DIGGING INTO LABOUR MARKET DYNAMICS: TOWARD A RECONCILIATION OF STOCK AND FLOWS SHORT TERM INDICATORS*

F. Rapiti, C. Baldi, D. Ichim, F. Pintaldi, M. E. Pontecorvo, R. Rizzi

Abstract

The recent availability of short-term job flows from Compulsory Communication (Cc) administrative database has increased the publicly disseminated information. While creating opportunities for a deeper labour market understanding, job flows measures might also generate misleading interpretations, especially when compared with stocks statistics. To address this issue and to fill data gaps, the Italian National Institute of Statistics (Istat) promoted an Agreement among Istat, Inps, Ministry of Labour and social policy, Inail and Anpal to produce harmonised, complementary and coherent statistics with three specific outputs: a quarterly joint press release, an annual report and a labour market statistical information system. The main challenge is to reconcile flows and stock data deriving from administrative and statistical sources. This paper illustrates the statistical treatment applied to the Cc database to achieve a better comparability with official statistics used in the quarterly joint press release disseminated since December 2016.

Sintesi

APPROFONDENDO LA DINAMICA DEL MERCATO DEL LAVORO: VERSO UNA RICONCILIAZIONE DEGLI INDICATORI CONGIUNTURALI DI STOCK E FLUSSI

La recente disponibilità di dati congiunturali sui flussi di posizioni lavorative da fonte amministrativa (Comunicazioni Obbligatorie-Co) ha aumentato le informazioni diffuse al pubblico. Pur creando opportunità per una più approfondita comprensione del mercato del lavoro, le misure relative ai flussi potrebbero anche generare interpretazioni fuorvianti, soprattutto se confrontate con le statistiche sugli stock. Per affrontare questo problema e colmare dei data gap nelle statistiche sul mercato del lavoro, l'Istituto Nazionale di Statistica (Istat) ha promosso un accordo con Ministero del lavoro e politica sociale, Inps, Inail e Anpal per produrre statistiche armonizzate, complementari e coerenti. La sfida principale è riconciliare i dati sui flussi e sugli stock derivanti da fonti amministrative e statistiche. Questo paper illustra il trattamento statistico applicato ai dati delle Co, nel contesto del comunicato trimestrale congiunto pubblicato a partire da dicembre 2016, per ottenere una migliore comparabilità con le statistiche ufficiali.

Key words: employment, labour market, stock and flows indicators

1 Introduction

In recent years in Italy there has been an increase in availability of short-term indicators from administrative sources aiming at monitoring the employment trends. Aspiring to favour a better understanding of the labour market, the Ministry of labour and social policy (Mlps) and the National social security institute (Inps) have been disseminating new short-term statistics based on jobs flows. The brand new and interesting job flows indicators have surely enriched the public debate on employment

---

. Sometimes they have also contributed to generate misleading interpretations, especially when compared with Istat traditional Labour force stock statistics (Anastasia, 2017). Different figures provided for (almost) the same phenomenon by various institutions turned out in an overload of information. The natural multidimensionality of the labour market and the consequent use of different indicators only partly explains the users' impression of experiencing an information "excess". The diffusion of data which are not completely coherent with the concepts, definitions and classifications of official statistics (Istat, 2015) has given rise to misinterpretation, exaggeration and inappropriate interpretation of technically correct numbers. Some media talked about a "war of numbers on employment" focused on the numerical assessment of the recent labour market reform (Giovannini, 2015).

Data, information and knowledge are three different concepts and moving from the first to the latter is a complex process in which official statistic have a responsibility. National statistical authorities have to promote whatever action is necessary to empower public by improving their capacity of extraction of useful knowledge from data. In turn, such actions will increase official statistics credibility. Istat, aware of the key importance of the labour market issue, decided to act for the sake of an informed, transparent and impartial public debate, and promoted a Coordination Agreement among five Institutions Istat, Inps, Mlps, the National Institute for Insurance against Accidents at Work (Inail) and subsequently National Agency for Active Labour Policies (Anpal). The scope is to cooperate to exploit the full potential of the available information assets, to fill information gaps and to disseminate integrated, harmonized, non-redundant and high quality data and analyses. The Agreement, started to be implemented in December 2016, relies on sharing different data sources, metadata and methodology to make, together, a step forward to understand the labour market functioning. The five institutions should jointly release information to empower public to comprehend the real state of employment. There are three specific outputs: a quarterly joint press release with short-term indicators, an annual joint structural report and an underlying statistical information system to be jointly used by the five institutions. By now, the first two outputs are released on a regular basis, while the third is under construction.

The paper focuses on the data integration process applied for the publication of the quarterly joint press release. In particular, the statistical treatment applied to the Compulsory Communication-Cc database (flows data) to achieve comparability with stock based official statistics (Istat) is illustrated. The main challenge is to reconcile flows and stock data derived from administrative and statistical sources.

The structure of the paper is the following. Section 2 synthetizes the recent debate on labour market trends and the relevance of stock and flows data. In Section 3, after a brief description of the Cc database, the statistical treatment of flows data is detailed. Several comparisons with other data sources are presented in Section 4. Some concluding remarks and future work are sketched in Section 5.

2 The recent debate on labour market trends and the importance of stock and flows information

In 2015-16 the public debate on labour market focused on the effects of the policy reform called "Jobs Act" and in particularly on the social security hiring incentive (so called "decontribuzione") and the new open-ended contract ''Contratto a tutele crescenti''. In the discussions, stock have been opposed to flow indicators, statistical surveys to administrative data and jobs against persons employed. Some national

newspapers even misinterpreted quarter-over-quarter activations changes with quarter-over-quarter employment changes (Anastasia, 2017).

It is useful to remind here that a stock quantity is measured at a specific point in time (or averaged over a period), whereas a flow quantity is measured over an interval. As reported in the next section, the level of employment in a given moment in time is a stock measure, while its change between two moments is a flow measure. More precisely, it is a net flow which can be obtained also from the balance between the gross flows in and out of employment. The analysis of the change starting from the gross flows provides further insights into the labour dynamics revealing the massive gross flows behind the much smaller net flows (highlighting the underlying heterogeneity).

Stock and flows labour indicators are currently produced from two different sources (and points of view): mainly household surveys and data collected from the side of employer. In the first case, they refer to persons (headcount), in the latter to jobs. The most important and used short-term indicators to monitor labour market are Istat Lfs's quarterly average stocks (employment, unemployment, etc.). Based on Oros survey, Istat also releases quarterly indicators of average stocks of jobs. It should be stressed that the Mlps and Inps releases both refer to flows (activations, cessations, transformations). However, the most extensive and long term experience in regularly releasing short- term gross and net flows of jobs is that of Veneto Lavoro-Labour market observatory[1]. As for other international experiences, only Statistics New Zealand and Bureau of labor statistics in USA release short-term quarterly indicators. The first institute releases results with a delay of 12 months and employs administrative data while the latter uses a traditional business survey[2].

The main objective of the quarterly joint press release (hereafter Note) is to present information from different sources in order to provide an integrated overview of employment trends. To reach this aim, a particular attention has been given to metadata information with a systematic review of the concepts, definitions, classifications, differences between data sources and implications of different survey techniques. The analyzed data are those from the Istat's Lfs and Oros processes, the Inps's Observatory on precariousness, the Inail's work injuries and the Mlps's Cc. The latter has been processed in a way to make these indicators as much comparable as possible with stocks statistics (see section 3).

The Note was welcomed by the media and succeeded in clarifying many issues related to the interpretation and comparison of different indicators. It has apparently achieved the result to calm the stormy waters of the public debate about employment trends. Notwithstanding, the complexity of the phenomena still leaded some analysts to require further simplification and invoke a single indicator, possibly "carved in stone". However, in a data driven society in which any private and public decision is based on information, more data and indicators, produced by different providers, means also more freedom to choose and to evaluate, which, in turn, strengthen a pluralistic and democratic society (Alleva, 2017).

3 Jobs, Activations and Cessations. A bit of stock-flow accounting

In this section, after a brief description of the Cc database, the definitions and computations needed to relate Cc flow statistics to stock statistics are illustrated. Since March 2008, the Compulsory Communications System (Cc) encloses all the

---

[1] http://www.venetolavoro.it.

[2] See https://stats.govt.nz/ and https://www.bls.gov/bdm/.

communications sent by public and private employers to notify activation, extension, transformation and cessation of jobs, as well as business transfers and change of company name. The Cc encloses different forms. The main one is Unilav, through which all private and public employers (excluding the temporary work agencies for agency workers) fulfil the obligation to communicate any job-related event, i.e. activations, extensions, transformations and cessations of jobs. The events must be communicated within the day before their entry into force.

Every form has a Linked Employer Employee Database information structure with individual data regarding both the employer (fiscal code, name, economic activity sector, flag public administration, registered office, place of work) and the worker (fiscal code, surname, name, sex, date and place of birth, citizenship, permit to stay and motivation, place of residence, education level). Jobs are characterized by a common set of information, i.e. contract type, working time, professional qualification, seasonal work, etc. Moreover, specific dates, depending on the kind of event (activation, cessation, training period cessation, extension, transformation), are useful to determine the duration of job. The coverage of the Cc source is really wide and for this reason very useful for statistical purposes. In terms of employers, it includes all public and private employers belonging to all economic sectors with the exception of armed forces. Referring to the contract type it includes: all regular employees; part of self-employed workers (project workers, occasional contractors, self-employed in entertainment sector, business agents, association in participation; stages and socially useful work. Cc excludes also some managerial contracts of public and private corporations. For further details on the Cc see Baldi *et al.* 2011.

In order to derive Cc-based statistical indicators comparable with the more traditional ones, the link between stocks and flows (Davis *et al.*, 2006) is exploited. The Cc's implicit statistical unit is the employer-employee work relationship, here called job and denoted by $\mathbb{J}_j$. A job is characterized by an activation date, $t_a(j)$, and, possibly, a cessation date, $t_c(j)$. Let us consider a fixed period of time q, for example a given quarter, identified by beginning, $b(q)$, and ending dates, *e(q)*. Considering a pair q - $\mathbb{J}_j$, in a given date *t*, a job may be active or not; in the period q, $\mathbb{J}_j$ may be active at least for one day or not active at all. In notation, the period-indicator function

$$I_j(q) = 0, if\ t_c(j) < b(q)\ or\ e(q) < t_a(j);\ 1, otherwise \qquad [1]$$

signals whether a given job is active during q. Similarly, the date-indicator function

$$I_j(t) = 1, if\ t_a(j) \leq t \leq t_c(j);\ 0, otherwise \qquad [2]$$

signals whether a given job is active at date *t*. It is straightforward to observe that the following identity holds:

$$I_j(t) = I_j(q) * I_j(t)\ \forall j, q\ and\ \forall\ b(q) \leq t \leq e(q)\ .$$

It follows that J_t, i.e. the number of active jobs at any date *t*, is given by $J_t = \sum_j I_j(t)$. Consequently, the average number of jobs over the period q may be expressed as

$$\bar{J}_q = \frac{\sum_{t=b(q)}^{e(q)} J_t}{e(q) - b(q)} = \frac{\sum_{t=b(q)}^{e(q)} \sum_j I_j(t)}{e(q) - b(q)} = \frac{\sum_j \sum_{t=b(q)}^{e(q)} I_j(t)}{e(q) - b(q)} \qquad [3]$$

$$= \frac{\sum_j \sum_{t=b(q)}^{e(q)} I_j(t) * I_j(q)}{e(q) - b(q)} = \sum_j \frac{\sum_{t=b(q)}^{e(q)} I_j(t)}{e(q) - b(q)} * I_j(q) = \sum_j D_{qj}^s * I_j(q)$$

where D_{qj}^s denotes the standard duration of a job $\mathbb{J}_j$ in a period q, i.e. $D_{qj}^s = \frac{\sum_{t=b(q)}^{e(q)} I_j(t)}{e(q) - b(q)}$ (the standard duration is defined as the number of days the job has been active during

q standardized by the length of (q). Equation [3] states that $\bar{J}_q$ equals the weighted sum of the active jobs where the weights equal the standard durations.

The stock-flow relationship relates the number of jobs, activations and cessations: the difference between the number of jobs at times t and $t\text{-}1$ equals the difference between the number of activations at time t and the number of cessations at time $t\text{-}1$:

$$J_t = J_{t-1} + A_t - C_{t-1} \qquad [4]$$

where A_t and C_t represent the number of job activations and cessations occurring at a given date t, respectively (Figure 1). The asymmetry between A_t and C_t is due to equation [2]. Indeed, a job $\mathbb{J}_j$ has to be considered "active" even in its end date $t_c(j)$. Consequently, a job will be excluded from the jobs stock only the day after $t_c(j)$.

It is easy to see that both J_t and $\bar{J}_q$ may be expressed in terms of a stock of jobs at a time 0 and a difference between activations and cessations. In fact, putting $t = 0$ in equation [4] and then iterating over time t, the number of jobs J_t may be expressed as:

$$J_t = J_0 + \sum_{s=1}^{t} A_s - \sum_{s=0}^{t-1} C_s = J_0 + A_t^u - C_t^u \qquad [5]$$

Averaging over the period q, the average number of jobs over q may be obtained:

$$\bar{J}_q = \frac{\sum_{t=b(q)}^{e(q)} J_t}{e(q) - b(q)} = J_0 + \frac{\sum_{t=b(q)}^{e(q)} \sum_{s=1}^{t} A_s}{e(q) - b(q)} - \frac{\sum_{t=b(q)}^{e(q)} \sum_{s=0}^{t-1} C_s}{e(q) - b(q)} = J_0 + \overline{A_q^u} - \overline{C_q^u} \qquad [6]$$

where $\overline{A_q^u}$ and $\overline{C_q^u}$ represent, respectively for the activations and cessations, the average over the period q of the flows cumulated from the starting point.

Figure 1: Cumulated daily activations, cessations, net flows and quarterly averages Q1.2015 –Q4.2017
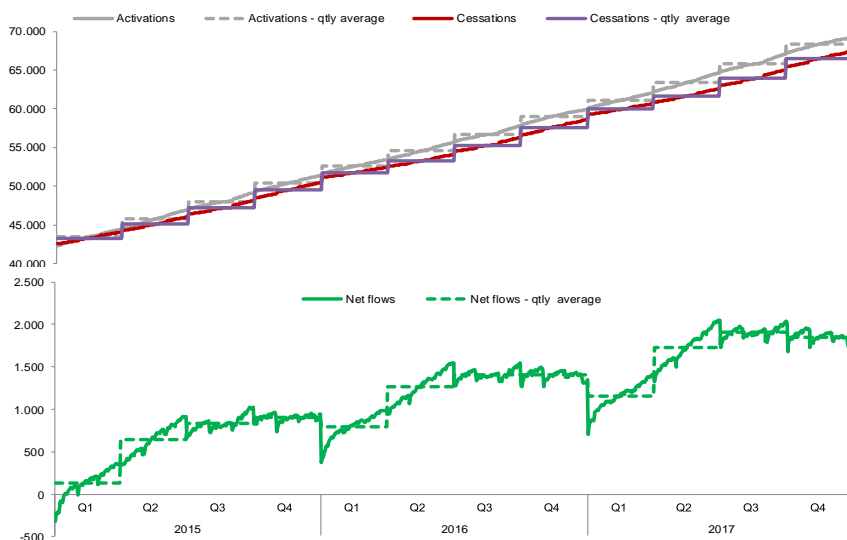
The Cc system registers only job activations and cessations. Thus, only the terms A_t^u and C_t^u in [5] and $\overline{A_q^u}$ and $\overline{C_q^u}$ in [6] may be derived using the Cc data. It follows that the number of jobs cannot be calculated within the Cc system. However, the changes of this level may be calculated. Indeed, using [5], the *change between pointwise number of jobs* (CPJ) at the end dates of two consecutive quarters becomes:

F. Rapiti, C. Baldi, D. Ichim, F. Pintaldi, M. E. Pontecorvo, R. Rizzi

$$J_{e(q)} - J_{e(q-1)} = \left(J_0 + \Sigma_{s=1}^{e(q)} A_s - \Sigma_{s=0}^{e(q)-1} C_s\right) - \left(J_0 + \Sigma_{s=1}^{e(q-1)} A_s - \Sigma_{s=0}^{e(q-1)-1} C_s\right) \qquad [7]$$
$$= \Sigma_{s=b(q)}^{e(q)} A_s - \Sigma_{s=b(q)-1}^{e(q)-1} C_s$$

Furthermore, using [6], the *change of the average number of jobs* (CAJ), becomes:

$$\bar{J}_q - \bar{J}_{q-1} = \left(J_0 + \overline{A_q^u} - \overline{C_q^u}\right) - \left(J_0 + \overline{A_{q-1}^u} - \overline{C_{q-1}^u}\right) = \left(\overline{A_q^u} - \overline{C_q^u}\right) - \left(\overline{A_{q-1}^u} - \overline{C_{q-1}^u}\right) \qquad [8]$$

Notice that CAJ in equation [8] takes into consideration the standard duration of each job in the two quarters, while CPJ in equation [7] does not. Indeed, in equation [7], only pointwise periods, i.e. the end dates of different quarters, are considered. On the contrary, in equation [8], the average number of jobs over a period is involved and $\bar{J}_q$ implicitly takes into account the standard durations, see [3]. It means that CAJ better measures the change of total amount of work between periods.

The two indicators CAJ and CPJ have both weaknesses and strengths. Being based on the average number of active jobs over a period, the CAJ allows for a better comparability to many labour market indicators, for instance, the Labour Force Survey and Oros. From the timeliness point of view, CPJ clearly outperforms CAJ. An unexpected surge of activations at the end of a quarter should be signaled much more evidently by CPJ than by CAJ. However, the meaningfulness of the CPJ signal depends on the duration of the created jobs. If the duration of the activated jobs is considerable, the CPJ could signal such labour input increase soon while CAJ would signal it only the following quarter. If, instead, the created jobs are very short lived, CPJ could give "too strong" and erratic signals that might be misinterpreted as a sign of durable increase in employment. This feature is crucial in a system with a very high share of short-term contracts (Istat et al. 2017). In other terms, average-based indicators, like CAJ, would provide a more accurate measure of the "long-term" impact of such events, but at the cost of a possible delay. The two ways of calculating changes thus prove complementary to each other and their joint use improves the interpretation of the labour market short-term evolution.

4 Comparison of different indicators

This section compares three of the sources analyzed in the quarterly joint Note, Cc (calculated as CAJ), Oros and Lfs, which are more comparable thanks to the length of their time series. These sources differ in several aspects: administrative vs. survey data; stock vs. flows measures, demand side and jobs (Cc and Oros) vs. supply side and persons (Lfs). Thus, different definitions of employment arise: Cc and Oros are based on the relationship between employer and worker established through a formal work contract; Lfs is based on individuals who claim to have done at least one hour of paid work in the reference week. Therefore, the difference between the coverage of the sources must be considered: Lfs includes dependent and independent, regular and irregular employment and all the economic activity sectors (A-U sections of Nace 2007); Cc includes regular employees and some dependent self-employment in the A-U sectors, while Oros includes regular employees in the industry and services sectors (from B to S, excluding O). Thus, in order to correctly compare the sources, the analysis was limited only to the employees in industry and private services (B to N) sectors, excluding temporary agency worker contracts (which are absent from Cc) and job-on-call work because this type of the contract do not imply an actual labor input. The Cc, Lfs and Oros labour market trends are very close. They unequivocally depict the decline in employment during the economic crisis and the troughs of the business cycle at the end of 2013. The subsequent employment recovery appears stronger in

the administrative source (Cc). In the year-on-year quarterly changes comparison, the dynamics of jobs in Oros and Cc exhibit a more similar trend. The Lfs trend is more irregular, showing opposite patterns with respect to Cc trend in few quarters and a forward shift in the 2016 employment growth peak (Figure 2). In the absolute cumulative changes, the Cc series shows a divergence with respect to the other two; the drift increases over time especially from the end of 2015 (Fig. 3).

Figure 2: Lfs, Cc (CAJ) and Oros absolute year-on-year absolute quarterly changes in industry and private service (B-N). Q1.2011-Q4.2017. Raw data
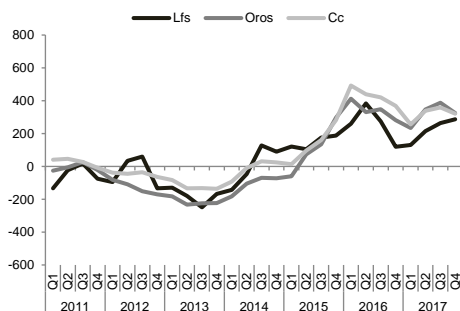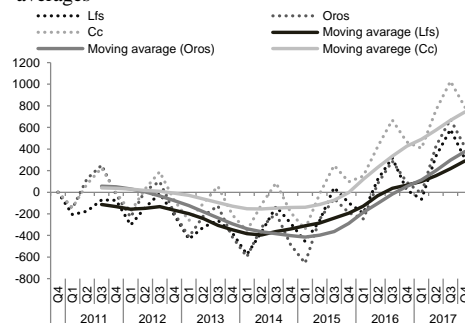
Figure 3: Lfs, Cc (CAJ) and Oros absolute cumulative changes in industry and private service (B-N). Q1.2011-Q4.2017 (Q4.2010=0). Raw data and moving averages

On the other hand, the series of cumulative changes of Lfs and Oros are closer in the last period and slightly compensate the diverging trends in previous years. The differences may be due to the previous definition and measurement issues, as well as to over-coverage and under-coverage errors (De Gregorio *et al.*, 2014). Indeed, administrative data generally tend to have over-coverage errors, e.g. false activations, while survey data more often have under-coverage errors, e.g. sampling and non-sampling errors, (Statistics Canada 2009).

To further investigate the determinants of such differences between trends, a microdata level record-linkage was performed. The Cc and Lfs 2011-2015 microdata were integrated by linking the records corresponding to the individuals having at least one activation in the Lfs reference week. The very preliminary results show that there are cases of employed persons without an employment relationship (persons absent in the Cc data; probably irregular workers) and cases of contracts without work (i.e. unemployed or inactive persons in Lfs with formal contract in Cc[3]). Differences were also found with respect to job duration. There are fixed-term contracts from the Lfs side which are open-ended in the Cc data. On the contrary, a continuous and uninterrupted succession of fixed-term contracts with the same employer can be considered "de-facto open-ended contracts" (Istat-Cnel, 2017). In general, individuals tend to declare their own employment status rather than the formal one. These first results suggest that the differences between the de-facto (from Lfs) and the formal (Cc) conditions should be considered. In some cases, such differences may indicate different and more complex phenomena rather than simple inconsistencies.

5 Conclusions and future development

The Agreement signed in December 2015 represents a real step forward in the collaboration among the five Italian main producer of labour market statistical

---

[3] This may be caused by temporal misalignment between work and contract, especially for short-term jobs and different irregular practices (fictitious work i.e. false assumptions only to get social security benefits).

information. In the quarterly joint Note the combination of comparable variables on clear and unambiguous domains and use of precise definitions and transparent methodology enhance the release of more coherent statistics. The statistical treatment of the Cc data source allows the estimation of the change of average number of active jobs over a period (CAJ), an indicator more comparable and consistent with traditional statistics on stocks. However, comparing the sources, non-negligible deviations in trends remain. Such deviations are due to definitional and conceptual differences and to many other factors that must be further investigated. A more systematic micro integration of the sources is going to be carried on. This is one of the tasks of the inter-institutional technical working group created to implement the Agreement.

The integration of sources at micro level involves a new phase of literacy to statistical data, oriented to the complexity of the modern labour market phenomena. Sometimes, it is not much a question of distinguishing between "right" and "wrong" data but rather the correct interpretation of its meaning. The differences between the sources must be carefully analyzed, being able to discern between: i) unwanted and (theoretically) avoidable errors (e.g. non-sample survey errors, wrong reporting in questionnaire or in administrative form); ii) irregular or false declarations for interest of the individual or firm; iii) diversity of meaning of the phenomenon. The non-complete comparability between data from different sources has always been one of the main obstacles to their integrated use. This issue becomes even more relevant when developing an integrated statistical system based on microdata. In this context, consistency at micro level becomes crucial. Furthermore, the ability to distinguish the reasons causing the differences and the importance of define hierarchies among sources are fundamental. The final objective is to derive an integrated and validated micro dataset which enables deeper labour market analyses founded on the full exploitation of the richness of integrated variables (profession, education level, etc.).

References

Alleva G., "Troppe statistiche? Inevitabile", Il Messaggero, 14 March 2017.

Anastasia B., (2017), "Le statistiche aiutano a capire il mercato del lavoro o servono a confondere?" in Economia e Società Regionale n. 3.

Baldi C., De Blasio G., Di Bella G., Lucarelli A., Rizzi R.: *Turning the Compulsory Communication data into a statistical system,* Springer International Publishing Switzerland 2014 "Studies in Theoretical and Applied Statistics" (p.215-224).

Davis S. J., Faberman R. J., Haltiwanger J., The Flow Approach to Labor Markets: New Data Sources and Micro–Macro Links, Journal of Economic Perspectives, 2006, 20 (3), pp 3–26.

De Gregorio C., Filipponi D., Martini A., Rocchetti I., (2014), *A comparison of sample and register based survey: the case of labour market data. Proceedings of the European conference on quality in official statistics* (Quality 2014), Vienna.

Giovannini R., "La guerra dei numeri sul Jobs Act è appena cominciata", in Internazionale, 23 April 2015.

Istat 2015, "Audizione del Presidente dell'Istat G. Alleva all'Ufficio di Presidenza della Commissione Lavoro del Senato", 30 September.

Istat, Ministry of labour and social policy, Inps, Inail, Anpal, 2017, "Il mercato del lavoro. Verso una lettura integrata", https://www.istat.it/it/archivio/206846.

Istat-Cnel (2017), Bes 2017. *Il benessere equo e sostenibile in Italia*, Rome, Istat.

Paolini O., Fortis M. (2017) "Una proposta per fermare l'overdose di statistiche", Il Messaggero, 4 March.

Statistics Canada 2009 "Statistics Canada Quality Guidelines". Third Edition - October 1998.