

**26th Meeting of the Wiesbaden Group on Business Registers
- Neuchâtel, 24 – 27 September 2018**

Motseoa Molahlehi

Lesotho Bureau of Statistics

**Topic: The Administrative data as a source of updating Statistical
Business Register**

Abstract

The aim of the industrial statistics programme is to obtain comprehensive and accurate statistical information on the industrial activity within the economy. Unless the Statistical Business Register (SBR) is maintained on a regular basis, it quickly loses its value by becoming invalid and ceases to effectively produce true picture of the economic activity in the country.

This information may be obtained either through the statistical surveys or through institutional links with the data sets which can be obtained from different administrative sources. A commonly applied definition for “administrative data” in Europe is sources with information that is not primarily collected for statistical purposes and also another definition is raw data from administrative source, before any processing or validation by NSI.

Among three data sources (Administrative statistical sources including feedback from economic surveys, profiling and SBR improvement surveys, other surveys for example, data from private data suppliers, telephone directories and the internet, the guidelines recommend that SBR’s be created and maintained primarily using administrative data sources). This approach is in line with Principle 5 of the United Nations Fundamental Principles of official statistics, which states that “Data for statistical purposes may be drawn from all types of sources, be the statistical surveys or administrative records. The statistical agencies are to choose the source with regard to quality, timeliness, costs and burden on respondents. However, the best choice of data sources depends on the specific situation in

any given country, including the availability of administrative data and the scope and complexity of the National Statistical System itself.

Data on industrial activities are required for a variety of purposes by different users including the government, the business community, researchers and others. Countries with a developed statistical system always make more use of administrative sources for coverage of industrial activities. In a society that is more influenced by globalisation, International comparability of economic statistics is becoming increasingly important. This will be due to indicators which are defined on international basis on harmonized populations of statistical units. As a result, this leads to more comparable statistics.

In countries with less advanced statistical systems, the statistical business register will be incomplete because their micro- and small enterprises are not included in the register. It may not always be possible in practice, often for cost reasons, to cover all micro- and small units; therefore, some sort of cut-off is usually applied in practice. The proportion of the GDP covered by the units in the frame is often a more useful cut-off measure than the proportion of units covered.

A business register of good quality helps improve the efficiency of the national statistical system, which in turn helps reduce the response burden imposed on businesses. Therefore Policy Makers can easily access quality data for enhanced decision making.

Keywords: Industrial Statistics, Statistical Business Register, Administrative Data, Economic Activity

Topic: The Administrative data as a source of updating Statistical Business Register

Background

The Background of UNECE Statistical Business Register (SBR) over the last decade had growing demand for better and more detailed economic statistics and has put focus on SBR and their role in the production of economic statistics. It was discovered that SBR do not only serve as a sampling frame, but also as a source on which to derive economic statistics. SBR's have also proved to be central for statistical office's efforts to reduce response burden and utilise administrative data sources and also for the combination of survey data and administrative data.

The conference of European statisticians (CES) has a vision for the production of official statistics and raised a number of challenges for the development of the statistical production and services. The need for international guidance and recommendations on good practices on SBR's has been raised by countries participating in the Wiesbaden group on business registers. The proposal for international guidelines was supported by the steering committee of the Wiesbaden group in 2011. The final version for European guidelines was done in June, 2025

The background of the session paper is in relation to the status of Statistical Business Register (SBR) in Africa. There was a concern about quality of economic statistics in African countries. The consultant was engaged to assist in developing guidelines for construction of statistical business registers. The decision was made in 2011 to promote development of SBRs as part of Statistical Capacity Building Program being a component of the labour statistics. This was implemented by the mission of experts from SADC and COMESA countries which was held in South Africa in Pretoria. There was a publication of guidelines in 2014 after the meetings of experts and technical assistance. The Generic SBR project was initiated in 2014 after consultant's mission.

Concerns on the quality of economic statistics in Africa due to Lack of up-to-date business registers identified as the fundamental problem, Central registers mainly designed for employment purposes are maintained and updated using area enumeration and are very expensive to maintain and gets outdated. Units and divisions maintain their own registers on various machines and in different packages (Excel, IMPS, CSPro, etc.).

Status of Generic SBR in African countries show Mauritius is fully operational since the system was initiated first in the country. Zambia, Botswana and Swaziland have adapted generic SBR system. Other countries will adapt generic SBR system in the near future. Lesotho which is my country, fall under the countries which are still to develop the generic SBR system. Lesotho conducted economic census in 2012, in which census of establishments was the first stage which was the baseline for SBR and administrative data was used to update the SBR. In 2015 census of establishment was conducted again to update the SBR. The coverage of the country's SBR is incomplete due to gaps in both primary and administrative data.

Introduction

Industrial statistics is required for a many purposes by a variety of users including the government, the business community, researchers and others. Industrial statistics may be defined as a science that employs the techniques and methods of general statistical theory to develop a system of indexes that describe the composition and distribution of industries. (<https://encyclopedia2.thefreedictionary.com>)

This information may be obtained either through the statistical surveys or through institutional links with the data sets available elsewhere, in administrative sources. Generally, a mix of the two sources is used for the collection of industrial statistics. The extent of the use of one source rather than the other depends upon the statistical system of a particular country. Countries with a developed statistical system make progressively more use of administrative sources for coverage of industrial activities. Administrative

data refers to information collected primarily for administrative (not research) purposes. This type of data is collected by government departments and other organizations for the purposes of registration, transaction and record keeping, usually during the deliverance of service. (www.adls.ac.uk >guidance>introduction)

Administrative sources of information are progressively used by member states to compile and maintain statistical registers. Some are integrating the information held in the two types of register with the aim of producing a multipurpose register. The VAT register is one source of administrative information used by most statistical institutes while registers maintained by other taxation authorities, social security administrations and chambers of commerce are used by other statistical offices. Where these exchanges of information occur, the shape and content of the administrative and statistical registers can influence each other.

The Maintenance of SBR

The maintenance of statistical registers should not be regarded as an isolated operation but as part of a co-ordinated approach towards the joint development of statistical and administrative registers, although care must always be taken that the interests of a business will not be harmed by the transfer to other authorities of information it has given to the statistical institute. The document "Fundamental Principles of Official Statistics in the Region of the Economic Commission for Europe" adopted by the Member States of the Economic Commission for Europe contains the following two principles that are relevant in this context:

Data for statistical purposes may be drawn from all types of sources, be they statistical surveys or administrative records. Statistical agencies are to choose the source with regard to quality, timeliness, costs and the burden on respondents. Individual data collected by statistical agencies for statistical compilation, whether they refer to natural or legal persons, are to be strictly confidential and used exclusively for statistical purposes.

The list of all economic units in the survey target population is known as the sampling frame which is used for conducting sample surveys for data collection. The business register (BR), which is the sampling frame should include all accurate and up-to-date data items associated with units that are required for stratification, sample selection and contact purposes, for example, industrial, geographical and size codes, name, address and description of the unit, telephone number, and, preferably, a contact name.

The frame for a survey should contain all the units, without omission or duplication, that are in the survey target population and that contribute to the gross domestic product (GDP) of the national economy. It may not always be possible in practice, often for cost reasons, to cover all micro- and small units; therefore, some sort of cut-off is usually applied in practice. The proportion of the GDP covered by the units in the frame is often a more useful cut-off measure than the proportion of units covered. That is the sampling frame must contain information with a realistic GDP instead of many businesses which do not account for a required turnover.

The statistical business register is an essential tool for data collection and it is a register of enterprises or establishments engaged in the production of goods and/or services. The enterprises in the statistical business register have identifiable links to their establishments and are classified by economic activity. In countries with less advanced statistical systems, the statistical business register will be incomplete because their micro- and small enterprises are not included in the register, given the sheer number of enterprises in this segment of the total universe of enterprises. However most of African countries are implementing the system of using administrative data especially from the tax authority to build the sampling frame.

The statistical business register should be complemented by an area frame designed to cover the enterprises not included in the register to cover the industrial sector of the economy as a whole.

Purpose of the business register

The business register is an important statistical tool which provides not only the sampling frame needed for conducting the sample survey for the collection of data but also the basis for grossing up results from sample surveys so as to produce business population estimates.

A business register of good quality helps improve the efficiency of the national statistical system, which in turn helps reduce the response burden imposed on businesses. That is data on establishments can only be collected by National Statistics offices and the frame can be used by the country as a whole.

A business register can open up possibilities for electronic data interchange for statistical work, including transfer of data on a regular basis between national statistical offices, and business and other national organizations, especially in countries where the system of business register is advanced.

It is desirable, as the best option that the frame for every list-based enterprise survey undertaken as an inquiry on industrial activity is derived from a single general-purpose activity business register maintained by the statistical office, rather than that stand-alone registers be used for each individual survey.

Two basic reasons for using a single business register.

First, and most importantly, the business register operationalizes the selected model of statistical units and facilitates classification of units according to the agreed conceptual standards for all surveys. If survey frames are independently created and maintained, there is no means of guaranteeing that the surveys are properly coordinated with respect to the coverage that they provide.

Second, it is more efficient for a single organizational unit within the national statistical office to be responsible for frame maintenance than for each survey unit to create the frames for each of its surveys. For instance,

the economics division in National Statistical offices must be responsible for the frame of Business register since it deals with industrial statistics.

The roles of SBR

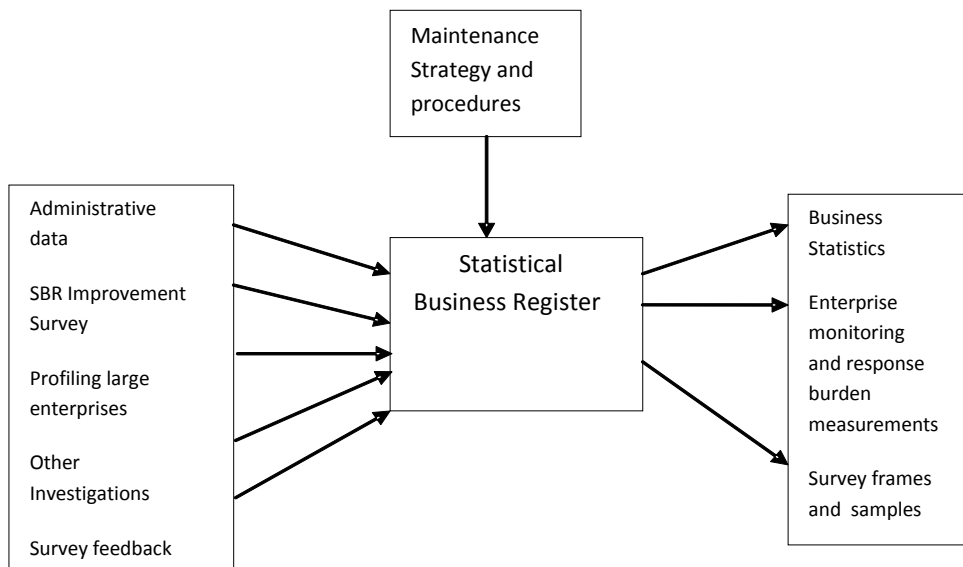
- ✚ SBR live register provide the gateway between data from various input sources and statistical units.
- ✚ SBR register snapshot and frozen frame provide populations of statistical (and possibly administrative) units at fixed points in time.
- ✚ SBR survey frame give a set of statistical units for a survey, valid for a specified reference period, with all characteristics required.
- ✚ SBR survey supports monitor survey response and measure and control response burden.
- ✚ SBR statistics produce statistics based directly on the SBR.
- ✚ SBR information source provide lists of enterprises and their locations and possibly other characteristics.
- ✚ SBR international data exchange assist coherence in international statistics.
- ✚ SBR in modernisation of statistical production and services advance integration of SBR within the production processes for economic statistics.

Summary of SBR components:

The SBR is fed by the inputs from the sources of data shown on the diagram and these are used for its update and maintenance. SBR maintenance is when changes have to be detected and the SBR has to be updated. The changes can be new businesses formed, change of production activities, existing businesses merge, or location, go bankrupt est.

The SBR then produce the outputs like business statistics, enterprise monitoring and response burden measurements and survey frames and samples.

Inputs and Outputs of SBR



Source: *Business Register Recommendation Manual*

Advantages of using administrative data

The advantages of using administrative data to supplement or replace survey data wherever possible are well known. They are particularly related to SBR construction and maintenance, as outlined in the following paragraphs.

Coverage: Use of administrative data in place of survey data reduce sampling error, remove or significantly reduces non-response and provides more accurate and detailed estimates for various sub-populations, e.g. small geographic areas. This is because administrative sources often give complete, or almost complete, coverage of a target population, whereas sample surveys often cover only a relatively small proportion directly. Coverage is of great significance from an SBR viewpoint, given its aim of including all economically active units. Thus, the main advantage of using an administrative source is the level of coverage it provides, which is either complete, or, if not complete, at least well defined, so that it can be assessed against the target population.

Costs: Another advantage administrative data offer over survey data is comparatively lower cost. Surveys are expensive, particularly if they are conducted as censuses or involve the use of personal interviews. Administrative data are often available free of charge, or for the marginal cost of extraction, particularly if they originate from the public sector. Even if there is a charge, it is often cheaper to use administrative data than collect the same information by survey. Fewer staff is usually needed to process the data and there is no need for non-response follow-up. The size and scope of an SBR makes it very difficult and expensive to populate and maintain solely by statistical data. A periodic economic census would be required to

build an SBR, which would be very expensive, as would be the complementary intercensal maintenance procedures.

Response burden: Using data from administrative sources involves no additional response burden. While businesses usually understand the reasons for supplying data for registration and taxation purposes, even if they do not like doing so, they may see statistical data requests as an extra, less necessary, burden. If they have already provided details to other government departments, they may become annoyed at receiving requests for similar information from the NSI. An associated advantage is that the use of administrative data may, in some cases, allow statistics to be produced more frequently, with no extra cost to businesses.

Content and timeliness:

The use of administrative sources may increase the quality of the SBR by providing access to more up to-date information on key characteristics, such as: name and address, births and deaths of units, the dates of these events, economic activity, code, location, size, in terms of number of employees and/or turnover.

As well as improving the timeliness of SBR data in the context of their use for survey frames, administrative data can improve the timeliness of statistics that are derived from the SBR. This is because surveys take time to plan, to design, to pilot questionnaires, to analyse the population and optimise the sample, to collect and process the data, etc. Access to a suitable administrative source via the SBR may provide a quicker and more efficient solution.

Disadvantages of using administrative data

Administrative data are collected for a specific administrative purpose, and the corresponding needs and priorities are likely to be different from those of the statistical system. Thus administrative units may not coincide with statistical units, and their characteristics may be differently defined. For example, the tax authorities may permit a single enterprise to have multiple value added tax (VAT) accounts. Turnover collected for VAT purposes may not include turnover related to the sales of VAT exempt goods and services, whereas the statistical system wants total turnover.

Similarly, the classification systems used within administrative sources may be different to those used by the NSI. For instance, a register within a food safety authority might classify a department store according to its food department. Even if the classification systems are the same, they may be applied differently, depending on the primary purpose of the administrative source. Extra, less necessary, burden. If they have already provided details to other government departments, they may become annoyed at receiving requests for similar information from the NSI.

Another common problem encountered when using data from administrative sources relates to timeliness. Administrative data may not be available in time to meet statistical needs, or may relate to a period which does not coincide with that required for statistical purposes. For income tax year may not coincide with the calendar year required for structural business statistics.

There is generally some sort of delay between an event happening in the real economic world and it being recorded by an administrative source. This is then followed by a further lag before the administrative data become available to the SBR. Lags in recording births and deaths

of enterprises are a major source of SBR coverage errors. If these gaps can be measured, allowance can be made for them in statistics based on SBR data.

Changes to administrative regulations or procedures

Public sector administrative sources are set up for the purposes of administering regulations, for example collecting taxes or for monitoring government policies. This means that they are susceptible to political changes. If a policy changes, administrative sources may be affected in terms of coverage, definitions, thresholds, etc. They may even be abolished completely. Such changes may happen quite suddenly, with little warning. Thus, reliance on administrative sources always carries a certain degree of risk in terms of continuity of supply.

These risks can be mitigated by legal or contractual provisions, by obtaining early warning of impending changes through regular contact with those responsible for the sources, and by drawing up (and implementing where necessary) contingency plans.

Monitoring the quality of administrative data create knowledge of the administrative sources

Although administrative data have many advantages, it is important to invest time in understanding and monitoring their quality. It can be useful to compare sources in terms of coverage of their units and accuracy of their characteristics. The closeness of administrative units and characteristics to what are required for an SBR is an important factor in determining the quality of an administrative source. An SBR improvement survey may be required to determine the values of certain characteristics for which values provided by administrative sources are not considered appropriate.

The starting point in assessing the quality of an administrative source is to build up a thorough knowledge of the source, including its primary purpose

and the way the data are collected and processed. Thorough understanding of a source allows a more accurate assessment of its strengths and weaknesses. To help develop and document this knowledge, particularly for the benefit of future SBR staff, it is useful to develop some form of template to record information from the source on contacts, units, characteristics, quality and formats.

Quality indicators examples to be reviewed regularly and be discussed with the administrative authorities are:

- ✚ The number and proportion of enterprises lacking a valid and complete economic activity code.
- ✚ The number and proportion of enterprises for which the activity status (active, dormant, dead, etc.) is unknown.
- ✚ The number and proportion of enterprises lacking a complete address.

Feedback from surveys based on the SBR and/or from SBR improvement surveys.

In the case where an administrative source and the SBR do not agree, the reasons should be investigated with the aim of gaining a better understanding of SBR quality. Surveys may be used to investigate such discrepancies. They may be conducted, either specifically for this purpose or as part of some other data collection exercise. The investigations should help in determining the appropriate rules and priorities for updating from different sources.

There may also be conflict in data from different administrative sources, for example conflicting industry codes. Procedures and rules need to be

developed to resolve these problems. This could involve verifying the data by contacting the enterprise, or undertaking analytical work to determine which source is most reliable. The goal is a set of general rules to deal with conflicts.

Concerning metadata on industrial statistics, it provides a mechanism for comparing national practices in the compilation process. This may help and encourage countries to implement international standards and to adopt best practices in the compilation of particular statistics.

Metadata

The term metadata defines all information used to describe other data. A very short definition of metadata, then, is “data about data”. Metadata descriptions go beyond the pure form and content of data and to encompass administrative facts about data (who has created them and when) and how data were collected and processed before they were disseminated or stored in a database. In addition, metadata facilitate an efficient search for and location of data.

Metadata is also used for better harmonization of approaches adopted by different countries which will improve general quality and coverage of key statistical indicators.

The most fundamental purpose of metadata is to help the users of industrial statistics understand, analyse and interpret the data, even if they have not themselves participated in the process of the production of those data.

They facilitate sharing, querying and understanding of statistical data over the lifetime of the data.

They also refer to any methodological descriptions on how data are collected and manipulated.

For industrial statistics data items, for example, metadata include the name of the data item, the unit from which the information has been collected, data sources, information about classifications used and series breaks, definitions and methodologies used in their compilation.

Metadata are essential for the interpretation of statistical data.

The metadata of the disseminated industrial statistics should encompass the following six main components:

- (a) Data coverage, periodicity and timeliness;
- (b) Access by the public;
- (c) Integrity of disseminated data;
- (d) Data quality;
- (e) *Summary* methodology; and
- (f) Dissemination formats.

Countries are encouraged to accord the development of metadata a high priority and to consider their dissemination an integral part of the dissemination of industrial statistics.

Recommendations

Administrative sources must agree on common variables which are statistical with NSO so that SBR can cover the missing variables. This can be done through Memorandum of Understanding's.

Just like when surveys/census are conducted, users/stakeholders must always make inputs in the statistical tool.

The NSO can even deploy some staff to help in the capturing of those crucial variables.

Conclusion

The administration data seems to be the best source of SBR, especially VAT data though there are some limitations in that data source. Countries must form a memorandum of understanding between their NSO and administration data, especially VAT data.

The generic SBR system has proven successful in inducing the faster adoption of concepts, structures and tools which can improve the quality of economic statistics in Africa.