

An Overview of the Use of Administrative Data for Business

Statistics in Europe

Costanzo, Luigi¹

ISTAT, Division of Statistical Registers, Administrative Data and Statistics on Public Administration

Via Adolfo Ravà 150

Rome (00144), Italy

E-mail: lucostan@istat.it

1. Introduction: source of information and contents

This paper presents some findings of an information collection carried out in 2009-2010 within the Admin Data ESSnet (Work Package 1), in order to provide an Overview of existing practices in the use of administrative data for producing business statistics over Europe². Thirty-one countries (the 27 EU members and the 4 EFTA countries) were covered, with regard to two main topics:

- a. *General issues regarding the use of administrative data for statistical purposes* (legal basis, cooperation with administrative data holders, organisational solutions for managing administrative data, projects to increase/ improve the use of administrative data);
- b. *Existing practices in the use of administrative data for producing business statistics*, in the domains of four major EU Regulations: Business Register (BR), Short-term statistics (STS), Structural business statistics (SBS) and PRODCOM. In particular: combination of sources used (and why administrative data are not used, where applicable), main features of the administrative sources currently in use, use of estimation methods.

Given the limits of this presentation, only a few key issues will be addressed. The whole dataset is available to NSI researchers at the Admin Data ESSnet Information Centre (<http://essnet.admindata.eu>).

2. General issues: legal basis and cooperation with administrative data holders

In principle, the use of administrative data for producing official statistics is recognised and fostered by all national statistical laws currently in force, since the few laws that still did not explicitly enable such use have been purposely revised or replaced in the last decade. Despite of this, in several countries, a variety of factors hamper the full implementation of such a principle, so that – in practice – not all NSIs are in the position of exploiting all the potential of administrative data. Some of these factors are intrinsic to the legal basis (points of weakness, restrictions due to confidentiality laws), while others concern the cooperation between NSIs and Administrative data providers, although in most cases the two aspects are closely linked.

As regards the legal basis, two major points of weakness can be identified in several statistical laws:

- a. *The right of the NSI to access administrative data is established only in principle* (so that further

¹ Presenting author. Co-authors: Di Bella, Grazia [ISTAT / Via A. Ravà 150 / Rome (00144), Italy / dibella@istat.it]; Hargreaves, Elizabeth [ONS / Government Buildings, Cardiff Road / Newport, South Wales (NP10 8XG), United Kingdom / elizabeth.hargreaves@ons.gsi.gov.uk]; Pereira, Humberto J. [INE / Avenida A.J. de Almeida / Lisbon (1000), Portugal / humberto.pereira@ine.pt]; Rodrigues, Sofia [INE / Avenida A.J. de Almeida / Lisbon (1000), Portugal / sofia.rodrigues@ine.pt].

² The Admin Data ESSnet is a project granted by Eurostat within the framework of the MEETS Programme. The information collection consisted in a literature/ web search (1st stage, November 2009-January 2010), whose findings were submitted to the NSIs in the form of partly pre-filled modules to be edited/ completed and validated (2nd stage, October-December 2010).

authorisations or even special legal provisions are needed case by case) or without clearly establishing a corresponding duty for the administrative data holders (so that data are supplied by virtue of provision agreements, whose terms must be negotiated from time to time)³;

- b. *The law does not give power to the NSI for the coordination of administrative data providers, e.g. in the design of administrative forms/ information systems, for the adoption of shared definitions/ classifications, etc. – which makes harder to obtain administrative data of suitable quality and/ or to implement integrated systems of data collection/ provision*⁴.

Furthermore, some countries report restrictions related to data confidentiality, but in most cases these appear to be rather linked to policy issues than to real legal barriers. In fact, current laws on data protection normally provide for the NSIs special rights of access to administrative sources, under the bond of statistical secrecy and other limitations such as the solely statistical purpose and the ban on transferring data to third parties. Such exceptions, however, can be questioned or not applied where the NSI is not an autonomous administration, but a department of a Governmental body⁵.

As already mentioned, the features of the legal basis explain only a part of the variability observed in the actual access to administrative data for statistical use – much of the rest being explained by something more elusive but equally relevant: *the quality of cooperation between NSIs and administrative data holders*. This is always indicated as a key factor in all most successful experiences made in this field, starting from the pioneering ones (and still the most advanced) of France and Nordic countries. Also data quality issues, which are often indicated as the reason why available administrative data are not being used for statistical purposes, appear, at a closer look, to be strictly related to cooperation. In most cases, in fact, the problems reported are those typical of a lack of coordination within the Public administration (such as incompatibilities of technical standards, mismatches of definitions/ classifications, non-shared criteria for the maintenance of data sources, etc.).

Although based on different approaches, France (on one side) and Denmark, Finland, Norway and Sweden clearly represent the European forefront in the statistical use of administrative data. In these countries, efficient integrated systems for producing business statistics have been implemented since long time, within which surveys are normally used – where necessary – only to supplement administrative data (and not vice-versa), so reducing to the minimum the statistical burden on enterprises.

- In the “*French model*”, highly centralised, the NSI is to directly manage a “hybrid” business register (SIRENE) serving for both administrative and statistical purposes, and also a wide range of sources used to update it – which places under its control much of the generation process of administrative data.
- In the “*Nordic model*”, highly participative, the emphasis is more on the interaction between the NSIs and the administrative data holders, which are involved in a regular consultation process through which common goals and strategies are being continuously defined and implemented.

Both models rely, however, on very strong legal frameworks, that enable the NSIs to effectively act as coordinators of the whole Public administration as regards the production of official statistics and both have been consolidated over decades of practice.

Compared with 1993, when the BR Regulation was issued by the European Community, the divide separating France and Nordic countries from the rest of Europe has greatly reduced. At that time, very few

³ The first case is reported, among others, in Germany. The second represents the standard situation in most of European countries.

⁴ Also this case can be considered as a standard situation. In most countries, the law identifies the NSI as the coordinator of a wider National statistical system, in which normally participate all major administrative data holders. Despite of this, the involvement of the NSI in the design or review of administrative information systems is very uncommon.

⁵ This case is reported, among others, in Belgium.

countries were using (mostly to a marginal extent) administrative data for producing business statistics. Today this practice has become quite common and in all observed countries (except Iceland) a statistical BR exists, serves as infrastructure for the whole production of business statistics and is regularly updated (in whole or in part) by using administrative data. Nevertheless, the overall picture remains quite uneven: in many countries there are still important gaps to be filled and, generally speaking, the above models appear to be hardly replicable in different legal and cultural contexts. For this reason, an enhancement of the statistical cooperation among public bodies appears to be – in the current situation – the main road (if not the only one) to increase the use of administrative data. In such direction, several initiatives have recently undertaken in various countries, indicating pragmatic routes and goals for the foreseeable future. To name a few:

- In *Ireland*, the NSI has launched, in cooperation with the Tax authorities, a vast programme of feasibility studies aimed at realising the Statistical Potential of Administrative Records (SPAR).
- In *Italy*, a recent decree (2010) assigns to the NSI the task of “defining methods and forms to be used by public bodies for the use/ exchange of statistical information, and to coordinate possible modifications of forms and information systems used by public bodies to collect information relevant to statistical purposes”.
- In *Portugal*, the implementation of the IES system (Simplified Business Information) – realised in cooperation by the NSI and the major holders of administrative data on businesses, has led to a complete replacement of SBS surveys with administrative sources.
- In *Slovenia*, a Statistical Advisory Committee has been purposely established to deal with administrative sources, mainly focusing on legal basis, acquisition, uses and harmonisation of administrative registers for statistical purposes.

3. Some snapshots from the Overview of existing practices

The information collection carried out within the Admin Data ESSnet aims to provide a progress report on the process through which the traditional ways of collecting data on business activities are being replaced with more efficient ones, entailing lesser costs for the NSIs and lesser burden for the respondents.

In this perspective, the *Business Register* offers perhaps the best viewpoint to understand to which extent, at a national level, administrative data are available and actually used. As shown in Fig. 1, in a slight majority of countries, the BR is updated also by means of regular surveys, carried out to collect information which are not available (or deemed not usable/reliable) from administrative sources. Normally, only little subsets/ samples of the business populations are covered, mostly to get information about local units of larger enterprises, newborn enterprises or belongings to enterprise groups, and sometimes simply to check/improve the quality of the characters assigned by the administrative sources. The use of survey data for updating the BR does not necessarily indicate a less advanced use of administrative sources: it rather marks, in many countries, the limit of the current possibilities to advance in the change over from statistical surveys to administrative sources for producing business statistics.

In the domain of *STS*, production processes that are entirely based on administrative data are, of course, quite rare. Some national experiences, however, demonstrate that there are possibilities to expand the use of administrative sources, especially with regard to *Turnover* and *Employment* variables. VAT data are already used as the main source for Turnover in Finland, France, Netherlands and Slovenia; while, for Employment, Social security data are used in Austria, Belgium, Cyprus, Italy and the Netherlands, and Personal income tax data (along with administrative registers) are used in Denmark and Finland. Moreover, several countries not currently using administrative data for these variables have reported that methodologies are under development or (more frequently) that relevant administrative data exist and would be available, but cannot be used for technical reasons/ quality issues (Tab. 1). Clearly, this seems to be a case where investing in the cooperation with administrative data holders (adoption of common standards, joint feasibility studies, etc.) could lead to reduce the recourse to surveys. Also in the sub-domain of *Building permits*, where administrative data are largely used (but often collected from local authorities through surveys), significant

improvements appear to be at hand, thanks to the spread of ICT at all levels of local government – which now makes easier to implement register-based systems, as recently made in Austria.

In the domain of *SBS* – consisting of 9 sub-domains, each covered by a specific Annex to the EU Regulation – analysis was conducted by Annex, and not by variable. As regards the *Annexes I-IV* (Common module, Industry, Trade and Construction), Portugal is the sole country where, as already mentioned, surveys have been completely replaced by administrative sources. In a large majority of cases (Fig. 1), a combination of survey data and administrative/ BR data is used (usually the surveyed units are those above a set threshold). Only 4 countries are reported as using only surveys as a direct source of input. Prospects for an expansion of the use of administrative data in this area appear to be interesting, as few countries have reported special obstacles to such a development (Tab. 1). The Portuguese example demonstrates that a complete replacement of survey inputs is possible, but only as a result of an extraordinary effort of the whole Public administration. In several countries, however, a wider use of administrative data, and then a substantial reduction of the statistical burden, seems to be – for the time being – a realistic goal.

Figure 1. Countries of EU & EFTA by combination of direct sources used for producing business statistics and business statistics domain (2010)



Elaboration from Admin Data ESSnet WP1, *Deliverable 1.2/2010. Database "Overview of Existing Practices in the Uses of Administrative Data for Producing Business Statistics in EU and EFTA"* (2011).

Administrative data are more widely used in the production of *Annexes V-VII* (Insurance, Banks and Pension funds). In this case, the same bodies that collect relevant administrative data (Central Banks or other supervisory Authorities) are often responsible for compiling statistics, either autonomously or in cooperation with the NSIs.

The situation is very different for *Annex VIII* (Business services), where no country is reported as using only administrative data, and a slight majority (Tab. 1) use only survey data. The leading experiences seem to be those reported in Austria, Finland, Netherlands, Norway and Sweden – where administrative data are used also as direct sources of input to cover some variables (namely turnover and employment, again).

Annex IX (Business demography) is a special case, because all information needed should be available

from the statistical BR. The countries reported as using survey data for producing these statistics are only 8, but all of them use also BR and/or administrative data (at least as an indirect source).

Finally, because of their subject, Prodcum statistics offer little possibility for using administrative sources in replacement of direct surveys, and the few practices reported are all of marginal importance.

Table 1. Countries of EU & EFTA not using administrative data as a direct source for producing business statistics by reason why administrative data are not used and business statistics domain (2010)

Domains	Why administrative data are not used (<i>Pct. of valid responses</i>)					Valid responses (<i>Absolute values</i>)
	A methodology is being developed	Admin. data available but not suitable, for various reasons	No admin. data exist which are relevant	Legal restrictions	All relevant data are available from BR	
Turnover	23.1	69.2	7.7	-	-	13
New orders	-	8.3	91.7	-	-	24
STS Prod. prices/costs	4.5	9.1	86.4	-	-	22
Building permits	14.3	14.3	71.4	-	-	7
Employment	23.1	30.8	30.8	7.7	7.7	13
Annexes I-IV	50.0	12.5	25.0	12.5	-	7
Annex V	-	50.0	-	50.0	-	2
SBS Annex VI	50.0	50.0	-	-	-	2
Annex VII	-	25.0	50.0	25.0	-	4
Annex VIII	7.1	14.3	78.6	-	-	13
Annex IX	-	-	8.3	-	91.7	12
PRODCOM	-	-	100.0	-	-	16

Elaboration from Admin Data ESSnet WP1, *Deliverable 1.2/2010. Database "Overview of Existing Practices in the Uses of Administrative Data for Producing Business Statistics in EU and EFTA"* (2011).

4. Some conclusions

Within the limits set for this presentation, it was possible to touch only a few, very general aspects of the use of administrative data in business statistics. This is a vast and complex field, refractory to any generalisation: by definition, the form and contents of administrative data depend on the national norms in accordance with which they are collected, and there is no guarantee that what works in one country can be successful replicated in another. Furthermore – and this can be seen as a first conclusion – higher standards of documentation are needed, which is paramount to an effective circulation of best practices: several NSIs produce poor or null documentation of the use of administrative data, and organisations such as Eurostat, the OECD or the UNSD, should do more to improve the quality of national metadata reporting in this area.

At the legislative level, all countries recognise the strategic value of administrative data for official statistics, and all laws enacted or amended in the last decade are aligned in establishing the principle of least statistical burden. On the implementation level, however, a lack of regulation on the obligations of data

holders is often reported by the NSIs as a serious point of weakness, and much remains to do to fully realise the law goals.

The potential of administrative data appears to be not yet fully exploited in several of the observed domains. This is especially true, within STS, for Turnover and Employment variables, and, as regards SBS, for the Annexes I-IV. In both cases, in fact, relevant administrative data (mostly from Tax sources and Social security) appears to be more widely available than they are actually used. Within STS, also the production of Building permits statistics seems to be mature to replace obsolete modes of data collection with more efficient register-based systems. In all of these cases, also based on the results of a specific query about why administrative data are not used, the next step for several countries could be to switch from using administrative sources as auxiliary sources (as sampling frames, for editing/ validation, for imputation of missing values, etc.) to using them as the main sources, i.e. in total or partial replacement of current surveys. To this goal – as taught by the longstanding experience of Nordic countries, but also by the recent innovations introduced in Portugal for SBS – investing in a closer cooperation with the Public administrations can be the key to improve the quality and availability of the existing sources and so reduce the statistical burden imposed on enterprises. Fewer opportunities for expansion of administrative data use appear to exist for SBS Annex VIII (Business services), and nearly none for STS-New orders, STS-Production prices/ costs, and Prodcom. On the other hand, in the domains of SBS Annexes V-VII (Insurance, Credit and Pension funds) and IX (Business demography), administrative data appear to be substantially used, except in a few cases, to the limit of current possibilities.

REFERENCES

Chumbau A., Pereira H.J., Rodrigues S. (2010), Simplified Business Information (IES): Impact of Admin Data in the production of Business Statistics, paper presented at the Admin Data ESSnet Seminar “Using administrative data in the production of business statistics. Member states experiences” (Rome, March 2010).

Consalvi M., Costanzo L., Filippini D. (2008), Evolution of Census Statistics on Enterprises in Italy 1996-2006: from the Traditional Census to a Register of Local Units, paper presented at the IAOS Conference “Reshaping Official Statistics” (Shanghai, October 2008).

Costanzo L. (2011), Use of Administrative Data and Use of Estimation Methods for Business Statistics in Europe: an Overview, paper presented at the Admin Data ESSnet Workshop “Using Admin Data - Estimation approaches” (Vilnius, May 2011).

CSO-Ireland (2010), Use of Administrative Data to compile Business Statistics – First implementation report, CSO.

INSEE (2008), The French Business Register: from a quality approach to a statistical register, paper presented at the 21st Meeting of the Wiesbaden Group on Business Registers - International Roundtable on Business Survey Frames (Paris, November 2008)

Orjala H. (2008), Potential of administrative data in business statistics. A special focus in improvements in short term statistics, paper presented at the IAOS Conference “Reshaping Official Statistics” (Shanghai, October 2008).

Rapiti F.M. et Al. (2010), What have we learned in almost 10-years experience in dealing with administrative data for short term employment and wages indicators?, paper presented at the Admin Data ESSnet Seminar “Using administrative data in the production of business statistics. Member states experiences” (Rome, March 2010).

Statistical Office of Slovenia (2009), Business Statistics With Strong Emphasis on Administrative Data Sources: Experience of Slovenia, paper presented at the UNECE Conference of European Statisticians, 57th plenary session (Geneva, June 2009).

Tolkki V. (2007), Finnish SBS system; use of administrative data, methods and process, paper presented at the INE Seminar “Reengineering Business Statistics” (Lisbon, October 2007).

Tønder J.-K. (2008), The Register-based Statistical System. Preconditions and Processes, paper presented at the IAOS Conference “Reshaping Official Statistics” (Shanghai, October 2008).

UNECE (2007), Register-based statistics in the Nordic countries, United Nations, Geneva.