to meet Istat, Eurostat and other international organizations requests on quality reporting. Given the absence of a common framework on the assessment of administrative data those quality requirements are often rather different. Diverse views and practices regard also the basic format of quality reporting. For instance, Eurostat (2003) recommends to produce a source-specific report and a product-specific one because data from each administrative source may be used as an input into numerous statistical products and, conversely, each statistical product may use, as inputs, data from numerous administrative (and statistical) sources. But this distinction, correct from a general point of view, does not work in the Oros case, because there is a very non-conventional use of administrative data and this implies that the two reports overlap.

A short review of the Oros quality reporting experience is illustrated in the fifth paragraph.

4. A quality assessment from the Oros processing view

The quality strategy along the whole Oros production process has been implemented to identify the quality aspects relevant for the Oros statistical use of administrative data and to find proper measures for their assessment. As a consequence of this effort, every quarter quality indicators are measured at each stage of the production process. They are monitored to keep under control data quality and to provide information to improve the process. Some of them are rather traditional indicators, useful mainly as documentation to assess the quality of process and output. Others are more oriented to carefully monitor every step of the quarterly process. The latter can signal decisive problems or detect sources of error helping survey managers to react quickly, solve problems and, if necessary, to run again the procedures.

Table 1: *Quality dimensions, indicators and measures for the different phases of the Oros production process*

Quality aspect/ dimension	Indicator	Measure/method
DATA CAPTURING		
Accessibility	Privacy and security assurance in the administrative file transmission	- Is the file crypted by INPS? Y/N - Successful decrypting? Y/N
Consistency	Integrity of the administrative file (presence of wrong characters invalidating the file processing)	- Rate of wrong characters - Rate of characters converted - Rate of residual critical characters
Timeliness	Frequency and target timing of deliveries	- Days from the end of reference period for the DM10 "provisional population". - Days from the end of reference period for DM10 population - Days from the end of reference period for AR
Punctuality	Delay from the target delivery date	- Time lag of delay in days
Consistency	Expected number of records and of DM10 forms (each DM10 lays on several records)	- Absolute number of records - Absolute number of units (DM10) - Ratio records/units
	Reference month of each DM10 form	- Rate of units with wrong reference period
METADATA		
Clarity	Availability of administrative metadata	- Number of official INPS acts to be collected, analysed and stored in the metadata database - Number of INAIL acts to be collected and analysed - Information on other sources
Timeliness	Availability of updated administrative metadata	- Last updating of basic metadata on INPS web-site - Last updating of metadata on INAIL web-site

Comparability	Stability (changes) over time of administrative concepts and definitions	- Number of new and expired DM10 codes
		- Rate of DM10 codes by type (debts and credits)
		- Rate of codes to include/exclude for the retrieval of the statistical variables
		- Number of changes of INPS employer contribution rates
		- Number of changes of INPS employee contribution rates
	Expected quantitative impact of administrative metadata changes over time	- Y-on-Y monthly changes of INPS employee contribution rates
		- Annual change of INAIL premium rates by sector

ADMINISTRATIVE DATA TREATMENT

Accuracy	Coherence of administrative data (errors on DM10 codes-coherence with current legislation) Correct retrieval of statistical variables Correct aggregation of DM10 data Duplicate units treatment	- Rate of DM10 codes with formal error
		- Rate of edited DM10 codes
		- Rate of invalid DM10 codes
		- Rate of duplicate DM10 codes
		- Weighted rate of codes to include/exclude for the computation of the statistical variables
		- Rate of non-coherent values by codes
		- Rate of duplicate units (DM10)
	Quality of identification and classification administrative information (INPS-AR)	- Rate of wrong business id-code (fiscal code)
		- Rate of missing business id-code (fiscal code)
		- Rate of edited business id-code (fiscal code)
		- Rate of units with a "first best" economic activity code (drawn from BR-ASIA)
		- Rate of residual units with a "second best" economic activity code (INPS-AR)
		- Rate of units without economic activity code
	Overcoverage due to units outside the survey scope	- Rate of out-of-scope units (outside C-K sections, public units)

EDITING AND IMPUTATION

Accuracy	Unit response rate for the DM10 "provisional population"	- Response rate=Units of the DM10 "provisional population" on units of the DM10 population (unweighted and weighted)
	Cross-sectional and longitudinal coherence of the target variables at unit level (microediting)	- Rate of units with violated edit rules
		- Rate of units imputed
		- Rate of items imputed
		- Total contribution to key estimates from imputed values
		- Rate of units interactively checked by sub-population

	Unit non-response imputation	- Rate of units imputed - Total contribution to key estimates from imputed units
	Presence of influential errors (selective editing by sub-population)	- Rate of units imputed - Rate of items imputed - Total contribution to key estimates from imputed values
Accuracy/ Comparability	Coherence of macro data (macroediting)	- Number of time series comparable with other sources - Rate of suspicious domains identified by analytical checks - Rate of suspicious domains identified by automatic detection (TRAMO-Error) - Rate of domains checked - Rate of units imputed after drill-down checks - Total contribution to key estimates from imputed values
INTEGRATION WITH SURVEY DATA		
Coherence	Correct matching between administrative and large enterprises survey (LES) micro data	- False negative mach rate - False positive mach rate - Rate of units interactively checked (new entry, exit, big changer, incorrect fiscal code)
RELEASE		
Accuracy/ Reliability	Revision policy of target estimates	- Mean Revision - Mean Absolute Revision
Timeliness	Time lag between the end of the reference quarter and the date of the release	- Days from the end of the reference quarter to the date of the national press release - Days from the end of the reference quarter to the date of the transmission to Eurostat
Punctuality	Time lag between the scheduled and the actual date of the release	- Days from the scheduled press release date and the actual one - Days from the scheduled date of the transmission to Eurostat and the actual one
Comparability	Comparability over time	- Length of times series comparable over time
Accessibility	Forms of data dissemination	- National press release (Y/N) - Yearbooks (Y/N) - On-line time series data base - Eurostat press release - Yearly country report
Clarity	Metadata dissemination	- National Press release notes - Metadata for Eurostat - SDDS metadata for international institutional users - Metadata in the System on the Quality on the Istat website (SiQual)

	Relevance	Contacts with users and media	- Number of quarterly access to the Istat on-line time series data base - Number of quarterly phone or e-mail contacts with users - Numbers of quarterly media releases and articles on Oros indicators

In Table 1 the quality dimensions and indicators monitored from the Oros processing view are shown. They are measured mainly with a quantitative approach while for a few only a qualitative assessment has been possible. They are classified by the six different production phases. On the whole, the largest number of indicators has been monitored to assess the accuracy quality dimension. Also comparability and timeliness components have been assessed through several indicators. Overall, the quality indicators produced in the entire Oros production process are 28 (Table 2).

Table 2: *Number of quality indicators for the different phases of the Oros production process by different quality dimensions*

	Accessibi-lity	Clarity	Timeli-ness	Punctua-lity	Consisten-cy	Compara-bility	Accuracy	Cohere-nce	Relevan-ce	TOTAL
1 Data capturing	1		1	1	2					**5**
2 Metadata		1	1			2				**4**
3 Administrative data treatment							6			**6**
4 Editing and imputation						1	4			**5**
5 Integration with survey data								1		**1**
6 Release	1	1	1	1		1	1		1	**6**
TOTAL	**2**	**2**	**3**	**2**	**2**	**3**	**11**	**1**	**1**	**28**

Different dimensions of quality are involved in different stages of the Oros process. Consistency, accessibility and punctuality are important quality aspects affecting data capturing, while the metadata phase is more related to comparability, clarity and timeliness. Accuracy dimension is a fundamental concern in both administrative data treatment and editing and imputation phases. Integration with survey data involves mainly the coherence dimension assessment, while the release stage is interested in the evaluation of almost all the quality dimensions.

5. Oros experience in producing standard quality indicators and reports

Since 2003, when the release of Oros quarterly indicators has started, different quality reporting needs have emerged. Quality reports can satisfy different target groups: producers, top management, central quality managers, international organizations and final users. In the last years many different kinds of standard reports has been proposed (Eurostat 2009, Laliberté et al. 2004, Brancato et al.2004).
Among the different quality indicators monitored in the Oros production process, only a little sub set is directly suitable for reporting data quality to both external and internal users in a standard way. Quality requirements initially were built from a theoretical point of view, generally referring to a typical sample survey, mainly cross-section. Moreover, they focused mainly on final output quality assessment.

At Istat level, after fruitful interactions between short-term survey managers and the Istat central quality management, more suitable indicators have been identified through a better reconciliation between theory and survey practices.

Nevertheless, many important quality indicators related to short-term survey based on administrative data still have little or no role in quality reporting existing guidelines.

An exhaustive overview of the Oros quality reporting is given along two dimensions in Figure 2: the frequency of the reporting production and the different target users.

At the beginning, a basic assessment of general quality and administrative data suitability has been produced ("Survey documentation and methodology handbook"). This initial quality assessment describes in details the survey methods and the production of quarterly indices, including the suitability of concepts, definitions, translation scheme of administrative information into statistical variables, coverage, reference time, stability over time. This kind of documentation is produced in the survey design and implementation phases, and updated and modified when relevant methodological changes occur. This documentation is mainly qualitative and it is addressed to all users.

Then, every quarter the internal "Process monitoring report" is carried out, as described in § 3. Starting from these quality indicators produced quarterly, it has been possible selecting a proper set of indicators, to populate the Istat SIDI Information System for Survey Documentation and all other recurring reports required. A short description of the most important characteristics of the different reports, referring to the two main indicators produced (Oros per FTE, hourly LCI), follows:

1. SIDI (Information System for Survey Documentation)
    - developed mainly to provide the top management and the central quality managers with qualitative and quantitative information on survey quality;
    - the basic qualitative documentation on each survey is partially disseminated on the System on the Quality (SIQual) available on Istat website;
    - required with reference to the Oros per FTE indicators released at national level;
    - annually updated;
    - includes a qualitative description of the survey features and methodology as well as a range of standard quality indicators on:
        o Coverage
        o Response rate
        o Comparability
        o Timeliness
        o Revision policy
        o Human resources costs

2. Istat Quality Report:
    - external-user oriented;
    - should be disseminated within the System on the Quality (SIQual) available on Istat website, but it is yet in an experimental phase;
    - produced as pilot report with reference to the Oros per FTE indicators released at national level;
    - based on both quantitative and qualitative approach within a framework coherent with Eurostat quality components;
    - consists of a subset of standard quality indicators appropriately chosen within those available from the SIDI:
        o Response rate
        o Indicators on the revision policy (MR, MAR)
        o Timeliness for provisional data release
        o Timeliness for definitive data release
        o Length of the homogeneous time series
    - contains also qualitative description of non-sampling error, relevance and accessibility dimensions.

Oros quality reporting: an overview

+ PROCESS

+ PRODUCT

Target group / Frequency	Oros Producers	Internal users (of micro and macro data)	Top managers and central quality managers	External Expert users (Eurostat, IMF, BCE,..)	General Public
	Survey Documentation and Methodological Handbook				
Once				Metadata in SDDS	
					Oros PR explanatory notes
Annually			SIDI information system for survey documentation	LCI Quality Report	Istat Quality Report
Quarterly	Oros Process Monitoring Report			Quarterly LCI meta information	

+ QUALITATIVE

+ QUANTITATIVE

3.   Labour Cost Index Quality Report:
   - required by Eurostat to evaluate the quality of national hourly LCI used to produce the European aggregate index;
   - annually produced and transmitted since 2004;
   - based on a standard structure according to Eurostat dimensions of quality, with a further aspect related to the "completeness" (compliance) about the technical Regulation requirements:
     - Impact of revisions (MR, MAR)
     - Timeliness release
     - Coherence with Quarterly National Accounts
     - Longitudinal coherence
   - particular attention is paid to the description of the method for compiling hours worked (LCI denominator).

4.   Quarterly LCI meta information:
   - a standard template, release-specific, required by Eurostat togheter with the quarterly LCI data transmission;
   - mainly based on qualitative indicators, contains a description of changes in the labour market (collective agreements, laws) which have an impact on quarterly data;
   - particular attention is paid to reasons of revisions in NSA, WDA and SA data and details on changes of modelling seasonal adjustment.

5.   Sdds Metadata:
   - required periodically by Eurostat as basic documentation of both Oros per FTE indicators and hourly LCI;
   - based on a standard template proposed by IMF (International Monetary Fund);
   - contains mainly qualitative indicators and is divided in two separate parts:
     - Basic information that identifies four dimensions of data dissemination (the data: coverage, periodicity, and timeliness; Access by the public; Integrity of the disseminated data; Quality of the disseminated data)
     - Summary methodology with details on analytical framework, concepts, definitions and classifications

6.   Oros press release explanatory notes:
   - a very short standard document attached to the Oros per FTE quarterly press release;
   - describes general aspects of the survey, methods and definitions;
   - contains only qualitative description
   - addressed to all users to inform timely about any relevant change of methodology

This overview shows a variety of requirements that have to be met with different reporting tools. Moreover some quality dimensions and indicators are rather heterogeneous.
Little attention is paid on administrative data and no specific quality indicators are required. In the Oros case the assessment of crucial quality aspects related to administrative data use is not reported to users. Finally, a better integration and systematizing of all the reviewed quality reporting tools is desirable but only partially achievable.

6. Conclusions

The paper has reviewed the different reporting tools used to assess quality of the quarterly Oros survey and the numerous quality indicators used mainly from a producer point of view have been described. The original use of administrative data has forced the Oros survey managers to monitor peculiar

aspects of the quality not usually taken into consideration in the standard quality assessment approaches suggested by Eurostat or other international organizations. In fact, this quality assessment mainly derive from a bottom up approach, learning by doing, by monitoring and improving process quality, rather than applying general framework to the Oros case.

This confirms that the quality reporting of surveys based on administrative data is difficult to standardise but some advancements can be made carefully choosing the proper indicators and the frequency of quality assessment, taking into account both users needs and production process monitoring requirements.

There are several different ways of approaching administrative data quality assessment but it doesn't exist a clear best approach as each one can be more or less useful to producer or users according to the specific context.

The studies to get to better and generalised frameworks should go on combining theories and practices, "bottom up" and "top-down" strategies. The most interesting advancement in literature, not by chance has been written by NSI statisticians with high practical experience in management of administrative data (Daas et al. 2008, Thomas 2005).

In the Oros case several specific indicators to assess the quality of the process, in particular the metadata updating and the translation/aggregation of raw INPS data, have been successfully implemented. These specific indicators are essential from the producer point of view, but they could also be used to report to the users some key quality issues. On the other hand, the Oros survey satisfies the Istat internal (SIDI) and external (Eurostat, IMF) requests of standard quality reports using more traditional indicators. At Istat level, some advancements have been made in choosing more appropriate indicators for documenting quality of short-term business surveys. But some more efforts are needed to get to a higher harmonization to avoid producing too many, slightly different, indicators.

Moreover, to improve Oros quality reporting more work must be done in two directions:
- the Oros more specific indicators need to be further systematized and maybe combined with the more traditional ones;
- given the trade-off between the need to monitor many different dimensions and indicators and the timeliness, a right balance must be found choosing an appropriate subset of indicators to be produced and used on a short-term basis.

References

Baldi C., E. Cimino, F. Rapiti, P. Minicucci, R. Succi, D. Tuzi (2001), *L'utilizzo dei dati INPS per la stima trimestrale del numero dei dipendenti, le retribuzioni, il costo del lavoro e le ore lavorate*, Documenti Istat, 14, Istat, Rome.

Baldi, C., and F. Ceccato, E. Cimino, M.C. Congia, S. Pacini, F. Rapiti, D. Tuzi (2004), *Use of Administrative Data to produce Short Term Statistics on Employment, Wages and Labour Cost*, Essays, 15, Istat, Rome.

Brancato, G. et al. (2004), *Standardising, Evaluating and Documenting Quality: the Implementation of ISTAT Information System for Survey Documentation – SIDI*, paper presented at the European Conference on Quality and Methodology in Official Statistics (Q2004), 24-26 May 2004, Mainz, Germany.

Congia, M.C., Pacini S., Tuzi D. (2008), *Quality Challenges in Processing Administrative Data to Produce Short-Term Labour Cost Statistics,* paper presented at the European Conference on Quality in Official Statistics (Q2008), 8-11 July 2008, Rome, Italy.

Daas, P.J.H., and J. Arends-Tóth, B. Schouten, L. Kuivenhoven (2008), *Proposal for a quality framework for the evaluation of administrative and survey dat*a, paper presented at the Workshop on the Combination of surveys and administrative data, Vienna, Austria.

Eurostat (2009), *ESS Standards for Quality Reports*, Luxembourg.

Eurostat (2003), *Quality Assessment of Administrative Data for Statistical Purposes*, Luxemburg.

Laliberté L, W. Grünewald and L. Probst (2004): *Data Quality: a comparison of IMF's Data Quality Assessment Framework and Eurostat's Quality Definition,* January, 2004. http://dsbb.imf.org/vgn/images/pdfs/dqaf_eurostat.pdf

ONS (2007), *Guidelines for Measuring Statistical Quality*, Version 3.1, London.

Platek, R., and C.-E. Sarndal (2001), *Can a Statistician Deliver?*, Journal of Official Statistics, Vol.17, No.1, pp. 1-20.

Rapiti F. (2006), *Short-term statistics quality reporting: the LCI National Quality Report 2004*, Documenti Istat, 16, Istat, Rome.

Thomas M. (2005), *Assessing the Quality of Administrative Data*, Survey Methodology Bulletin, No 56, pp. 77-84.

Wallgren, A., and B. Wallgren (2007), *Register-based Statistics. Administrative Data for Statistical Purposes*, West Sussex: Wiley.

Documenti ISTAT(*)

1/2006 – Roberta Roncati – *www.istat.it (versione 3.0) Il nuovo piano di navigazione*

2/2006 – Maura Seri e Annamaria Urbano – *Sistema Informativo Territoriale sulla Giustizia: la sezione sui confronti internazionali*

3/2006 – Giovanna Brancato, Riccardo Carbini e Concetta Pellegrini – *SIQual: il sistema informativo sulla qualità per gli utenti esterni*

4/2006 – Concetta Pellegrini – *Soluzioni tecnologiche a supporto dello sviluppo di sistemi informativi sulla qualità: l'esperienza SIDI*

5/2006 – Maurizio Lucarelli – *Una valutazione critica dei modelli di accesso remoto nella comunicazione di informazione statistica*

6/2006 – Natale Renato Fazio – *La ricostruzione storica delle statistiche del commercio con l'estero per gli anni 1970-1990*

7/2006 – Emilia D'Acunto – *L'evoluzione delle statistiche ufficiali sugli indici dei prezzi al consumo*

8/2006 – Ugo Guarnera, Orietta Luzi e Stefano Salvi – *Indagine struttura e produzioni delle aziende agricole: la nuova procedura di controllo e correzione automatica per le variabili su superfici aziendali e consistenza degli allevamenti*

9/2006 – Maurizio Lucarelli – *La regionalizzazione del Laboratorio ADELE: un'ipotesi di sistema distribuito per l'accesso ai dati elementari*

10/2006 – Alessandra Bugio, Claudia De Vitiis, Stefano Falorsi, Lidia Gargiulo, Emilio Gianicolo e Alessandro Pallara – *La stima di indicatori per domini sub-regionali con i dati dell'indagine: condizioni di salute e ricorso ai servizi sanitari*

11/2006 – Sonia Vittozzi, Paola Giacchè, Achille Zuchegna, Piero Crivelli, Patrizia Collesi, Valerio Tiberi, Alexia Sasso, Maurizio Bonsignori, Giuseppe Stassi e Giovanni A. Barbieri – *Progetto di articolazione della produzione editoriale in collane e settori*

12/2006 – Alessandra Coli, Francesca Tartamella, Giuseppe Sacco, Ivan Faiella, Marcello D'Orazio, Marco Di Zio, Mauro Scanu, Isabella Siciliani, Sara Colombini e Alessandra Masi – *La costruzione di un Archivio di microdati sulle famiglie italiane ottenuto integrando l'indagine ISTAT sui consumi delle famiglie italiane e l'Indagine Banca d'Italia sui bilanci delle famiglie italiane*

13/2006 – Ersilia Di Pietro – *Le statistiche del commercio estero dell'Istat: rilevazione Intrastat*

14/2006 – Ersilia Di Pietro – *Le statistiche del commercio estero dell'Istat: rilevazione Extrastat*

15/2006 – Ersilia Di Pietro – *Le statistiche del commercio estero dell'Istat: comparazione tra rilevazione Intrastat ed Extrastat*

16/2006 – Fabio M. Rapiti – *Short term statistics quality Reporting: the LCI National Quality Report 2004*

17/2006 – Giampiero Siesto, Franco Branchi, Cristina Casciano, Tiziana Di Francescantonio, Piero Demetrio Falorsi, Salvatore Filiberti, Gianfranco Marsigliesi, Umberto Sansone, Ennio Santi, Roberto Sanzo e Alessandro Zeli – *Valutazione delle possibilità di uso di dati fiscali a supporto della rilevazione PMI*

18/2006 – Mauro Albani – *La nuova procedura per il trattamento dei dati dell'indagine Istat sulla criminalità*

19/2006 – Alessandra Capobianchi – *Review dei sistemi di accesso remoto: schematizzazione e analisi comparativa*

20/2006 – Francesco Altarocca – *Gli strumenti informatici nella raccolta dei dati di indagini statistiche: il caso della Rilevazione sperimentale delle tecnologie informatiche e della comunicazione nelle Pubbliche Amministrazioni locali*

1/2007 – Giuseppe Stassi – *La politica editoriale dell'Istat nel periodo 1996-2004: collane, settori, modalità di diffusione*

2/2007 – Daniela Ichim – *Microdata anonymisation of the Community Innovation Survey data: a density based clustering approach for risk assessment*

3/2007 – Ugo Guarnera, Orietta Luzi e Irene Tommasi – *La nuova procedura di controllo e correzione degli errori e delle mancate risposte parziali nell'indagine sui Risultati Economici delle Aziende Agricole (REA)*

4/2007 – Vincenzo Spinelli – *Processo di Acquisizione e Trattamento Informatico degli Archivi relativi al Modello di Dichiarazione 770*

5/2007 – Anna Di Carlo, Maria Picci, Laura Posta, Michaela Raffone, Giuseppe Stassi e Fiorella Tortora – *La progettazione dei Censimenti generali 2010-2011: 1 - Analisi, valutazione e proposte in merito ad atti di normazione e finanziamento*

6/2007 – Silvia Bruzzone, Atonia Manzari, Marilena Pappagallo e Alessandra Reale – *Indagine sulle Cause di Morte: Nuova procedura automatica per il controllo e la correzione delle variabili demo-sociali*

7/2007 – Maura Giacummo, Carlo Vaccari e Monica Scannapieco – *Indagine sulle Scelte Tecnologiche degli Istituti Nazionali di Statistica*

8/2007 – Lamberto Pizzicannella – *Sviluppo del processo di acquisizione e trattamento informatico degli archivi relativi al modello di dichiarazione 770. Anni 2004 – 2005*

9/2007 – Damiano Abbatini, Lorenzo Cassata, Fabrizio Martire, Alessandra Reale, Giuseppina Ruocco e Donatella Zindato – *La progettazione dei Censimenti generali 2010-2011 2 - Analisi comparativa di esperienze censuarie estere e valutazione di applicabilità di metodi e tecniche ai censimenti italiani*

10/2007 – Marco Fortini, Gerardo Gallo, Evelina Paluzzi, Alessandra Reale e Angela Silvestrini – *La progettazione dei censimenti generali 2010 – 2011 3 – Criticità di processo e di prodotto nel 14° Censimento generale della popolazione e delle abitazioni: aspetti rilevanti per la progettazione del 15° Censimento*

11/2007 – Domenico Adamo, Damiana Cardoni, Valeria Greco, Silvia Montecolle, Sante Orsini, Alessandro Ortensi e Miria Savioli – *Strategie di correzione del questionario sulla qualità della vita dell'infanzia e dell'adolescenza. Indagine multiscopo sulle famiglie. Aspetti della vita quotidiana 2005*

12/2007 – Carlo Nappi – *Manuale per la preparazione di originali "ready to print"*

1/2008 – Franco Lorenzini – *Indagine sulle unità locali delle imprese: la flessibilità organizzativa e il ruolo degli uffici regionali come strategia per la riduzione del disturbo statistico e il raggiungimento di elevati tassi di risposta*

2/2008 – Elisa Berntsen, Simone De Angelis, Simona Mastroluca – *La progettazione dei Censimenti generali 2010-2011 4-L'uso dei dati censuari del 2000-2001: alcune evidenze empiriche*

3/2008 – Marina Peci – *Progetto SCQ -Scuola Conoscenza Qualità-Statistica e Studenti*

4/2008 – Giampiero Siesto, Franco Branchi, Cristina Casciano, Tiziana Di Francescantonio, Piero Demetrio Falorsi, Salvatore Filiberti, Gianfranco Marsigliesi, Umberto Sansone, Ennio Santi, Roberto Sanzo e Alessandro Zeli – *Messa a regime dell'uso dei dati fiscali (Modelli UNICO) per l'integrazione delle mancate risposte e la riduzione del numero delle unità campione della rilevazione PMI*

5/2008 – Giovanni Seri e Maurizio Lucarelli – *A.D.ELE. Il laboratorio per l'Analisi dei Dati ELEmentari. Monitoraggio dell'attività Anni 2004-2007*

---

* ultimi cinque anni

6/2008 – Francesco Altarocca – *Strumenti informatici innovativi nella conduzione di indagini statistiche*

1/2009 – Silvia Dardanelli, Simona Mastroluca, Alessandro Sasso e Mariangela Verrascina – *La progettazione dei censimenti generali 2010 – 2011*
5 - Novità di regolamentazione internazionale per il 15° Censimento generale della popolazione e delle abitazioni

2/2009 – Rossana Balestrino e Alberto Gaucci – *Tecniche di cattura dati nei processi di produzione statistica*

3/2009 – Barbara Fiocco – *Le "misure" dell'Italia nell'Annuario Statistico Italiano*

4/2009 – Daniela Pagliuca, Raffaella Cianchetta, Marco Broccoli, Teresa Buglielli, Roberto Di Giuseppe e Diego Zardetto – *L'Osservatorio Tecnologico per i Software generalizzati (OTS) nel 2008*

5/2009 – Silvia Losco – *Il riuso informatico nelle Pubbliche Amministrazioni: normativa e prime esperienze in Istat*

6/2009 – Fabio Crescenzi Marco Fortini, Gerardo Gallo e Andrea Mancini – *La progettazione dei censimento generali 2010 – 2011*
6 - Linee generali di impostazione metodologica, tecnica e organizzativa del 15° Censimento generale della popolazione

7/2009 – Silvia Losco – *Gli standard informatici dell'Istat*

8/2009 – Alfredo Roncaccia e Roberto Iannaccone – *L'indagine sulle Opere Pubbliche dalla costituzione dell'Istituto Centrale di Statistica ai giorni nostri*

9/2009 – Ugo Guarnera, Orietta Luzi e Massimo Greco – *La procedura automatica di controllo e correzione dell'indagine SPA 2007: aggiornamenti e integrazioni*

10/2009 – Francesca Brait, Claudia De Vitiis, Roberto Petrillo, Monica Russo, Massimo Strozza e Paola Ungaro – *L'indagine sui dottori di ricerca: un'esperienza pilota*

11/2009 – Marco Grazzi, Roberto Sanzo, Angelo Secchi e Alessandro Zeli – *ISTAT - MICRO.3 A new integrated system of business micro-data 1989–04*

12/2009 – Daniela Ichim – *Italian Farm Structure Survey: SDC Methodology for an MFR Dissemination*

1/2010 – Patrizia Grossi, Silvio Stoppoloni – *La progettazione dei censimenti generali 2010 – 2011 Informazione e formazione ai comuni sulle innovazioni di metodi e tecniche per il 15° Censimento generale della popolazione e delle abitazioni*

2/2010 – Donatella Tuzi – *L'individuazione di dati anomali nella Rilevazione Oros: un approccio in serie storiche basato su "Tramo for Errors"*

3/2010 – Lidia Brondi – *L'evoluzione del Sistema Statistico Nazionale nel periodo 1992-2008 attraverso l'analisi statistica delle schede identificative dei progetti Psn*

4/2010 – Andrea Brancatello, Luisa Ciardelli, Susi Osti, Raffaella Succi, Marina Attili e Claudia Iaccarino – *Il processo produttivo delle statistiche demografiche: le procedure e i controlli di qualità*

5/2010 – Maria Carla Congia e Fabio Massimo Rapiti – *Quality assessment and reporting in a short-term business survey based on administrative data*